L2 revision and post-task anticipation during text-based synchronous computer-mediated communication (SCMC) tasks

Vararin Charoenchaikorn

This thesis is submitted for the degree of
Doctor of Philosophy in Linguistics

May 2019

Department of Linguistics and English Language

This thesis is dedicated to my parents,
with love and gratitude.

Declaration

This thesis has not been submitted in support of an application for another degree at this or any other university. It is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated. Many of the ideas in this thesis were the product of discussion with my supervisor, Prof. Judit Kormos.

Vararin Charoenchaikorn
Lancaster University, UK

Abstract

The current research investigates L2 revision during text-based synchronous computer-mediated communication (SCMC) and its relationships with the accuracy of text in chat logs and typing ability. Another main aim of this study is to explore how tasks can be implemented to facilitate learning in this medium. In particular, the effects of post-task anticipation (± post-task anticipation) and its type (anticipation of an individual vs a collaborative language correction post-task) on learners' main task performance in terms of revision, speed fluency and accuracy are examined.

The study is primarily motivated by the methodological shortcomings of previous research exploring L2 changes during text-based SCMC, and the scarcity and limited scope of investigation of post-task anticipation studies. Various data collection methods were utilized to gain rich research data. Performance data were obtained from computer screen recordings, keystroke logs and chat logs by means of two text-based SCMC tasks involving picture description and decision-making. Stimulated recall interviews were carried out to gauge participants' thoughts during revision in order to ensure the reliability of the coding of revisions; and an exit questionnaire and a follow-up interview were administered to elicit their responses pertaining to different aspects of the research including their experience of post-task anticipation. This study manipulated both between- (± post-task anticipation) and within-participant (two types of post-task anticipation; anticipation of an individual and a collaborative language correction post-task) factors. Eighty-four Thai learners of English were randomly assigned to either a control (N = 28) or an experimental (N = 56) condition. While the control group carried out two main tasks without any post-task anticipation, the experimental group was informed about a post-task before each of the main tasks.

Keystroke logs were examined for linguistic errors and evidence of revisions made during drafting or to the already-sent text. Revisions were coded based on criteria adapted from the revision taxonomies of previous writing research and aided by the data from computer screen recordings and stimulated recall interviews. The variables investigated included quantity, linguistic units, focus and triggers of revision, and rates of error revision success and error corrections. Accuracy was gauged in terms of both accuracy during

writing and final text accuracy. Speed fluency was assessed by process-based measures, and typing ability was operationalized as typing speed adjusted for typing accuracy during a typing test. Qualitative data from the exit questionnaire and follow-up interviews were used in conjunction with quantitative data during the analysis.

The results showed a high total revision frequency and a high rate of error revision success, suggesting that learners paid close attention to their L2 output and could successfully draw on their L2 knowledge to improve form-related errors in this medium. There was evidence that participants attended more to grammatical features than lexical ones, noticed and corrected more grammatical mistakes compared to lexical ones, and tended to correct grammatical errors more successfully than lexical ones. However, although students attended to grammatical items and revised frequently, the observed dominance of content revisions over form-related revisions indicated that their attention was primarily devoted to the meaning-related aspects of language, rather than to form. This finding does not support previous claims regarding the benefit of text-based SCMC, which argue that this medium is suitable for promoting learners' attention to form. In addition, local revisions occurred very frequently, suggesting that learners' attention might be restricted to short stretches of text at the letter, word or phrase level. Regarding the relationship between revision and final text accuracy, error correction rates were found to be the best predictors of final text accuracy out of all the revision measures. The results of follow-up analyses showed that proficiency potentially influenced final text accuracy and error correction rates; higher proficiency was significantly correlated with increased error correction rates and final text accuracy. Although the correlations between typing ability and most L2 revision measures were not significant, significant relationships were observed between typing ability and 1) error correction rates and 2) accuracy. These findings indicate that learners with better L2 typing ability may have more attentional resources available for attending to L2 output, resulting in increased detection and correction of their linguistic errors and increased internal L2 monitoring.

As far as post-task anticipation is concerned, the findings do not support Skehan's (1998) hypothesis about the potential of post-task anticipation for enhancing attention to form and accuracy during the main task performance. No significant effect of post-task

anticipation was detected on revision, accuracy or speed fluency. The non-significant effect found on fluency is consistent with the findings of previous post-task anticipation research which did not detect a clear influence of post-task anticipation on this performance aspect. In addition, no significant effect of type of post-task anticipation was observed.

Acknowledgements

First and foremost, I would like to express my thanks to my supervisor, Prof. Judit Kormos, for her unwavering support, patience and expert guidance during my PhD. Also, I would like to acknowledge the insightful feedback and advice given to me by Asst. Prof. Marije Michel.

This study would not have been possible without the cooperation of the participants and so I thank Karuna Naphon, Kietnawin Sridhanyarat, Dr Kriangkrai Vathanalaoha and Munchuree Kaosayapandhu for their kind assistance in recruiting them.

Next, I would like to thank all past and present Thai research students and PhD students in the Department of Linguistics and English Language for their friendship and encouragement throughout my time at Lancaster University. Specifically, I am grateful to Dr Samuel Earp for his understanding, support and consideration.

Lastly, I would like to express my deepest gratitude to my parents, Visit Charoenchaikorn and Phenchun Charoenchaikorn, and my sisters, Jeerapat Charoenchaikorn and Varitta Charoenchaikorn, for all the love and support they have given me throughout my life.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Background and motivation for the research

Today, the Internet has become an essential part of humans' life across most of the globe. New generations are born *digital natives* (Prensky, 2001) and their life, which is interconnected with technology, shapes their literacies and cognitive and learning processes (Thorne, 2013). Thus, it is not surprising that technology plays an important role in learning and teaching, including in the area of L2 where technology-assisted and online pedagogic tasks are widely used by educators.

There are several benefits of performing L2 tasks online, the most obvious of which is the increased opportunities for learners to learn and practise the target language outside the classroom. Many challenges caused by the physical and temporal constraints of traditional classroom settings can also be overcome in the online context. These challenges include, for instance, learners' overreliance on the teacher and crowded or mixed-proficiency classrooms which are difficult to manage (Lai & Li, 2011). According to Lai and Li (2011), the use of technology for online L2 teaching creates a sense of authenticity that can promote learners' motivation and expands the range of tasks that can be implemented due to the abundance of online resources.

Despite the potential of teaching and learning L2 through online tasks, many researchers have voiced their concerns over the amount of L2 research on the use of technology in language instruction (Motteram & Thomas, 2010; Ortega & González-Lloret, 2015; Thomas & Peterson, 2014; Thomas & Reinders, 2010). Thomas and Peterson (2014) observe that there is an "evangelical tone" (p. i) in the discussion surrounding educational technology, even though, in reality, the issue remains under-researched. In line with this, Ortega and González-Lloret (2015) warn educators to be careful about the "euphoric assertions" (p. 80) of the potential of technology and the "idyllic images of effortless, black-and-white benefits" (p. 80) portrayed regarding the role of technology in L2 education.

In addition, many researchers have highlighted the necessity for the reconceptualization of task design and implementation so that learning opportunities are maximized through the unique affordances of technology (e.g. Adams & Nik, 2014; Baralt, 2010; Chapelle, 1998, 2001; Doughty & Long, 2003; Skehan, 2003a). For instance, Chapelle (2001) points to the importance of finding tasks that are suitable for the nature of technology-assisted L2 learning in the excerpt below:

> ...anyone concerned with second language teaching and learning in the 21st century needs to grasp the nature of the unique technology-mediated tasks learners can engage in for language acquisition and how such tasks can be used for assessment ... 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. (p. 2)

Unlike performing tasks in a face-to-face context, performing online tasks involves not only linguistic performance, but also a complex system of communicative competency, digital literacy and intercultural understanding (Lai & Li, 2011). Thus, it is important for language educators and researchers to understand the nature of L2 learning through online tasks, which is different from that of L2 learning in traditional classrooms. It may be that without appropriate implementation of the right tasks the use of technology-assisted L2 tasks alone may not yield many benefits or solve L2 pedagogic problems.

The present study is motivated by the call for research on technology-mediated L2 instruction and the importance of the reconceptualization of task design and its implementation in online L2 learning discussed above. In this research, I aim to investigate the potential of L2 learning through text-based synchronous computer-mediated communication (SCMC), or computer-based text chat, and how tasks can be effectively implemented in this medium. The focus of my research will be presented in greater detail in the next section.

1.2 Research focus

The focus of this study is on online computer-based text chat, a type of computer-mediated communication (CMC). Computer-mediated communication refers to human-human interaction taking place via the instrumentality of computers (Herring, 1996). It can take various forms, such as text-based, voice-based, video-based or a combination of these. With its unique characteristics, CMC is considered a revolution in knowledge production and a modern mode of communication (Warschauer, 1997). CMC offers a virtual space for language learners to communicate and collaborate with one another or with native speakers of the target language, which has been assumed to foster language learning (Blake, 2007; Chapelle, 2004; Kern & Warschauer, 2000).

To date, CMC has been widely researched, especially in the domain of text chat. The central claim regarding the benefit of text chat revolves around how it can draw learners' attention to language form, a process that has been hypothesized to promote L2 development. However, the results of previous L2 text chat research have yielded unclear findings concerning this benefit. Hence, in this research, my main aim is to answer the question of whether computer-based text chat can serve as a medium that encourages learners' attention to form. My study explores learners' revision behaviours, which are indicators of learners' attention to their own language output, during L2 task performance in this medium. Another main goal of this study is to explore how tasks might be implemented to raise learners' attention to form during L2 computer-based text chat. In particular, it investigates the potential for promoting learners' attention to language form through post-task anticipation. It has been hypothesized that knowledge of post-tasks might encourage learners to be more accurate and avoid language errors during the main task (Skehan, 1998). However, research investigating the effects of post-task anticipation is scarce and limited to the context of oral production tasks.

The research questions of this study are listed below. The first two research questions investigate 1) L2 revision behaviours during computer-based text chat and 2) the effects of post-task anticipation and type of post-task anticipation on revision behaviours, speed fluency and accuracy. The other questions explore the relationships between L2 revision behaviours during text chat and 1) final text accuracy and 2) learners' L2 typing ability.

Essentially, this research was conducted within a foreign language (FL) learning context. However, throughout this thesis, the term second language (L2) is used synonymously with foreign language.

RQ1 What are the L2 revision behaviours of Thai learners of English with A2 to B2 proficiency levels during text chat?

RQ2 What are the effects of post-task anticipation (± post-task anticipation) and type of post-task anticipation (anticipation of an individual vs a collaborative language correction post-task) on revision behaviours, speed fluency and accuracy?

RQ3 How are L2 revision behaviours related to text accuracy in chat scripts?

RQ4 How is learners' typing ability related to L2 revision in text chat?

My research uses both quantitative and qualitative data collected using various methods. Performance data were gathered from the participants' screen recordings, keystroke logs and chat scripts. They were collected by means of two computer-based text chat tasks involving picture description and decision-making. Stimulated recall interviews cued by replaying screen recordings of learners' performance were also conducted to explore participants' thoughts during revision. This type of data was collected in an attempt to ensure the reliability of the coding of revisions. To gauge typing ability, a factor which may influence revision, a typing test was administered. Finally, an exit questionnaire provided additional insights into participants' perspectives.

In essence, the current research is situated in the domain of technology-assisted task instruction. The work informs how task medium and implementation methodology might facilitate language learning. However, despite focusing on the use of tasks for language learning and teaching, this study is not specifically related to *task-based* language teaching (TBLT). This term refers to an approach that assumes the importance of communicative tasks, adopting them as the central unit for curriculum design, classroom activities and assessment. It is distinct from *task-supported* instruction, in which tasks are only utilized to complement a form-focused syllabus, offering learners the opportunity to practise target language features (Samuda & Bygate, 2008) (see the review in Chapter 4, Section 4.2).

This research takes a more general perspective by exploring how tasks could be used in any type of instruction, either task-based or task-supported.

1.3 Overview of the thesis

This thesis comprises eight chapters. Chapter 2, which follows this introductory chapter, conducts a review of writing and revision processes and empirical studies carried out to explore L2 revision. Chapter 3 focuses on literature relevant to computer-based text chat, or text-based synchronous computer-mediated communication. It offers an overview of the characteristics of this medium and discusses how learners monitor their language when communicating in computer-based text chat. In addition, it describes how the medium may facilitate L2 learning by focusing on its potential for enhancing students' attention to form. Reviews of empirical studies exploring attention to form and L2 revision in this mode are also provided in this chapter. Chapter 4 of this thesis discusses the definition of task in SLA, the utilization of tasks in L2 pedagogy, the importance of attention to form when teaching and learning through L2 tasks, and how educators can promote attention to form when using tasks for L2 instruction. Moreover, task implementation methodologies designed to foster attention to form in different task stages and post-task anticipation research are reviewed in this chapter. Following the review of literature in Chapters 2–4, Chapter 5 presents the methodology of the present research. It describes the rationale of the study, the operationalization of revision and post-task anticipation, my hypotheses, the research design, participants and research instruments, and the data collection and analysis procedures. Chapter 6 reports the results of the study, and Chapter 7 subsequently discusses the findings related to each research question. Finally, Chapter 8 concludes the thesis by summarizing the findings, discussing the contributions and implications of this research, outlining the limitations of this study, and making suggestions for future research.

Chapter 2: L2 revision

Revision plays an important role in writing processes. It entails evaluating and making changes during writing (Piolat, 1997). Initially, revision was considered to be a "tidying-up activity" (Faigley & Witte, 1981, p. 400) aimed at correcting surface errors such as lexical, grammatical and spelling mistakes in produced text. In this view, revision is the final stage of writing processes which comprise prewriting, writing and rewriting stages (Faigley & Witte, 1981). Bartlett (1982), for instance, distinguishes revision from text generation by explaining that, unlike text production, revision involves comparing some parts of a text to the writer's knowledge or intention, a process that then leads to making changes to the "existing text". However, this definition of revision is challenged in more recent research. More recent works do not consider revision as the final stage of text production but define it as any change made at any point during the course of writing (e.g. Chenoweth & Hayes, 2001, 2003; Fitzgerald, 1987; Hayes, 1996, 2012). In this view, revision can occur before the writer finishes writing a segment of text, e.g. a word, phrase or sentence, or even before the text is actually produced. In this research, this definition of revision is adopted.

Due to its importance in the writing process, revision has been included in just about all cognitive models of writing. These models are constructed to explain the psycholinguistic processes involved in L1 writing. To the best of my knowledge, to date, no cognitive models that are specific to L2 writing processes have been devised, and most L2 writing and revision studies have explained L2 writing processes by drawing on the processes included in L1 writing models. The section below will discuss writing and revision processes by outlining an early proposal for L1 revision processes (Bartlett, 1982) and influential cognitive models of writing and revision (Bereiter & Scardamalia, 1987; Chenoweth & Hayes, 2001; Galbraith, 1999, 2009a, 2009b; Hayes, 1996, 2012; Hayes & Flower, 1980; Hayes, Flower, Schriver, Stratman, & Carey, 1987; Kellogg, 1996). At the end of this chapter, empirical studies carried out to explore L2 revision will be reviewed.

2.1 Writing and revision processes

2.1.1 Bartlett's (1982) proposal for revision processes

In 1982, Bartlett proposed different components of L1 revision processes. She explains that revision comprises three processes: detection, identification and correction. The first process, detection, refers to the process in which writers make a conscious comparison between an existing text and their own goals, intentions and knowledge, e.g. knowledge of writing or genre conventions. Bartlett believes that writers must detach themselves from their text when rereading it in order to be able to detect any discrepancies. The second component involves the identification of problems. According to Bartlett, although detection and identification usually occur together, there may be times when writers sense that something is wrong in the text they write but are unable to identify why it seems wrong to them. She points out that identification depends on the writer's conceptualization of their own knowledge, writing strategies, goals and skills, which informs them of what they should do and of the nature of problems that may occur. Finally, the last process, which follows detection and identification, is correction. The success of this process, according to Bartlett, is based on how complete the previous two processes are. If a writer is able to detect problems and identify them properly, he/she may be able to engage in appropriate corrections. Revision strategies may depend on various factors including the writer's linguistic knowledge and stylistic judgements. The writer's choice of revision strategy is assumed to reflect their knowledge of the syntactic and semantic properties of particular linguistic features and knowledge of the context in which they occur. Furthermore, Bartlett hypothesizes that stylistic judgements play an important role in determining the choices of experienced writers, while novice writers rely on strategies they are familiar with, regardless of their appropriateness in some contexts.

Overall, Bartlett's proposal is an early attempt to explain revision processes. Reflected in this proposal is Bartlett's adoption of the early definition of revision explained in the previous section. In other words, Bartlett makes a clear distinction between revision and text production, with the latter preceding the former. In the sections that follow, influential

cognitive models that outline the psycholinguistic components of writing and revision processes will be reviewed. These models consist of more complex processes and assume less clear-cut stages between revision and text production.

2.1.2 Hayes' and Hayes and colleagues' cognitive models of writing and revision

Hayes (1996, 2012) and Hayes and colleagues (Chenoweth & Hayes, 2001; Hayes & Flower, 1980; Hayes, Flower, Schriver, Stratman, & Carey, 1987) have proposed several cognitive models of L1 writing and a model that specifically explains the processes involved in revision. These models outline several processes involved in writing and revision and the factors that may influence them. Over the years, the initial writing model (Hayes & Flower, 1980) has evolved to become more complex, consisting of many processes and subprocesses. Based on the findings of empirical studies, the original model has undergone several reorganizations and reconceptualizations of constructs, which will be explained in this section.

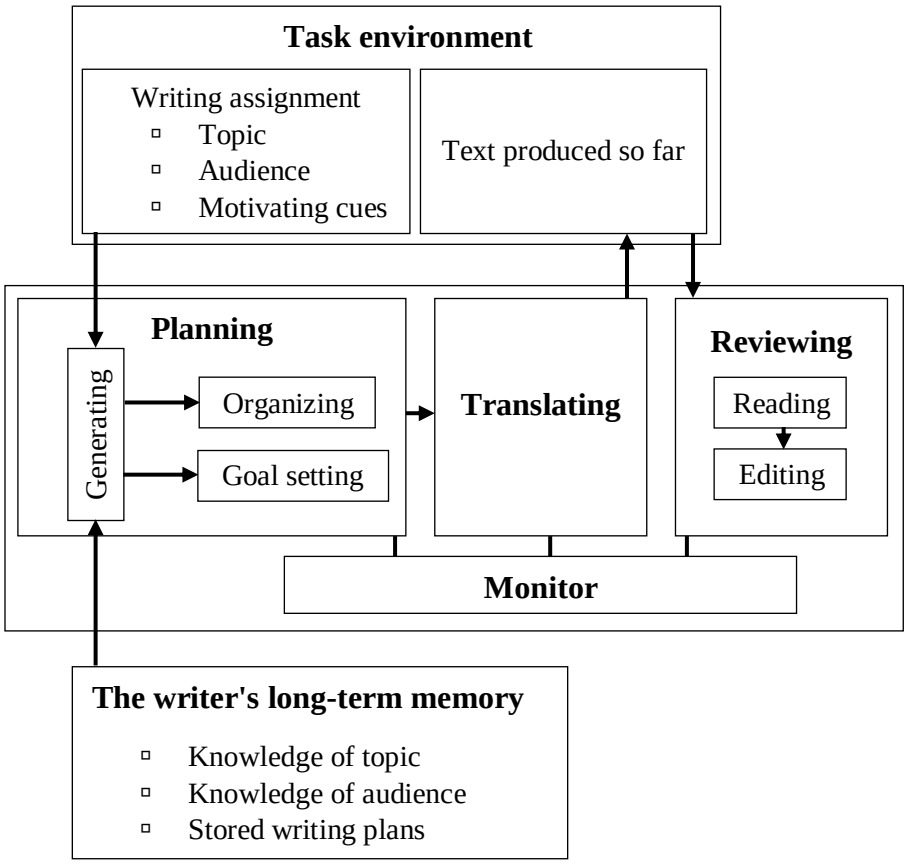
2.1.2.1 Hayes and Flower's (1980) model of writing

Hayes and Flower's (1980) original model of writing is a three-component model comprised of task environment, long-term memory and writing process (Fig. 1). The first two components are the context and writing resource, while the last explains writing processes.

Task environment is defined as "everything outside the writer's skin that influences the performance of the task" (Hayes & Flower, 1980, p. 12). It includes elements of the writing assignment, such as the topic of writing and the target audience, and the text produced so far. The second component, long-term memory, includes the writer's knowledge about the topic and the audience and their general writing plan, and it is hypothesized to interact with all writing processes during the course of writing. The writing process, the last component, is made up of three major processes, namely planning, translating and reviewing. Planning entails the use of information from the task environment and the writer's long-term memory

to set goals for writing and creating a writing plan to achieve them. Translating involves carrying out the writing plan generated in the planning process. Reviewing is the process which has the function of improving the quality of the text produced during the translation process. Its subprocesses are reading and editing. During editing, writers evaluate whether their writing goals are met, detect problems such as linguistic inaccuracies and unintended or unclear meanings, and make changes to improve the text. Importantly, Hayes and Flower (1980) make a distinction between editing and reviewing. While reviewing is a conscious attempt to improve the text after the translation process has ended, editing occurs automatically at any point during writing and briefly interrupts other ongoing processes. According to this model, writing processes are recursive in nature. Hayes and Flower emphasize that although the three writing processes generally occur in a linear fashion, all of them may reoccur at any moment when editing takes place.

Figure 1: Hayes and Flower's model of writing (1980, p. 11)



In addition to the three main cognitive processes, this model includes another process – the monitor. Although the illustration (Fig. 1) of the model presented in Hayes and Flower's (1980) publication might suggest that the monitor controls the performance of other processes, the monitor was not designed to be the master process (Hayes, 2012). Instead, its function is to account for individual differences among writers. According to Hayes (2012), some writers may prefer to start writing after they have finished planning, while others tend to plan while writing. Taking this into consideration, Hayes and Flower (1980) include the monitor in the model to represent each writer's predisposition to sequence the processes of writing in a certain way (Hayes, 2012).

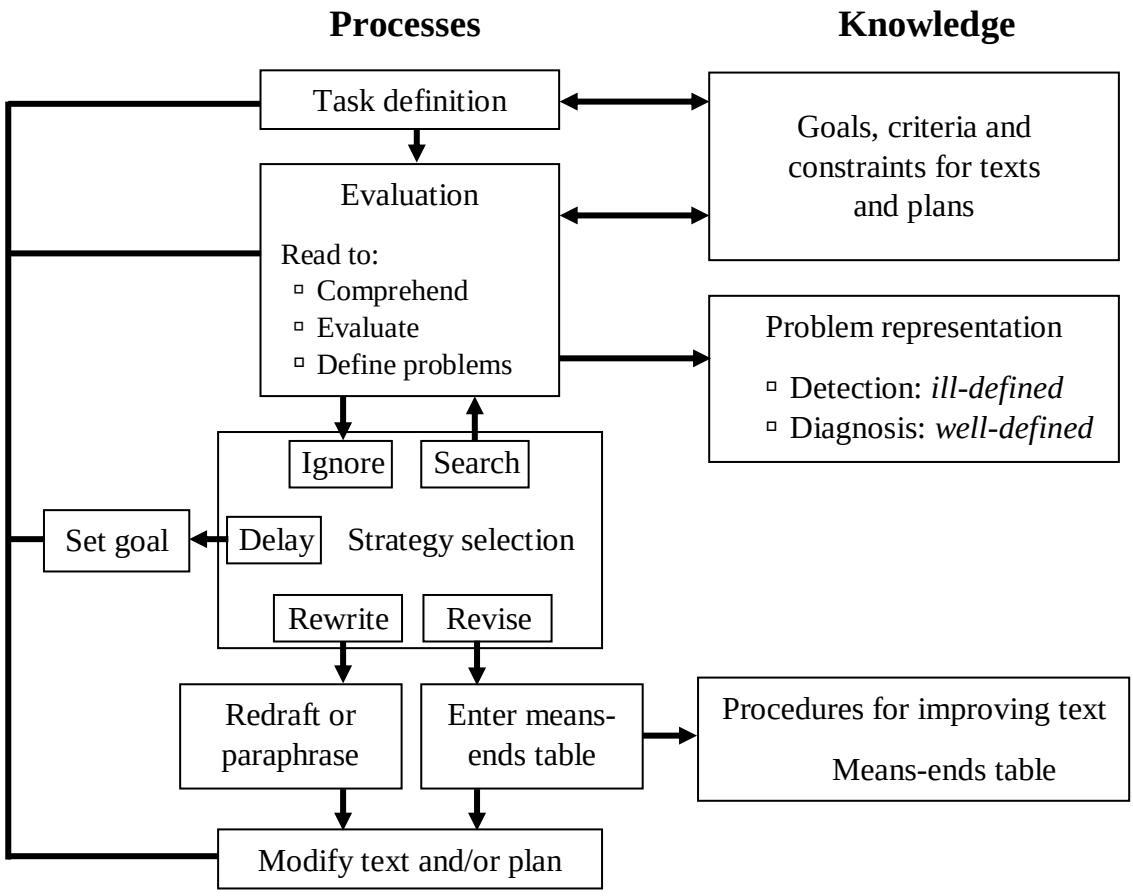
2.1.2.2 Hayes, Flower, Schriver, Stratman, and Carey's (1987) model of revision processes

Following Hayes and Flower's (1980) proposal of a model of writing, Hayes et al. (1987) devised a model for cognitive processes of revision (Fig. 2) based on the findings of empirical studies. This model provides a detailed description of the reviewing process in the earlier model (Hayes & Flower, 1980). In this model, revision is viewed as a complex process consisting of many subprocesses: evaluation, strategy selection and modification of text. The model is divided into two main parts – processes and knowledge – which explain the processes of revision and the types of knowledge associated with each process.

In this model, revision processes are guided by a plan in one's long-term memory called *task definition*, which specifies the goal of revisions and how revision should be carried out. Task definition can vary across individuals depending on their writing skills, and within an individual across different situations. The major revision process in this model is evaluation, which is guided by information relating to task definition. During evaluation, the writer reads to comprehend the text, and detects and characterizes problems. This process results in problem representation, which may range from very ill-defined to very well-defined problems. Problems that are ill-defined are those just detected, while well-defined problems are detected problems that have been properly diagnosed. Based on the definition of task and how well-defined the problem is, the writer chooses revision

strategies that can involve ignoring or delaying fixing problems, searching for more information regarding the nature of problems, or modifying the text. Text modification may range from rewriting to revising based on how much the surface structure of the original text is preserved. Rewriting refers to discarding the whole surface structure of the text to rewrite it in a different way. Revising, in a restricted sense, means fixing specific problems in the text with the fewest changes possible. According to Hayes et al. (1987), during revision processes, there may be variations in the individual's writing goals and criteria for revision, the types of problems each writer can identify, and the sophistication of the revision procedures they can carry out. All of these influence the processes of revision and determine the target of revisions.

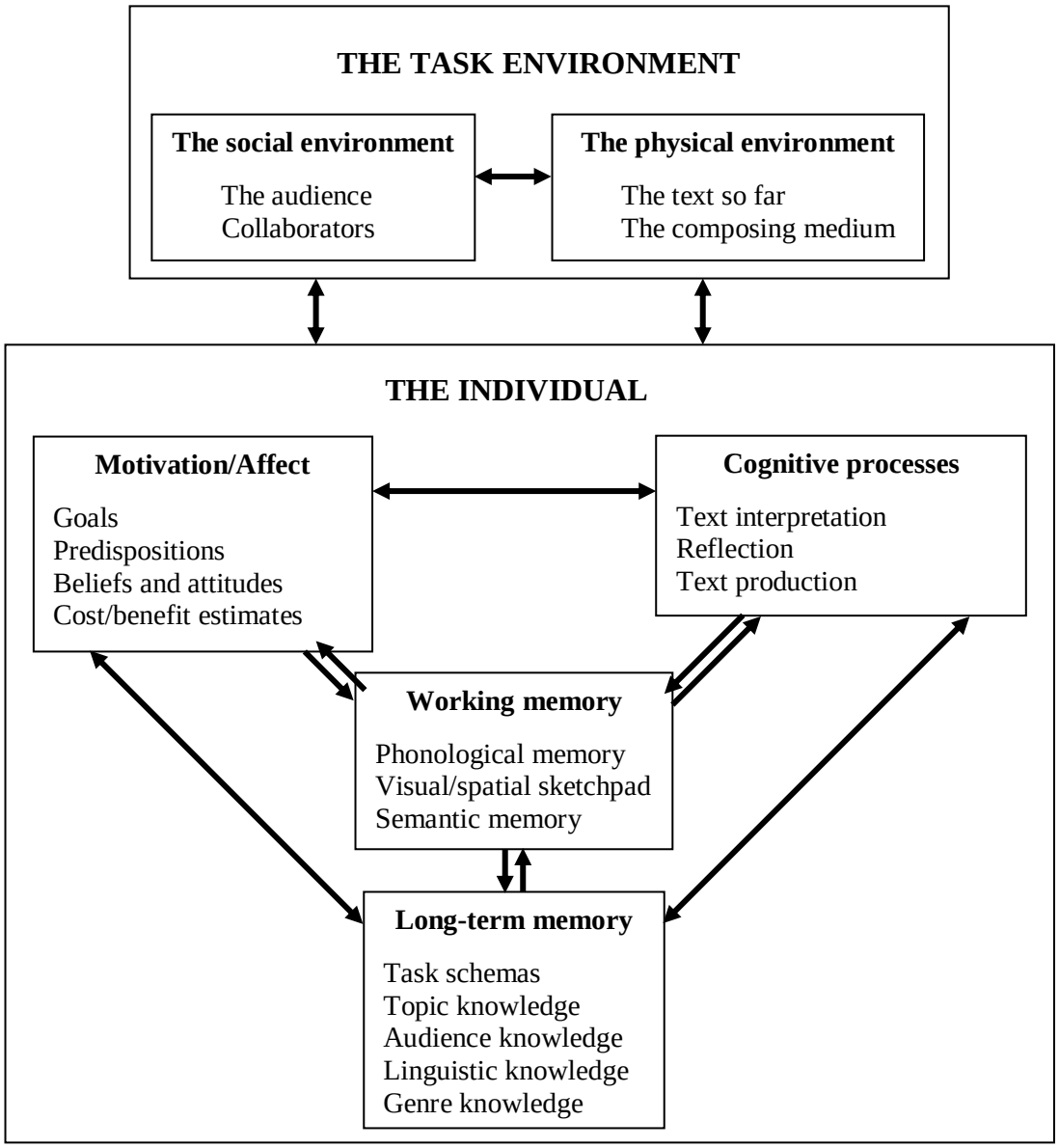
Figure 2: Hayes, Flower, Schriver, Stratman, and Carey's model of revision processes (1987, p. 185)



2.1.2.3 Hayes' (1996) model of writing

In 1996, Hayes proposed a revised version (Fig. 3) of the original model of cognitive processes of writing (Hayes & Flower, 1980). The new model is made up of two main components – the task environment and the individual – and is more elaborate, consisting of more elements than the original.

Figure 3: Hayes' model of writing (1996, p. 4)



There are three major differences that distinguish this model from the original version. First, the components of the task environment are reorganized into two categories: the social (the audience and collaborators) and the physical (the text produced so far and the composing medium) environment. The composing medium is a new element added to the original version and is assumed to influence the processes of writing, such as revising and planning. It is hypothesized that different kinds of composing media could offer different types of support for writing and involve different writing processes. For instance, it may be easier to check spelling mistakes and move chunks of text around when composing with a word processor than with pen and paper (Hayes, 1996). The language review process involved in pen-and-paper writing and during composing via a dictating machine are also different, as the former requires reading while the latter entails rewinding and replaying the parts that the writer wants to review (Hayes, 1996).

Second, apart from the composing medium, Hayes also added the influence of the writer's working memory and motivation and affective factors (e.g. the writer's goals, predispositions and beliefs) on writing processes in the new model. The model assumes that all writing processes have access to the cognitive resources in working memory and that these are used when carrying out non-automated activities. In terms of motivation and affect, Hayes explains that affective factors such as self-perceived writing competence may influence writing, and motivation affects the writer's response to their goals and the choices they make regarding the use of strategies to achieve those goals.

Third, the three cognitive processes of writing in the original model are reconceptualized in the new one. Planning is replaced by a more general process, reflection, which includes various subprocesses: problem-solving, inferencing and decision-making. The translation process is replaced by text production, a more general process responsible for generating oral, textual and graphic output. Finally, revision is replaced by text interpretation. According to Hayes, text interpretation has the function of generating internal representations relying on the graphic and linguistic input obtained from, for instance, reading and graphic scanning. This change is made based on the notion presented in Hayes

et al.'s (1987) revision model, which views the process of critical reading – reading to detect and diagnose problems in the text – as part of revision processes.

According to Hayes (1996, p. 15), "Recognizing that the revision model included reading as a subpart suggested that revision would more naturally be thought of as a composite of more basic processes". In this publication, Hayes (1996) conceptualizes revision as a composite of three main writing processes: reflection, text production and text interpretation, and provides an explanation of the factors influencing revision processes or what he calls the *control structure*. The control structure of revision is similar to the task definition presented in the model of revision (Hayes et al., 1987). It is a *task schema*, or a set of knowledge acquired through practice, which includes, for instance, knowledge about the actions needed to be taken when revising, criteria for assessing text quality, and the goal of revision.

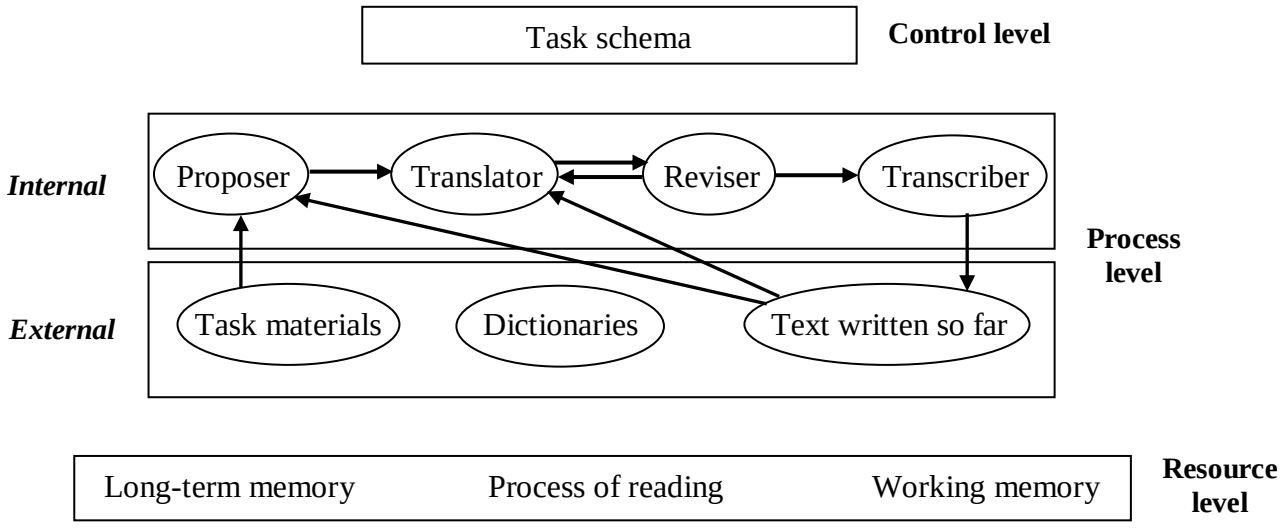
2.1.2.4 Chenoweth and Hayes' (2001) model of writing

Following Hayes' (1996) model, in 2001, Chenoweth and Hayes proposed a three-level model of writing (Fig. 4) containing a control, a resource and a process level. Similar to the control structure of revision presented in Hayes' (1996) previous work, the control level of writing in this model is the task schema, a package of information important for determining the interactions between writing processes. The resource level consists of long-term memory, working memory and the process of reading, all of which are resources for cognitive writing processes. The process of reading is added to this model as writers have to read the text they produce or source materials to check for problems or generate new content.

The main contribution of this model lies in the description of the process level. The process level is divided into two parts: internal processes and the external environment. The external environment includes elements described as the social and the physical environment in Hayes' (1996) previous model, e.g. the audience and the text produced so far. While there are three writing processes in the previous models (Hayes, 1996; Hayes &

Flower, 1980), this new model assumes four types of processors: the proposer, the translator, the reviser and the transcriber. The proposer is responsible for generating ideas and conceptualizing prelinguistic input for writing. The translator turns ideas into language through linguistic processing and stores its output in working memory for further evaluation by the reviser. If the language generated by the translator is judged acceptable by the reviser, it will be passed on to the transcriber, which is responsible for transcribing it into text. If the output of the translator is not acceptable, the proposer and/or translator will be reactivated. In addition to evaluating internal language, the reviser is also responsible for analyzing the language produced by the transcriber. Hence, in this model, the reviser can operate at various levels to monitor the output of other processes. Overall, this model is a recursive model of writing because it does not assume a one-way transfer of information from one cognitive process to another. Instead, a process can be influenced and reactivated by the process subsequent to it. For instance, when many ideas are generated by the proposer, the proposer may call on information from the translator to choose the option that is easiest to translate.

Figure 4: Chenoweth and Hayes' model of writing (2001, p. 84)



2.1.2.5 Hayes' (2012) model of writing

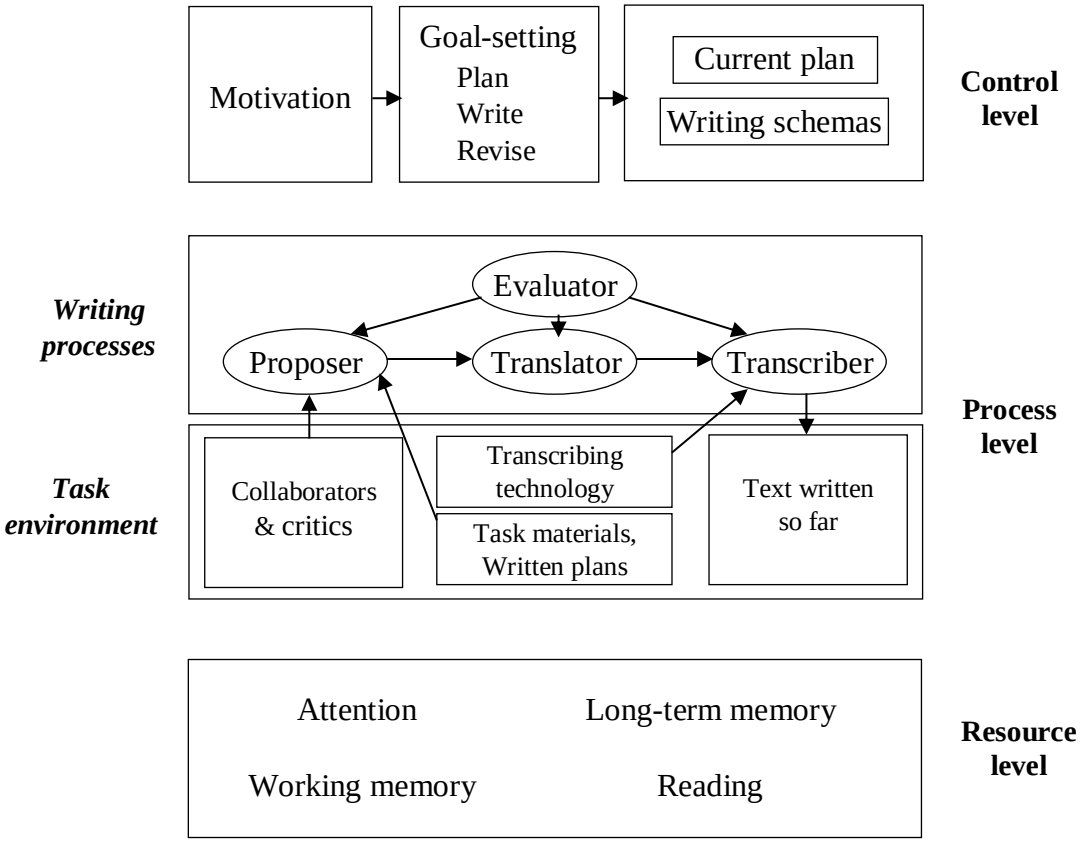
Like Chenoweth and Hayes' (2001) model, Hayes' (2012) new model of writing (Fig. 5) comprises a control, a resource and a process level. However, the resource and control levels of this model are more elaborate. Attention is added to the resource level and the control level consists not only of the task schema, but also motivation and goals for planning, writing and revising. Motivation, which appears as one of the main parts of the model Hayes proposed in 1996, is included in this level due to its influence on the writer's goals.

Regarding the process level, the major change lies in the reconceptualization of the reviser. The reviser in Chenoweth and Hayes' (2001) model is replaced by the evaluator. In this model, revision is viewed as a special writing activity rather than a separate writing process parallel to the proposer, the translator and the transcriber. Hayes points out that revision is a "special application" (2012, p. 376) of writing, which involves the integration of different writing processes. During the processes of revision, the writer engages in multiple writing processes: making plans to improve the text, translating ideas into language, and transcribing their internal language into text. Within this new model, the evaluator is responsible for monitoring the products of the other three processors (the proposer, the translator and the transcriber) and triggering revision when problems are detected.

Another change made to the process level is the incorporation of transcribing technology as part of the task environment. This technology is similar to the composing medium mentioned in Hayes' (1996) writing model. However, the use of the term "transcribing technology" puts more emphasis on the transcription process, a new cognitive writing process proposed in Chenoweth and Hayes' (2001) model. Based on the findings of several empirical studies (e.g. Bourdin & Fayol, 1994; Gee & Walsh, 2007; Hayes & Chenoweth, 2006; Jones & Christensen, 1999), Hayes (2012) hypothesizes that transcription may compete with other writing processes for attentional resources, and different transcribing methods (e.g. writing, typing and using voice-recognition technology) can influence

writing differently. For instance, they might affect the speed of composition and text quality.

Figure 5: Hayes' model of writing (2012, p. 371)



2.1.3 Kellogg's (1996) model of working memory in writing

Apart from the writing models from Hayes (1996, 2012) and Hayes and colleagues (Chenoweth & Hayes, 2001; Hayes & Flower, 1980), another influential cognitive model of writing that explains revision is Kellogg's (1996) model of working memory in writing. Kellogg proposed a three-process model of writing, consisting of formulation, execution and a monitoring process, and outlined the demands of these processes on the writer's working memory (Fig. 6). Drawing on Baddeley's (1986) working memory model, Kellogg assumes that working memory is limited in its capacity and includes different resources responsible for different types of cognitive processing. The phonological loop is

responsible for storing and processing auditory/ verbal information. The visuo-spatial sketchpad is dedicated to storing and processing visual-spatial information. These two resources are slave systems of the central executive, which is a more versatile resource responsible for processing central cognitive tasks (e.g. decision-making, problem-solving and reasoning), controlling slave systems and regulating attention to particular stimuli or activities.

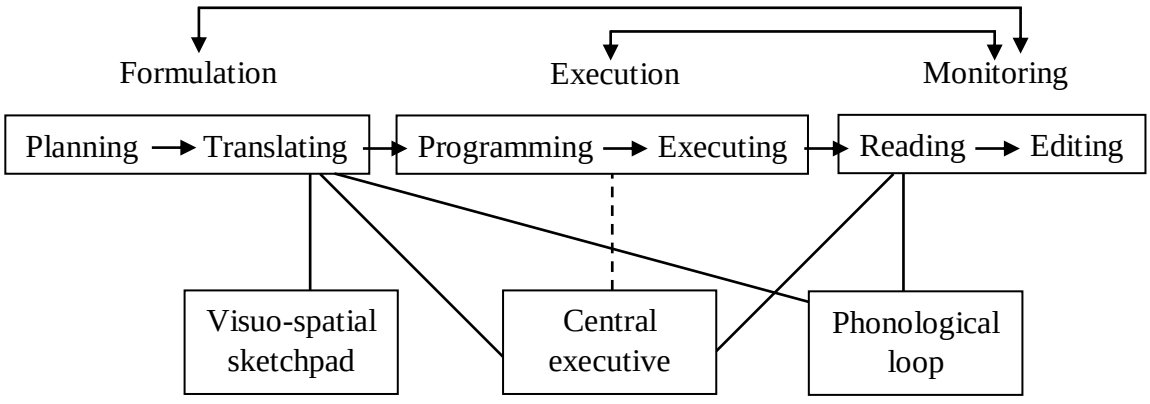
According to Kellogg, writing involves formulating, executing and monitoring. Each major process consists of two subordinate processes. Formulation entails planning and translating. During planning, the writer makes plans, sets writing goals and organizes ideas and content. The product of planning is fed into the translating process, which converts ideas into internal language through linguistic processing. The second main process, execution, includes programming and executing. During this process, the motor system is programmed for language transcription and is activated during execution. Execution can be carried out in various ways, e.g. by dictating, handwriting or typing. Finally, monitoring consists of reading comprehension and editing. Editing involves finding discrepancies between the writer's intentions and the products of the other major processes, and it can occur after both formulation or execution. The products of formulation, such as prelinguistic ideas and content organization, may be edited due to errors of representation or difficulty in translating these ideas. Editing occurs after execution when there are problems in the output of the execution process, such as typing mistakes. The interactions between monitoring and other processes – formulation and execution – are bidirectional. In other words, any problems in the output of the formulation and execution process can trigger monitoring, and the feedback from monitoring can be fed backwards to reactivate these processes.

With respect to the demands of writing processes on working memory, Kellogg speculates that formulation is the most cognitively demanding process. During planning, the visuo-spatial sketchpad and the central executive are required for the visualization of ideas and making plans. Translating involves the phonological loop, which is activated when the writer produces inner speech, and the central executive, which may be necessary during effortful translation. Regarding the execution process, Kellogg hypothesizes that the

cognitive load on the central executive will be high if the writer's motor skills, e.g. typing skills, are not well-practised. That is, when the motor system is less automatic, more demands are placed on the central executive during programming to regulate the writer's cognitive processes. If the motor system is highly automatized, the cognitive load during this process will be minimal. Finally, as far as monitoring is concerned, reading is assumed to activate the central executive and the phonological loop. Editing, which can have multiple forms, such as detecting errors in motor programming and revising ideas, only involves the central executive but is hypothesized to be more cognitively demanding than reading.

In essence, Kellogg's model is based on the assumption of parallel language processing. According to this model, multiple main processes can be activated at the same time as long as working memory is not overtaxed. For instance, a writer whose motor skills are advanced may have enough cognitive resources available to engage in the process of execution and formulation or monitoring simultaneously. When multiple systems operate in parallel, the central executive is responsible for the co-ordination of different processes and the allocation of attention. If the writer's cognitive resources are overloaded due to a particular process, the performance of other processes may suffer. As working memory plays an important role in this model, it is assumed that the writer's working memory span is related to writing skills (Kellogg, 1996). Writers with a greater working memory capacity may be able to carry out writing processes more quickly and efficiently.

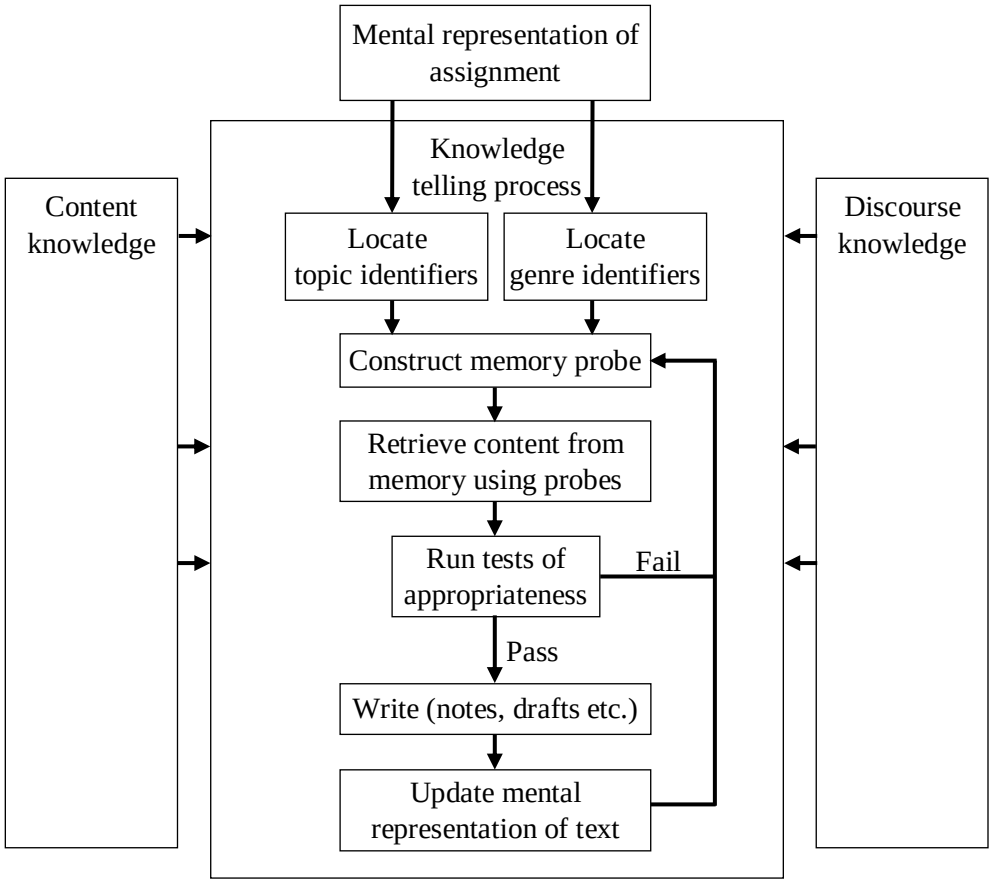
Figure 6: Kellogg's model of working memory in writing (1996, p. 59)



2.1.4 Bereiter and Scardamalia's (1987) models of writing

In addition to the models presented thus far in this chapter, two other influential models of text production are the *knowledge-telling* model (Fig. 7) and the *knowledge-transforming* model (Fig. 8) proposed by Bereiter and Scardamalia (1987). These models aim to describe the writing processes of novice and skilled writers, respectively. The knowledge-telling model is simpler than the knowledge-transforming model and it describes a writing approach that is not oriented towards goals. According to the knowledge-telling model, writing begins with the processes of constructing a mental representation of the assignment and locating identifiers of the topic and genre. Once the writer knows what the writing assignment involves and identifies the topic and genre of the assignment, topical and genre identifiers will trigger a memory probe for content and prime associated concepts. Content retrieved from long-term memory will then be judged for its appropriateness. If the content retrieved is not appropriate, memory probing will be reactivated to search for more appropriate content. Hence, this model presupposes a recursive nature of writing processes. Once text is produced, either internally or externally, the writer may use it as an additional source to further locate identifiers of the topic and genre, subsequently leading to more searching for and retrieval of content. When the information retrieved is influenced by the content of the current text, the text produced can become more coherent (Scardamalia & Bereiter, 1987). In essence, the knowledge-telling model assumes that the text production of novice writers simply involves the transcription of information retrieved from memory. Even though the writer judges this information for its appropriateness, the information retrieved is not modified during writing. The organization and coherence of the content are hypothesized to be dependent on the organization of the information in the writer's long-term memory (Bereiter, Burtis, & Scardamalia, 1988).

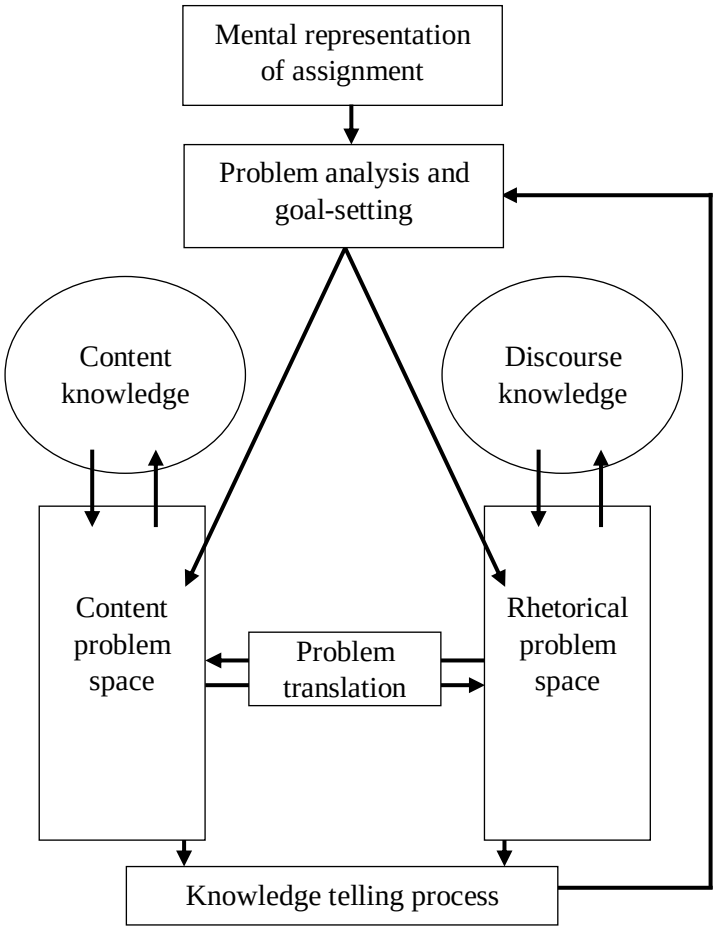
Figure 7: Bereiter and Scardamalia's knowledge-telling model of writing (1987, p. 8)



While the knowledge-telling model only entails the reproduction of knowledge, the knowledge-transforming model, which describes the writing processes of skilled writers, assumes that knowledge is refined during writing to satisfy writing goals and to meet the needs of the audience. Unlike the knowledge-telling model, in this model, writing is conceptualized as a complex problem-solving task in which the writer aims to achieve the goals of writing relating to both content and rhetoric. The model includes the process of knowledge-telling as a subprocess and the additional processes of problem analysis and goal-setting. According to this model, writing entails the construction of an elaborate mental representation of the assignment, which leads to the establishment of writing goals and analysis of rhetorical and communicative problems. As in the knowledge-telling model, this model includes the processes of content retrieval and evaluation. However, the difference is that, in this model, these processes are controlled by the writer's goals. Content is retrieved and evaluated not only to fit the requirements of the topic and the genre of the

assignment, as in the knowledge-telling approach, but also to solve rhetorical and communicative problems and accomplish writing goals. In addition to this, Bereiter and Scardamalia (1987) hypothesize that skilled writers usually adjust their thoughts (hence the term "knowledge-transforming") to satisfy their rhetorical goals as they write, for instance by generating more supporting arguments and examples to make the text more convincing. In doing so, the writers engage in reflective thinking, which may result in a new understanding of the topic or the reorganization of existing knowledge.

Figure 8: Bereiter and Scardamalia's knowledge-transforming model of writing (1987, p. 12)



As far as the revision of written text is concerned, Scardamalia and Bereiter (1987) hypothesize that "Significant revision presupposes a system of goals and goal-seeking procedures" (p. 156). In their view, the writing of a novice writer whose writing corresponds to the knowledge-telling approach is characterized by a paucity of revision and

surface revisions due to the lack of goals during writing. However, this type of writer may be able to carry out more revisions, especially those relating to content, after receiving instruction that aims to develop their knowledge-transforming ability. They propose that novice writers require more than just encouragement to revise or guidelines on how to revise effectively. Instead, these students need assistance with developing a writing system in which the processes of writing and revision are goal-oriented.

2.1.5 Galbraith's (1999, 2009a, 2009b) dual-process model of writing

Although Bereiter and Scardamalia's (1987) models offer important new insights into the writing processes of novice and skilled writers, and how writing can promote understanding of the writing topic, more recent hypotheses about writing processes have been put forth by Galbraith (Baaijen & Galbraith, 2018; Galbraith, 1999, 2009a, 2009b; Galbraith & Baaijen, 2018). Galbraith proposed a dual-process model of writing drawing on connectionist principles of information processing (Rogers & McClelland, 2014; Rumelhart, Smolensky, McClelland, & Hinton, 1986). Unlike Bereiter and Scardamalia (1987), Galbraith does not view writing only as a problem-solving process, but also as a knowledge-constituting one. The dual-process account assumes that there are two writing processes, with contrasting natures, and both are necessary for effective writing. The first process is a *knowledge-retrieval* process, which describes the writing system controlled by pre-determined plans (e.g. plans about content, content appropriateness and the organization of text). It entails the writing processes outlined in Bereiter and Scardamalia's (1987) knowledge-telling and knowledge-transforming models. In Galbraith's view, the knowledge-retrieval process is required for finding gaps in the global structure of the text and improving the appropriateness and organization of the content to ensure the achievement of rhetorical goals. The second system in the dual-process model is a *knowledge-constituting* process. It is a spontaneous text-generation process during which the writer finds out what to write as they are formulating the text. While the content of the text is retrieved from long-term memory in the knowledge-retrieval process, this process involves the synthesis of information. According to Galbraith, the ideas that emerge during the knowledge-constituting process correspond to the writer's implicit understanding of the

topic, and are synthesized within the writer's semantic memory, guided by the implicit organization of content in semantic memory. It is hypothesized that the knowledge-constituting process benefits writing because it ensures that the product of writing captures the writer's implicit understanding of the topic. Furthermore, the synthesis of ideas during this process potentially leads to the discovery of new knowledge and a better understanding of the writing topic. This assumption is different from that of Bereiter and Scardamalia (1987), who hypothesize that the knowledge-transforming strategy (i.e. knowledge retrieval) can develop the writer's understanding of the topic. While increased understanding is associated with more controlled text production (controlled by pre-determined plans) in Bereiter and Scardamalia's view, the dual-process account assumes that it is associated with more spontaneous text production.

2.1.6 Summary of the proposals for writing and revision processes

Thus far, the proposals for L1 revision processes (Bartlett, 1982; Hayes et al., 1987) and the models of cognitive writing processes proposed by Hayes (1996, 2012), Hayes and colleagues (Chenoweth & Hayes, 2001; Hayes & Flower, 1980), Kellogg (1996), Bereiter and Scardamalia (1987) and Galbraith (1999, 2009a, 2009b) have been reviewed. While Bartlett makes a clear distinction between the revision and text production stages, with the latter preceding the former, other researchers propose models that assume that revision can occur at any moment during writing. The output of different processes, e.g. ideas, organization of ideas, inner speech and text already produced, are constantly monitored to ensure the quality of the final text and achievement of the writer's goals. According to these models, the nature of writing processes is recursive. A writing process or sub-process may be reactivated when there is a mismatch between its output and the writer's goals, intentions or knowledge.

An important feature common to the cognitive models reviewed in this chapter is the acknowledgement of the influence of individual differences and the writing environment on writing and revision. Regarding individual differences, differences in terms of, for instance, working memory capacity, linguistic knowledge, motivation, beliefs and

attitudes, have been hypothesized to play an important role in the processes of writing and revision. There may be individual variations in writing goals, revision criteria, the types of problems each writer can identify and the sophistication of the revision procedures each writer can carry out which affect how writers engage in revision and writing activities (Hayes, 1996; Hayes et al., 1987). The writer's task schemas may also be determined by the writing instruction they have received in the past and their level of writing expertise (Hayes, 2012). In addition, writing skills may influence the processes of content generation and the extent to which revision is performed during writing (Bereiter & Scardamalia, 1987; Scardamalia & Bereiter, 1987); also, the writer's motor skills may determine the cognitive demands during the execution process (Kellogg, 1996). Interestingly, most of the writing models reviewed (Chenoweth & Hayes, 2001; Hayes 1996, 2012; Kellogg, 1996) assume the influence of the writer's working memory on writing. As working memory is the resource that co-ordinates multiple cognitive writing processes, the writer's attention capacity is potentially related to writing skill (Kellogg, 1996). Writers with larger working memory spans may be able to write and revise more efficiently. They may be better at processing different types of information simultaneously during composing and keeping information relating to problems in focus when revising.

With respect to the effect of the writing environment, task and other social and physical elements surrounding writing activities are hypothesized to influence the processes of writing and revision in many of the models reviewed. The influence of task demands on writing is presented in Hayes and Flower's (1980) early model and, despite the evolution of this original model, task remains one of the components of Hayes' (2012) most recent model. Similarly, Kellogg (1996) notes that, although L1 language processing is automatic, the translation process may be effortful, depending on tasks. For instance, writing which requires high precision in language could require effortful translation (Kellogg, 1996), which potentially places a high cognitive load on the writer's working memory and affects other ongoing cognitive writing processes (e.g. planning, executing and editing). In terms of other environmental factors, Hayes and his colleagues acknowledge the significance of the composing medium and transcribing technology (Hayes, 1996, 2012) and the social environment (e.g. audience and collaborators) (Hayes, 1996, 2012; Hayes & Flower, 1980)

in writing and revision. Writing may be affected by these factors because different kinds of composing media may afford different types of support for writing and revision (Hayes, 1996), and people write differently for different types of audiences, as writing is a social activity governed by social conventions (Hayes, 1996).

In closing, the cognitive models presented in this section are L1 writing models. As mentioned earlier, to date, no cognitive model has been created specifically to explain L2 writing or revision processes. Compared to L1 writing, the various stages of L2 writing may be more effortful and less efficient due to the lack of automatization in L2 processing. For instance, lexical retrieval and syntactic encoding are slower in L2 than L1 (Kormos, 2006). Reading, which is necessary in revision processes, is also less automatic in L2, and the transcribing process, which is largely automatized in L1, may be effortful in L2 if there are great differences between the writer's L1 and L2 spelling and orthographic systems (Kormos, 2012). This lack of automatization means that L2 writing requires more attention compared to L1 writing and, therefore, L2 writers' working memory may be more prone to being overtaxed during writing. Because of the absence of a cognitive model of L2 writing processes, L2 researchers have drawn on the cognitive processes outlined in L1 writing models when explaining L2 writing. To date, there is accumulating empirical research exploring L2 writing and revision. The section that follows will focus on a review of L2 revision research.

2.2 Empirical studies on L2 revision

To date, a number of L2 revision studies have been carried out to compare L1 revision with L2 revision (Hall, 1990; Lindgren & Sullivan, 2006; Stevenson, Schoonen, & de Glopper, 2006; Thorson, 2000; Whalen & Ménard, 1995). Other studies have investigated the effect of different variables, such as learners' linguistic experience (Chenoweth & Hayes, 2001; Spelman Miller, Lindgren, & Sullivan, 2008), writing skills (Zamel, 1983) and proficiency levels (Barkaoui, 2016; Choi, 2007; Manchón, Roca de Larios, & Murphy, 2009; Raimes, 1987) on L2 revision. Although many factors may potentially influence revision as reviewed in the previous section, the effects of a number of additional variables (e.g. task

characteristics, cognitive individual differences, typing ability) on L2 revision remain under-researched. This section will begin by conducting a brief overview of the findings of previous research comparing L1 and L2 revision and the L2 revision of learners with different proficiency levels. Then, the literature relating to the investigation of L2 revision behaviours in computer-assisted writing will be reviewed, followed by research exploring the connection between L2 revision and different variables relevant to this study: typing ability, task, text quality and writing fluency.

2.2.1 L1 vs L2 revision

A common finding of studies comparing L1 and L2 revision is that writers revise more when composing texts in L2 than L1 (Hall, 1990; Thorson, 2000; Zimmerman, 2000), which could be due to the fact that L2 writers make more errors as they write. More immediate revisions – revisions at the point of inscription – have also been found during L2 than L1 writing (Lindgren & Sullivan, 2006). However, synthesizing the research comparing L1 and L2 writing, Silva (1993) concludes that L2 writers reread and reflect on their text less and are less capable of intuitive revision than L1 writers. In addition, the study conducted by Hall (1990) reveals that, compared to L1 revision, L2 revision is more time-consuming. Hall notes that this could be because simultaneous engagement in revision and text generation is more difficult for L2 writers. With respect to the target of revisions, Whalen and Ménard's (1995) study found that L2 writers were more likely to make linguistic revisions in comparison to L1 writers, while pragmatic and textual revisions (e.g. changes made to content, coherence and text organization) were detected more frequently in L1 than L2 writing. It was hypothesized that the effortfulness of L2 linguistic processing may hinder processing at more global levels of discourse, e.g. ideas generation and content organization (Whalen & Ménard, 1995), therefore higher-level[1] revisions could be impeded

[1] In my study, the terms *high-* and *low-level revisions* strictly refer to the target of revisions. Linguistic and surface revisions, such as grammatical and typographical revisions, are considered to be at low levels. Higher-level revisions are revisions that involve processing at more global levels of language, such as revisions relating to content and organization. Importantly, these terms are not associated with linguistic units of revision, which concern the linguistic boundaries of changes. In this study, these boundaries are referred to explicitly as *linguistic units of revisions*. For instance, sentence-level revisions are carried out on higher-level *linguistic units* than word-level revisions.

by lower-level linguistic revisions. Despite this hypothesis, Stevenson et al. (2006) found that, even though there were more linguistic and typing revisions in L2 writing than in L1 writing, content revisions were comparable across languages.

2.2.2 Influence of L2 proficiency on revision

The findings of research exploring the influence of proficiency on L2 revision show that learners with higher L2 proficiency carry out fewer internal revisions (changes made mentally prior to transcription) and revisions at the point of inscription, but produce more distant revisions than learners of lower proficiency (Barkaoui, 2016; Choi, 2007). Writers of better L2 proficiency also revise more at high levels (such as content, style and discourse) and at large linguistic units (such as the clausal level), while those of lower L2 proficiency tend to revise language form and typography (Barkaoui, 2016; Choi, 2007; Manchón et al., 2009). It can be hypothesized that writers with greater command of L2 produce it more automatically. Thus, they do not experience difficulties with language formulation processes and have more attentional resources available to attend to more global levels of discourse and revise larger linguistic units compared to those of lower L2 proficiency. Indeed, studies which explored the effect of L2 proficiency on time allocation during writing (Manchón et al., 2009; Roca de Larios, Manchón, Murphy, & Marín, 2008) have confirmed that, when proficiency increases, the allocation of time to different writing processes becomes more balanced and less time is devoted to the language formulation process. As far as total revision frequency is concerned, previous research findings have been contradictory. While Barkaoui (2016) found that L2 writers with higher proficiency revised significantly less than those with lower proficiency, Choi (2007) found more revisions in a higher proficiency group. In addition, Raimes' (1987) study did not detect a strong correspondence between L2 proficiency and the number of revisions. As noted in the previous section, cognitive models of L1 writing processes assume the influence of a variety of factors on writing and revision. Based on these models, it may be that L2 revision frequency is influenced by several additional factors, rather than by proficiency alone.

Thus far, an overview of the studies comparing L1 to L2 revision and the L2 revision of different levels of learners has been presented. The section that follows will conduct a detailed review of studies investigating L2 revision in computer-assisted writing.

2.2.3 L2 revision in computer-assisted writing

A number of studies have been carried out to explore the nature of L2 revision in computer-assisted writing (e.g. Barkaoui, 2016; Choi, 2007; Li, 2006; Li & Cumming, 2001; New, 1999; Révész, Michel, & Lee, 2017). The common findings of these studies indicate that, in computer-assisted L2 writing, learners make changes to form more frequently than to content and there is a prevalence of typographical revisions (Barkaoui, 2016; Choi, 2007; New, 1999). In addition, learners were found to make more revisions to lower-level linguistic units (e.g. below-word, word and phrasal levels) than higher ones (e.g. clausal, sentential and multi-sentential levels) (Barkaoui, 2016; Choi, 2007; Li, 2006; New, 1999; Révész, Michel, & Lee, 2017). New (1999) conducted a small-scale study to examine the L2 revision of five intermediate learners of French when writing a magazine article in a word-processing program, *Système-D*, which contains multiple referencing tools, such as a dictionary and a grammar reference. Data were gathered from video recordings of writing sessions and program logs that recorded information accessed by participants during writing. Participants were found to make changes to form much more frequently than to content, and learners with neither high nor low self-perceived writing skills revised content frequently. Regarding the linguistic unit of revisions, graphical revisions accounted for 61.6 to 82 per cent of the total revisions made by each participant, and most of these revisions were related to typography. Word- and phrase-level revisions occurred less frequently, ranging from 6.8 to 25.5 per cent and 3.4 to 14.9 per cent of total revisions, respectively. Revisions made at clausal, sentential and multi-sentential levels were infrequent. New suggests that, due to the predominance of revisions of linguistic form, explicit instruction might be needed to encourage learners to reread and revise for meaning.

In a more recent study, Choi (2007) examined the L2 revision of 12 Korean learners of English during two argumentative individual writing tasks. The data collection involved

the use of keystroke logging software, *Inputlog*, compose-aloud protocols and retrospective interviews. Revisions were coded for types of text production processes (e.g. prelinguistic revisions, revisions at the point of inscription, revisions of the written text) and purposes (e.g. content, vocabulary, grammar), actions (e.g. addition, deletion, substitution), linguistic units (e.g. word, phrase, clause) and remoteness (within the same sentence and across sentences) of revisions. The results revealed that all participants made more external than internal revisions. In other words, more changes were made to the visible text than mentally. Consistent with New's (1999) findings, learners with both high and low proficiency carried out more linguistic-form revisions than content revisions, and revisions related to grammar, vocabulary, punctuation, spelling and typography occurred frequently. Regarding the linguistic unit of revisions, the two most frequent types of revisions were graphical and word-level revisions, and revisions at or beyond the clausal level were detected much less often. Finally, the findings concerning the remoteness of revisions showed that learners revised much more frequently within the same sentence from the point of inscription than distantly across sentences.

In another study, Barkaoui (2016) examined the nature of L2 revision in timed computer-assisted individual writing tasks and the effects of task type, proficiency and the writer's keyboard skills on revision. Data were obtained from 54 learners of English through keystroke logging and computer screen recording. In line with the findings of New's (1999) and Choi's (2007) research, this study found that L2 writers produced more typographical and linguistic revisions than content-related ones. Typographical, linguistic and content revisions accounted for 46, 32 and 13 per cent of all revisions, respectively. Following Broekkamp and van den Bergh's (1996) hypothesis about the possible causes of frequent linguistic revisions during L2 writing, Barkaoui speculates that L2 writers' limited language skills and their preoccupation with linguistic correctness could contribute to the findings observed. L2 writers might revise language form frequently because they make many linguistic mistakes owing to their limited language ability or because they tend to be preoccupied with accuracy. As learners' attention capacity might be limited, the cognitive demands of linguistic revisions may impede content revisions (Barkaoui, 2016). In addition, Barkaoui suggests that another possible cause of the high occurrence of linguistic

revisions could be that learners believe grammatical accuracy and vocabulary are what the reader values. In terms of the linguistic unit of revisions, local revisions made at levels below the word and sentence were found to be more frequent than revisions above the sentence level. Below-sentence level revisions occurred most frequently, accounting for 64 per cent of all revisions, followed by below-word and above-sentence level revisions, which were observed at 29 and 7 per cent, respectively.

More recently, Révész, Michel, and Lee (2017) explored the L2 revision behaviour of 30 Mandarin speakers who were learners of English in a computer-based version of Task 2 of the IELTS academic writing test through keystroke logging. Revisions were examined in terms of ratio of words and ratio of characters in the final text to those actually typed, and classified based on these linguistic units: below-word, word, below-clause, clause and sentence. The findings concerning the product/ process word and character ratios revealed that 79 per cent of the words and 74 per cent of the characters participants typed remained in the final text. As with the findings of New's (1999), Choi's (2007) and Barkaoui's (2016) studies, the data showed a prevalence of local revisions. Most revisions occurred below the word level (M = 90.73 revisions per 100 words). Fewer revisions were found at the word (M = 40.07) and other sub-clausal (M = 43.97) levels, while clauses (M = 3.07) and sentences (M = 2.60) were rarely revised.

Apart from the research examining L2 revision behaviours during computer-assisted writing reviewed thus far, there are also other works that have compared L2 revision in pen-and-paper writing with those in the computer-assisted writing mode. One such work is Li and Cumming's (2001) small-scale longitudinal study, which investigated how a Taiwanese learner of English revised 14 essays over a period of eight months. Revisions during computer-assisted writing were investigated through keystroke logs, and those made during pen-and-paper sessions were examined visually, as the participant was not allowed to use an eraser. Revisions were divided into discourse, syntactic, lexical and morphological levels, which were operationalized as the changes made to complete sentences, sentence structures, words and phrases and morphemes, respectively. The results showed higher frequencies of all levels of revisions in the computer-assisted context

compared to pen-and-paper writing, and the authors observed that it was easier for their participant to rearrange sentences in the computer-assisted writing mode. Based on these findings, Li and Cumming hypothesize that the mechanical difficulty involved in making changes to the text, especially at the discourse level, may be reduced in the computer-assisted writing context, therefore extensive changes are fostered in this environment.

In another comparative study, Li (2006) asked 21 Chinese learners of advanced English proficiency to write two argumentative English essays, one on pen and paper, the other on a computer, while thinking aloud. Compared to pen-and-paper writing, there was higher frequency of revisions in the computer-assisted writing. Consistent with Li and Cumming's (2001) findings, significantly higher numbers of most types of revisions (revisions made at character, word, phrase and sentence levels) were detected in computer-assisted writing than in pen-and-paper writing. The only non-significant difference across mediums was found in the number of paragraph-level revisions, which rarely occurred in the data. Li explains that word processors may facilitate revisions, leading to an increased revision frequency in this mode, and that computer-assisted writing may encourage more revisions of large linguistic units beyond the phrase level than pen-and-paper writing. Yet, in line with the findings of the other studies reported in this section, Li found that most revisions in the computer-assisted context occurred locally at the character, word or phrase levels, rather than at the sentence or paragraph levels.

Taking the findings of these L2 revision studies together, it seems that while most of the revisions found in computer-assisted writing are performed on low-level linguistic units (Barkaoui, 2016; Choi, 2007; Li, 2006; New, 1999; Révész, Michel, & Lee, 2017), comparative research has observed higher frequencies of total revisions and revisions made to both high- and low-level linguistic units in the computer-assisted mode than in pen-and-paper writing (Li, 2006; Li & Cumming, 2001). These findings are in line with the common findings of L1 writing studies which find that writers revise more when writing using a computer than with pen on paper (Van Waes & Schellens, 2003).

2.2.4 Relationship between L2 revision and typing ability

The previous section has discussed studies carried out to explore L2 revision in computer-assisted writing. This and the next three sections will review the literature relating to the relationship between L2 revision and the four variables investigated in this study: typing ability, task, text quality and writing fluency. To the best of my knowledge, to date, very few studies have explored the relationship between typing ability and L2 writing (Barkaoui, 2014, 2016; Barkaoui & Knouzi, 2018) and only one of them has examined the link between this ability and revision (Barkaoui, 2016). The aim of Barkaoui's (2016) research was to explore how keyboard skills influence L2 revision behaviours during two timed writing tasks (writing about a general topic and summary writing) by operationalizing the skills narrowly as typing speed and accuracy. In the study, 54 learners of English were divided into two groups based on the results of two typing tests that measured participants' average typing speed in text-copying tasks. The findings showed that revision behaviours (i.e. the number, temporal location, action and target of revisions) of learners with low and high keyboard skills did not differ significantly. However, there were more precontextual (i.e. revisions made at the point of inscription), content and total revisions, and fewer linguistic revisions on average in the high keyboard skill group compared to the low skill group. The only significant difference detected across groups concerned text length, with significantly longer texts observed in the group with better keyboard ability. Even though the findings offer some insights into the relationship between learners' L2 typing skills and revision, due to the scarcity of research in this domain, more studies examining this relationship are needed to confirm the findings of this study.

2.2.5 Relationship between L2 revision and task

As with the investigation of the influence of typing ability on L2 revision, the domain of research that explores the relationship between L2 revision and task has also received little attention from researchers, with just a few studies investigating it (Barkaoui, 2016; Porte, 1996; Raimes, 1987; Révész, Kourtali, & Mazgutova, 2017). These studies compared L2 revision in different types of tasks and in tasks with varying degrees of task complexity.

Comparing L2 revision in a narrative task with a specified audience and goal to that in a more impersonal argumentative task without any specified audience or purpose using think-aloud protocols, Raimes (1987) found more revisions in the narrative task. Porte's (1996) work, which explored the L2 revision of underachieving learners of English in descriptive and argumentative writing tasks, found higher frequencies of meaning-related revisions in descriptive tasks and non-meaning-related ones in argumentative tasks. Barkaoui's (2016) research showed that the action and linguistic unit of revisions were significantly affected by task types operationalized as independent (writing about a general topic) and integrated (writing a summary of a lecture and a text) tasks. Barkaoui examined participants' keystroke logs and detected higher frequencies of revisions made to all linguistic units and substitutions, and fewer deletions in the integrated task than the independent task. In the field of task complexity research, Révész, Kourtali, and Mazgutova (2017) observed a significantly higher frequency of below-word revisions in a simple (+ content support) than a complex (- content support) argumentative writing task in the keystroke logs of advanced learners. No significant difference was found across tasks in terms of total revision frequency and frequencies of revisions made to other linguistic units. This finding is largely inconsistent with their prediction, which hypothesized that the lack of content support would increase the demands on the planning process and leave fewer cognitive resources for language translation and revision. However, in line with this prediction, the stimulated recall data of this study showed that there was a slightly higher percentage of revisions relating to the language translation process in the simple task condition than in the complex one.

Although the findings of these studies contribute some knowledge regarding the relationship between task and L2 revision, the shortcomings of most research in this domain render the findings tentative. First, while Raimes (1987) aimed to manipulate two task types in terms of the specificity of audience and goal, it was reported that most participants in the study ignored the audience and purpose of the narrative task, simply treating it as an impersonal task. In addition, Raimes' (1987) and Porte's (1996) studies merely compared the raw mean number of revisions and mean revision frequencies across tasks without conducting any test of significance. This lack of significance tests makes it difficult to draw

conclusions from the findings of these studies. Finally, Raimes' (1987) research was carried out with a very small sample size (N = 8) and only eight students in Révész, Kourtali, and Mazgutova's (2017) study participated in a stimulated recall interview. Owing to these limitations, it seems that more empirical studies examining the relationship between task and L2 revision should be conducted to verify the findings of these studies, and reveal how different tasks and different aspects of tasks might affect L2 revision behaviours.

2.2.6 Relationship between L2 revision and text quality

In the previous sections, the studies examining the influence of typing ability and task on L2 revision have been reviewed. This section will discuss research exploring the connection between L2 revision and text quality. To date, only a handful of studies have examined the link between text quality and L2 revision (Choi, 2007; Révész, Michel, & Lee, 2017; Spelman Miller et al., 2008; Stevenson et al., 2006; Tillema, 2012). These studies have operationalized text quality in different ways, including holistic essay scores given based on a combination of different criteria (such as content, organization, task response, coherence and cohesion, lexical resource, complexity, accuracy and fluency) (Choi, 2007; Révész, Michel, & Lee, 2017; Spelman Miller et al., 2008; Tillema, 2012), content and language quality (Stevenson et al., 2006), accuracy (Révész, Michel, & Lee, 2017) and linguistic complexity (Révész, Kourtali, & Mazgutova, 2017; Révész, Michel, & Lee, 2017). Despite different operationalizations of text quality, the common finding of these studies is that there seems to be no significant relationship between total revision frequency and text quality (Choi, 2007; Spelman Miller et al., 2008; Stevenson et al., 2006; Tillema, 2012). In addition, the studies exploring different subcategories of revisions (such as the target, linguistic unit and location of revisions) observed no significant relationship between these subcategories and most text quality measures (Choi, 2007; Révész, Michel, & Lee, 2017; Stevenson et al., 2006). The exceptions to this are the significant correlations found in Révész, Kourtali, and Mazgutova's (2017) and Révész, Michel, and Lee's (2017) studies. In Révész, Michel, and Lee's (2017) research, there was a strong and significant positive correlation between revisions made at the sentence level and lexical complexity. Révész, Kourtali, and Mazgutova's study (2017) detected a significant negative correlation

between revisions made at the clause-and-above level and lexical complexity within the complex task condition. Interestingly, in Choi's (2007) work, which examined the data for learners with high and low proficiency separately, it was found that proficiency potentially influences the relationship between L2 revision and text quality. While there was a connection between increased revision frequency and improved text quality operationalized as holistic essay scores in the data for all participants, the data for the high proficiency group showed an opposite trend, with a connection observed between lower revision frequency and better text quality.

As my study explores the relationship between L2 revision and accuracy, this section will review two relevant studies in detail. One explored the connection between L2 revision and language quality (Stevenson et al., 2006), the other between L2 revision and accuracy (Révész, Michel, & Lee, 2017). Stevenson et al.'s (2006) research investigated the revisions of 22 school students while writing argumentative texts in Dutch (L1) and English on a computer. The quality of the language and content of their texts was judged by a panel of four raters. The authors hypothesized that, rather than the total frequency of revisions, the outcome of revisions (i.e. whether revisions could improve the text) and the frequency of some specific types of revisions might be more important for the improvement of text quality. It was predicted that the attention required for L2 linguistic revisions would diminish the attention available for conceptual-level revising processes. As a result, linguistic revisions might negatively affect the quality of the content of English texts. The study explored both external and internal revisions (changes made to visible text and mentally before transcription) through think-aloud and keystroke-logging techniques. Revisions were coded in four dimensions, each consisting of multiple subcategories: orientation (content, language and typing), domain (e.g. clause-and-above and below-word levels), location (e.g. at the point of inscription and to previously written text) and action (e.g. addition and deletion). Correlation analyses were carried out to investigate the relationships between each aspect of text quality (language and content) and the frequency for the total and for each subcategory of revisions. It was found that, for both languages, total revision frequency did not significantly correlate with either language or content

quality. Contrary to the authors' hypothesis, the results regarding English writing revealed no significant relationship between any of the revision subcategories and text quality.

In a more recent study, Révész, Michel, and Lee (2017) investigated the relationship between text quality and online writing behaviours – fluency, pausing and revision – of 30 L2 learners of English utilizing a computer-based IELTS academic writing test. The revision data were obtained from keystroke logging and examined in terms of the ratio of words and the ratio of characters in the final text to those actually typed. Revisions were also classified based on linguistic units: below-word, word, below-clause, clause and sentence. Text quality was operationalized in various ways, including language accuracy scores (i.e. errors produced per 100 words). The output of correlation analyses showed that there was no significant relationship between revision behaviours and accuracy. The product/ process ratios and frequencies of revisions made to all linguistic units did not correlate significantly with accuracy scores.

Despite the non-significant overall findings, it should be noted that both of these studies investigated the relationship between text quality and L2 revision behaviours by carrying out correlation analyses using the data of all participants. The findings might have been different if the researchers had compared the quality of the text produced by each participant before and after revisions. Indeed, Stevenson et al. (2006) found that approximately two thirds of the error-triggered L2 revisions in their study were carried out successfully. Additionally, it might be more useful to perform correlation analyses to test the relationship between the success rate of error revisions and text quality, rather than merely investigating the relationship between text quality and revision frequencies. Although it is more than a decade ago since Stevenson et al. (2006) suggested a possible link between the outcome of revisions and text quality, to the best of my knowledge, there has been no attempt to test this hypothesis empirically. In this section, the literature relevant to the relationship between L2 revision and text quality has been reviewed. The next section will turn to the connection between L2 revision and writing fluency.

2.2.7 Relationship between L2 revision and writing fluency

To date, some researchers have associated revision with writing fluency. For instance, in Van Waes and Leijten's (2015) multidimensional model of writing fluency, revision is included as one of the four main components of the model along with production, process variance and pausing behaviour. A high rate of production, short pauses and low frequency of revisions are usually assumed to be the characteristics of fluent writing (MacArthur, Graham, & Fitzgerald, 2008).

In addition to this, to date, several L2 studies have explored revision and fluency in conjunction with each other as writing behaviours (e.g. Chenoweth & Hayes, 2001; Palviainen, Kalaja, & Mäntylä, 2012; Révész, Michel, & Lee, 2017; Spelman Miller et al., 2008; Van Waes & Leijten, 2015). First, Chenoweth and Hayes' (2001) study investigated the effect of language experience on fluency, burst length and revision, by comparing L1 and L2 writing and the writing of students in their third and fifth semester of L2 study. They found that participants wrote significantly less fluently, as measured by words written per minute, and produced significantly shorter bursts (i.e. uninterrupted stretches) of text when writing in their L2 than L1. A significantly higher percentage of R-bursts (i.e. stretches of text produced between revisions) was also found in L2 writing compared to L1 writing and during the third semester compared to the fifth semester. These results indicate that increased linguistic experience and/or proficiency could be related to an increase in fluency and a decrease in the frequency of revisions.

Consistent with these findings, Van Waes and Leijten (2015) compared L1 to L2 writing and found that higher proficiency might be related to improved fluency and lower revision frequency. Writers were less productive in L2 writing, as measured by words and characters per minute, and the production of L2 text was more fragmented, involving smaller units of text produced between pauses. In addition, during L2 text production, more time was spent on pausing and the length of between-word pauses was longer in comparison to L1 writing. In terms of revision, more revisions were observed in L2 than L1 text production, as assessed by the product/ process word and character ratios.

While the findings reported thus far suggest that better fluency may go hand in hand with fewer revisions, Palviainen et al.'s (2012) study found an opposite trend. Their research explored how L2 writing fluency is related to text quality, which was assessed based on the Common European Framework of Reference (CEFR) rating scale of language proficiency. The main finding of this research was that the production of advanced-level (i.e. C-level) texts was more fluent than that of intermediate-level (i.e. B-level) ones. They observed longer pausing time and significantly fewer characters produced per minute, shorter bursts and decreased writing speed during bursts in the composition process of intermediate-level texts compared to that of advanced-level texts. Participants also wrote more fluently at B2 level than at B1 level, but no significant difference was observed with respect to fluency between C1 and C2 levels. Regarding revision, Palviainen et al. found that revisions, measured in terms of deletions, occurred less frequently at intermediate (B1/B2) levels than at advanced (C1/C2) ones. In sum, these findings suggest that higher proficiency levels could be associated with improved fluency and more revisions.

Finally, while other studies showed either an increase or a decrease in revision frequency as writing fluency increases, Spelman Miller et al. (2008) detected no connection between how frequently writers revise and the speed of writing. They examined the effect of year of study on L2 writing performance by asking each student to compose an L2 essay each year for a period of three years. The study found that year of study significantly affected writing fluency, but not revision. Over a three-year data collection period, learners became more fluent (i.e. produced significantly more characters per minute, wrote significantly longer stretches of text between pauses and/or revisions and became significantly more fluent during bursts), but the number of revisions did not change significantly. In addition, although increased fluency was a strong predictor of improved text quality, the number of revisions was not. These findings suggest that revision may not be related to fluency. The lack of association between revision and fluency observed in this research is also supported by the findings of Révész, Michel, and Lee (2017) who explored how online writing behaviours relate to text quality and IELTS scores. They found that some measures of IELTS scores and lexical, syntactic and discourse complexity (e.g. total writing scores, use

of New-GSL 1000 words, clauses per t-unit and use of causal connectives) were significantly associated with fluency, but not with any measure of revision. In addition, a measure of lexical complexity that was found to significantly correlate with revision (i.e. use of off-list words) was not associated with any fluency measure.

In sum, although it is commonly believed that revision is closely related to fluency, previous L2 findings have not shown a clear connection between these two constructs. While some studies observed higher fluency with fewer revisions, others found a simultaneous increase in both constructs or no association between them.

2.3 Summary

Revision, i.e. the process of making changes that can occur at any point during writing, plays an important role in writing and is a key component of most cognitive models of writing. These models assume that writing and revision are influenced by various factors relating to individual differences and the writing environment. Nonetheless, the effects of many variables that may influence revision have remained under-researched in L2 revision studies. Of particular interest to this thesis are the effects of typing ability and task on L2 revision, and how learners revise during computer-assisted writing. First, previous research found that typing ability did not significantly affect L2 revision behaviours. Since only one study has examined this effect, more studies in this domain are needed to confirm this finding. Second, in terms of task effects, some research has been carried out to investigate L2 revision in different types of tasks and in tasks with varying degrees of task complexity. However, the number of studies in this domain is small and the shortcomings of most of these studies render their findings tentative. Third, previous research comparing L2 revision in pen-and-paper writing and computer-assisted writing has detected higher frequencies of total revisions and revisions made to both high- and low-level linguistic units in the computer-assisted mode. Other studies that explored L2 revision during computer-assisted writing found a prevalence of typographical revisions and that L2 revisions in this mode tended to occur to low-level linguistic units (e.g. below-word, word and phrasal levels) and target linguistic form.

As with research exploring the influence of different variables on L2 revision, research that examines the connection between L2 revision and text quality is still lacking. Previous research has found no significant relationship between L2 revision behaviours and accuracy (Révész, Michel, & Lee, 2017) or language quality (Stevenson et al., 2006). As discussed in this chapter, these non-significant findings might be due to the method of analysis and how revision behaviour was conceptualized. In analyzing the relationship between L2 revision and text quality, it may be important to compare the quality of the text produced by each participant before and after revisions and to examine the success rate of error revisions.

Finally, in terms of the relationship between revision and fluency, while assumptions have been made regarding the association between these two constructs, the findings of previous L2 writing studies are inconclusive.

In this chapter, I have conducted a review of writing and revision processes and empirical studies on L2 revision. Overall, there is still room for further research in this area. The next chapter will discuss another topic relevant to my research – text-based synchronous computer-mediated communication.

Chapter 3: Text-based synchronous computer-mediated communication (SCMC)

As mentioned in the introduction to this study, computer-mediated communication (CMC) refers to human-human interaction taking place via the instrumentality of computers (Herring, 1996); and it can take various forms, such as text-based, voice-based, video-based or a combination of these. Generally, CMC is categorized into two types – synchronous and asynchronous CMC – based on the simultaneity of communication. Interaction in asynchronous computer-mediated communication (ACMC) is not real-time and users do not need to be online at the same time. Once a person sends his/her message, it will be saved and the recipient(s) can choose to view and respond to it whenever they want. Examples of ACMC are e-mails, blogs and online discussion boards. As opposed to ACMC, synchronous computer-mediated communication (SCMC) allows interlocutors to communicate in real time. Chat rooms, instant messaging and video-conferencing are examples of SCMC. However, as noted by Murray (2000), SCMC may not be entirely synchronous because the speed of typing and the Internet can cause delays during communication.

Because the focus of this research is on computer-based text chat, or text-based SCMC, this chapter will provide an overview of the issues relating to this medium. It will begin with an overview of the characteristics of text-based SCMC and language monitoring in this medium. Then, it proceeds to a discussion of how text-based SCMC may facilitate L2 learning by focusing on its potential for enhancing students' attention to form. At the end of the chapter, empirical studies carried out to explore attention to form and L2 revision in this mode will be reviewed. Henceforth, I will refer to text-based SCMC interchangeably with the terms *online chat*, *text chat* and *chat*.

3.1 Characteristics of text-based SCMC

Text-based SCMC refers to communication between humans that utilizes computers as the means to transfer text messages between users in real time. While in other types of SCMC,

such as voice- or video-based SCMC, users can express meanings through auditory cues and/or gestures, text-based SCMC relies solely on text. This form of communication is commonly viewed as the integration of oral and written communication, and many properties of online chat have been found to resemble those of speaking and writing.

First, text-based SCMC has a resemblance to oral communication, in that it requires rapid and spontaneous information exchange between interlocutors. Online chat users have less time to plan, edit and process language in this medium compared to ACMC and writing. Second, similar to spoken discourse, language in text-based SCMC tends to be casual and contain short turns (Kern, 1995), a first-person perspective (Yates, 1996) and features such as jokes, sarcastic comments and requests for personal information (Darhower, 2002; Sotillo, 2000). As in oral interaction, in text-based SCMC, interlocutors negotiate for meaning (Arslanyilmaz, 2012; Rouhshad, Wigglesworth, & Storch, 2016; Smith, 2003, 2005; Tudini, 2003; Yuksel & Inan, 2014) and use various communication strategies, such as word invention, circumlocution and requests for assistance (Lee, 2001, 2002). Finally, in the same way that speakers can stress parts of their speech, text-based SCMC users can emphasize parts of their message by manipulating the text format, e.g. by italicizing and bolding. As text-based SCMC is similar to oral interaction, Razagifard (2013) argues that text-based SCMC might involve the same cognitive processing as required in face-to-face conversation, and the "only" (p. 271) difference between these two kinds of communication lies in the fact that speech organs are replaced by hands and fingers when producing language in text-based SCMC.

However, this statement ignores the fact that language in text-based SCMC shares many characteristics with written discourse. The output produced during both text-based SCMC and writing can be seen and kept as permanent records, and both modes require the use of punctuation and text formatting (Smith, 2003). Moreover, like written language, language in online chat can be characterized by high lexical density (Smith, 2003; Yates, 1996). It has been reported that language in online chat tends to be both lexically and syntactically more complex than spoken discourse (Kern, 1995; Warschauer, 1996).

Apart from having similar features to speaking and writing, text-based SCMC also has its own unique characteristics. First, it has been found that people tend to ignore surface errors and use abbreviations and formulaic phrases in this medium (Murray, 2000). Second, while spoken interaction is carried out in real time on the spot, the speed of the Internet and typing can cause time lags during chat. In addition, online chat users can compose messages in their private dialogue boxes before sending them to interlocutors, which means that composers can edit and review their messages before they are seen by other people, and many interlocutors are able to compose their messages at the same time. This is unlike face-to-face communication in which all produced utterances are perceived instantly by the interlocutors, and interlocutors usually take turns to speak during conversation. Apart from the above unique qualities of text-based SCMC, this medium is also considered a "lean" medium (Smith, 2010, p. 79) because it lacks paralinguistic cues, such as facial expressions, body gestures and tones. Thus, signalling problems that occur during communication can be difficult (Ortega, 2009; Smith, 2005, 2010) and users are required to indicate these problems explicitly through the use of language. Without paralinguistic cues, chat users have to resort to expressing emotions through the use of emoticons – graphical representations of facial expressions.

Having discussed the general characteristics of online chat in this section, I will now turn to an explanation of language monitoring in this medium in the next section.

3.2 Monitoring in text-based SCMC

As mentioned in the previous section, dialogue boxes allow chat users to view and edit messages before sending them to interlocutors. As a result, unlike in oral communication, not all of the language output produced in online chat is public, and monitoring in this type of communication is different from monitoring in face-to-face interaction. O'Rourke (2008) compared monitoring in both modes and distinguished three stages of monitoring in online chat (Table 1). The first stage concerns monitoring before drafting. This type of monitoring resembles covert repairs in speech, which occur when the speakers monitor their language internally during language formulation. The subsequent stage of monitoring in online chat takes place when chat users are drafting their messages. During drafting, users can monitor

their output as it is typed on the screen and edit their messages before sending them to interlocutors. In his publication, Smith (2008) coined the term *CMCovert repairs* to refer to the adjustments made at this stage. They are defined as changes that are overt from the perspective of the composers but remain covert to interlocutors, and they are found specifically in online chat. Similar to O'Rourke (2008), Smith makes a connection between text chat and speaking by adopting the term *repair,* which is generally used to describe language adjustments made during oral communication. Smith points out the advantage of this kind of self-repair for SLA by claiming that, in making CMCovert repairs, learners can benefit from monitoring the actual output of their language production. In particular, he states that "This type of self-repair is different from true covert repair since, in the latter, we may not expect the same output-related benefit that may only be present upon actually producing the target language" (p. 90). Likewise, O'Rourke (2008) also notes that "the drafting process (i.e., typing, optionally re-reading, and optionally altering) is obviously central to questions of learners' attention to the form of their own output and their ability to identify and amend inaccuracies" (p. 235). Finally, the last stage of monitoring in text chat occurs after messages are sent. Due to the permanent nature of chat records, chat users can always re-read their previous messages and send further messages to repair them if they find any errors. Because this type of adjustment is made after erroneous text has been sent, it is discernible to all participants of a chat conversation. Therefore, it is similar to overt repairs in speech, which occur when speakers overtly repair their utterances while they are speaking.

Although O'Rourke's (2008) work compares language monitoring in text-based SCMC to that in speech, monitoring during online chat bears a close resemblance to monitoring in writing. As explained in the previous chapter, revisions can occur at any moment during writing and the output of various cognitive processes (e.g. ideas, organization of ideas, inner speech and text already produced) is constantly monitored. Similarly, during online chat communication, ideas, the organization of ideas and inner speech can be repaired during the first stage of online chat monitoring outlined by O'Rourke (2008), and adjustments to previously written text are made in the second and third stages.

Table 1: Monitoring stages in speech and text-based SCMC (from O'Rourke, 2008, p. 235)

	Internal	External	
		Private	**Public**
SCMC	Monitoring internal formulation	- Monitoring typed output while typing - Re-reading typed output after typing, before sending	Re-reading typed output after sending
Speech	Monitoring internal formulation		Monitoring spoken output while speaking

3.3 Text-based SCMC in SLA

Thus far, I have explained the properties of text-based SCMC and language monitoring in this medium. In this section, I will discuss how text-based SCMC can be a useful tool for L2 learning. At the end of this section, I will discuss the notion of attention to form, which is central to the claims made regarding the benefits of online chat and the potential of text-based SCMC for enhancing this attention.

3.3.1 Writing to learn L2 through text-based SCMC

Manchón (2011) notes the significance of writing as a means to promote L2 learning. She distinguishes between *learning to write* and *writing to learn*. The first refers to the process of developing learners' writing skills, while the latter entails how one learns a second language (*writing to learn language*) or content knowledge (*writing to learn content*) by engaging in writing activities. Drawing on Cumming (1990), who hypothesized that writing may facilitate the analysis and consolidation of L2 knowledge and form-meaning mappings, she argues that it is possible to utilize writing as a means to learn a second language. Reviewing the research on various dimensions of writing and L2 learning, Manchón (2011) concludes that writing, especially when performed collaboratively, may

facilitate language development. In addition, based on the synthesis of these research findings, she concludes that learners engage in various types of linguistic processing while writing, e.g. noticing and attending to linguistic form, formulating and testing L2 hypotheses, assessing linguistic options by utilizing both explicit and implicit knowledge, engaging in metalinguistic reflection. Further to this, the findings of research comparing speaking to writing (Adams, 2006; Adams & Ross-Feldman, 2008; Niu, 2009) indicate more linguistic processing in the written mode. For instance, in Niu's (2009) study, more attention to grammar and discourse features and more extensive discussions of language-related issues were observed in a writing task compared to an oral task.

Considering the above, text-based SCMC, as a type of communication that involves writing, is potentially useful for promoting language development. Similar to traditional writing, text-based SCMC may encourage a higher degree of L2 processing when compared to oral interaction (Chapelle, 2001). Learners may, for example, attend more to linguistic form and reflect more on input and output. Attention to form in online chat is the focus of this study and will be discussed in detail in the next section. Essentially, as text-based SCMC is a hybrid of writing and speaking, it is possible that this type of online communication could play a facilitative role in language learning by combining the advantages of speaking and writing. In other words, by engaging in online chat, learners may benefit from opportunities to negotiate meaning and receive immediate feedback, as in oral interaction, and opportunities for intensive L2 processing, as in writing.

Indeed, text-based SCMC could be a unique medium for writing to learn due to its distinctive characteristics and affordances, which may have important implications for L2 learning (Adams, Nik, & Newton, 2015). The reduced paralinguistic cues and relatively more time for information processing, for instance, have been hypothesized to create a less stressful environment for learners compared to face-to-face interaction. According to Beauvois (1997), CMC is a "less pressured environment that tends to lower the affective filter" (p. 171), and she clarifies this with an extract from a student interview:

> "In class, when the teacher asks a question and looks at you, and there is this big silence, and everyone is waiting, and you're sitting there ... But in the lab, you can take your

time, write your answer, or whatever you want to say, and it doesn't get sent until you are ready to send it. And nobody is waiting and wishing you'd hurry up!" (pp. 171–172)

Because text-based SCMC may be a less stressful environment compared to face-to-face communication, it is possible that this online mode can reduce learners' anxiety (Kelm, 1996; Kern, 1995; Satar & Özdener, 2008). Ultimately, online chat may be a medium suitable for promoting learners' participation (Chun, 1994; Kelm, 1992; Kern, 1995; Sullivan & Pratt, 1996) and motivating them to learn (Beauvois, 1992; Kern, 1995).

In addition to these characteristics and affordances, text chat might be an ideal mode for writing to learn L2, because it allows learners to engage in authentic communication outside the classroom without being constrained by time or space. The opportunities afforded by this medium mean that learning becomes more autonomous, as learners are less dependent on the teacher. This increased learner autonomy has been recognized as being beneficial for learning (Allford & Pachler, 2007; Benson, 2000; Conacher & Kelly-Holmes, 2007). In addition, because learners can interact with non-native peers and native speakers who may be living around the globe at any given time, they can benefit from increased opportunites for L2 interaction. Interaction is crucial for L2 learning, based on two theories of SLA. First, sociocultural theory (Vygotsky, 1978) assumes that learning entails the co-construction of knowledge through social interaction. More proficient interlocutors can assist, or *scaffold*, those with lower language ability to produce the L2 they cannot produce on their own during social interaction (Lantolf & Thorne, 2007). Second, the interaction hypothesis (Long, 1996) argues that interaction is important for L2 learning, because it creates opportunities for learners to negotiate meaning when problems arise during communication. When there is a communication breakdown, learners may, for instance, rephrase themselves, request further clarification or ask questions to confirm their understanding. Interactionist theory claims that the negotiation of meaning potentially benefits learners because it may increase mutual comprehension, trigger the provision of comprehensible input and create opportunities to receive corrective feedback and attend to linguistic form (Long, 1983, 1996). In addition, during interaction, there could be instances where learners talk about, correct or question the use of some L2 features, i.e. language-related episodes (LREs) (Swain & Lapkin, 1998), which direct learners' attention to these

language elements. The interaction hypothesis is based on Swain's (1985, 1995, 2005) comprehensible output hypothesis, which views output as an important factor for L2 development. Interaction results in opportunities to produce output, which is assumed by Swain (1995) to have three beneficial functions. By producing L2 output, learners can test hypotheses relating to the target language, notice gaps in their L2 knowledge and build their metalinguistic awareness by discussing linguistic problems with their interlocutors (Swain, 1995). In essence, the interactionist perspective views interaction as important for language learning and it emphasizes the potential of interaction for promoting attention to form. Because online chat enables learners to interact in L2 and attend to linguistic form, it can facilitate language learning.

With the possible benefits mentioned above, text-based SCMC may be a compelling medium for writing to learn L2. In the following section, I will discuss the notion of attention to form in SLA and the potential of text chat for promoting students' attention to form.

3.3.2 Attention to form in text-based SCMC

According to Robinson (2003), attention is defined as:

> …the process that encodes language input, keeps it active in working and short-term memory, and retrieves it from long-term memory. Attention and memory structures can be viewed hierarchically. The focus of attention is a subset of short-term memory, and short-term memory is that part of long-term memory in a currently heightened state of activation. (p. 631)

The notion of attention to form is central to SLA theories. As mentioned in Section 3.3.1, the benefits of L2 interaction revolve around how it can make students attend to linguistic form. To date, many researchers in the field of task-based language teaching (TBLT) have also highlighted the importance of attention to form for promoting L2 development (Ellis, Basturkmen, & Loewen, 2001; Long, 2015; Skehan, 1996, 1998). In addition, the findings of studies in immersion contexts have shown that instruction which focuses solely on

meaning may not lead to gains in sociolinguistic or grammatical competence (Genesee, 1987; Swain, 1985).

The benefit of enhancing learners' attention to form is supported by Schmidt's (1990, 1994a) noticing hypothesis, which assumes the importance of attention for SLA, and VanPatten's (1994, 1996, 2004) input processing theory, which hypothesizes the primacy of meaning during input processing. With regard to Schmidt's noticing hypothesis, Schmidt (1994b, 2001) posits that the conscious noticing of certain L2 elements facilitates learning of them. In addition, he hypothesizes that noticing is crucial for learning language features that are redundant, non-salient and infrequent (Schmidt, 2001). According to Schmidt (1994b), noticing is "registration of the occurrence of a stimulus event in conscious awareness and its subsequent storage in long-term memory" (p. 166). Within this perspective, noticing involves both attention and awareness. The assumption about the importance of attention and awareness for L2 learning has also been supported by Gass (1988) and Robinson (1995). Gass (1988) hypothesizes that in the initial stage of information processing, attention and awareness are involved during the selection of input. She proposes that input is selected based on learners' *apperception*, the process of noticing a certain aspect of language due to its relevance to prior knowledge. Like Schmidt, Gass believes that noticing entails both attention and awareness. Apperceived language elements are those elements that are recognizable to learners because they are linked to learners' prior knowledge, e.g. existing L1 or L2 knowledge. According to Gass (1988), input must first be apperceived in order to be processed further. Similarly, Robinson (1995) also emphasizes the significance of both attention and awareness. He proposes that, in order for input encoding to occur in one's long-term memory, input must first be detected and rehearsed in short-term memory (i.e. the activated part of long-term memory). In this view, detection requires attention and rehearsal involves awareness, and stimuli that are detected but not rehearsed can be retained in one's short-term memory for only a few seconds. Therefore, for Robinson, both attention and awareness are crucial in the process of information processing. Regarding VanPatten's (1994, 1996, 2004) input processing theory, it supports the notion of attention to form because it assumes that attention to meaning competes with attention to form and that meaning is primary during L2 input

processing. This hypothesis was formed based on VanPatten's (1990) empirical findings and the view that one's attentional capacity is limited (Kahneman, 1973; Wickens, 1984). According to VanPatten (1990, 1994, 1996, 2004), learners have difficulty in attending to form and meaning at the same time, and their attentional resources are directed first to meaning before communicatively redundant linguistic forms.

To date, researchers have proposed different ways to draw attention to form during meaningful L2 interaction, including those relating to instructional procedures (e.g. pre-task instruction), task variables (e.g. task modality) and the characteristics of corrective feedback (e.g. salience). In terms of task modality, it has been hypothesized that text-based SCMC can foster learners' attention to form. Indeed, attention to form is central to claims regarding the benefits of online chat for language learning. Text-based SCMC has been proposed as a medium that "provides a natural way to link a focus on meaning with a focus on form" (Salaberry, 2000, p.6). During online chat, learners can attend to both form and content, while the flow of conversation is still maintained (Smith, 2003). Smith and Renaud (2013) highlight the potential of text-based SCMC for enhancing attention to form, stating that "Essentially, then, the potential advantages afforded SLA by text-based SCMC come down to the construct of attention and noticing" (p. 149). Shekary and Tahririan (2006) also note that SCMC, with its unique characteristics, is considered "an intellectual amplifier" (p. 567) which promotes noticing.

Text-based SCMC has been widely claimed to foster attention to form due to its distinctive characteristics and affordances – its salient and written nature, re-readability and permanent records of chat conversations, and increased online planning time during interaction. First, attention to form can be fostered in text-based SCMC due to its written nature. Input and output are displayed clearly on the screen in the form of text in this medium, therefore the visual saliency of linguistic form is increased compared to that in spoken interaction (Chapelle, 2001). Because the salience of linguistic form can promote noticing (Gass, 1997), it may be easier for learners to notice L2 linguistic form, particularly that which is not perceptually salient in the oral mode, during online chat (Sauro, 2009). The salience of their own L2 output may also make learners conscious about their language use and notice

gaps in their output, which can lead to self-corrections (Lai & Zhao, 2006). Furthermore, it has been hypothesized that the salient nature of text-based SCMC may promote learners' noticing of corrective feedback (Sauro, 2009). When learners receive L2 feedback during chat, feedback and erroneous output are shown on the screen and can be compared visually, not mentally, as in the case of oral communication. Related to this, Warschauer (1999) posits that modified input – input that has been simplified or modified – is potentially more obvious and memorable in text-based SCMC because of its salient nature.

In addition to the salient nature of text-based SCMC, online chat has also been claimed to promote attention to form due to its re-readability and permanent record. In this medium, learners can scroll up and down the dialogue window to read and re-read chat messages, allowing them to reflect on both their interlocutor's and their own output. Kitade (2000) states that the ability to scroll back through conversations allows students to review their messages and notice their own errors, and think about what has been discussed. Sauro (2009) notes that the re-readability of chat messages provides learners with opportunities to review the use of linguistic structures available in previous turns when learners are uncertain about the correct formulation of a given construction. Therefore, it may afford a quick confirmation of a hypothesis and encourage the reuse of a linguistic form (Sauro, 2009). In terms of permanent records, all messages in a chat conversation are archived as written records and can be accessed by learners for language review after a chat session. These records provide additional opportunities to reflect on input and problematic linguistic output, and they are a useful tool for documenting L2 performance in a less invasive way, compared to recordings of oral interaction (Smith, 2003).

Finally, attention to form may be engendered by the increased online planning time afforded by online chat. Online planning refers to instances when learners use available time during task performance to plan "on the fly" (Skehan, 2007, p. 57). Even though the feel of the interactional flow during text-based SCMC remains similar to that in oral communication (Smith, 2003), pauses are better tolerated in online chat due to the time lag in message transmission and the typing speed relative to the speed of speaking (Sauro & Smith, 2010). This slower speed of communication allows for more online planning time,

hence more time for input processing and reflecting on output. Payne and Whitney (2002) point out that the slower speed of online chat interaction in comparison to face-to-face conversation potentially reduces the processing demand, because "the amount of language that an individual has to parse, comprehend, and respond to is lower for a given time period" (p. 14). Drawing on Williams' (2005) hypothesis, which assumes the influence of time pressure on noticing, Sauro (2009) hypothesizes that learners may notice a wider range of linguistic form during online chat compared to oral communication, because there are less demanding time constraints for input processing in this online mode. More online planning time has also been assumed to free up learners' memory and the attentional resources required during interaction and make more attention available for pre- and post-production monitoring (Yuan & Ellis, 2003). In other words, the reduced sense of urgency in SCMC may facilitate better monitoring, because it allows learners time to think about the content and linguistic form of the message they are going to convey and notice errors in their output.

Despite these possible benefits of text-based SCMC, some concerns have been raised about its potential to enhance learners' attention to form. First, it has been claimed that text-based SCMC might impose higher cognitive demands on learners than face-to-face interaction. As typing requires a certain amount of learners' attention, they may have fewer attentional resources available for noticing (Lai & Zhao, 2006). In line with this, Murray (2000) observes that learners may ignore minor errors, e.g. typographical and spelling mistakes, to compensate for their speed of typing, which is slower than speaking. Second, the discontinuity of communication is another cause for concern regarding attention to form during online chat. Smith (2003) observes that the disruption of turn adjacency caused by the time lag during message transmission and the lack of a strict turn-taking mechanism in chat can lead to discontinuities in negotiation routines. When this happens, it may become more difficult for learners to notice the problematic form which is the target of negotiation (Rouhshad et al., 2016). Lastly, although it has been claimed that online chat provides more time for information processing and reflecting on form, some researchers disagree. They have hypothesized that the relatively rapid nature of text-based SCMC predisposes learners to attend to meaning rather than form. In other words, the need to communicate rapidly may override the need to communicate using linguistically accurate and sophisticated

language. Rouhshad et al. (2016), for example, point out that the meaning-oriented nature of online chat tasks probably cancels out the positive effects of the extra language processing time afforded by text-based SCMC, and that this additional time might be used for typing rather than attending to form. Similarly, Iwasaki and Oliver (2003) suggest that learners do not have much time to review their messages because the communicative demands of chat require them to interact relatively quickly.

3.4 Empirical studies on attention to form in text-based SCMC

As discussed in the previous section, the central claims regarding the advantages of online chat relate to its ability to foster learners' attention to form. This section will present a brief overview of the research findings on attention to form in online chat. In previous L2 chat studies, attention to form and noticing have been operationalized in different ways, such as language-related episodes (LREs), noticing of corrective feedback, corrections and noticing of negotiation of meaning episodes. Overall, the findings of these studies seem to be mixed, with some indicating increased attention to form in this medium, while others do not show such benefits.

Regarding the positive findings, some researchers have found high numbers of LREs in online chat (Chen, 2008; Shekary & Tahririan, 2006), and that students notice more negotiation episodes when communicating via text chat than face-to-face (Lai & Zhao, 2006; Yuksel & Inan, 2014). Yilmaz and Yuksel (2011) found that learners benefited more from receiving recasts through text chat than orally, suggesting that this online medium may facilitate the noticing of recasts. In addition, in terms of corrections, Lai and Zhao (2006) detected significantly more self-corrections during online chat compared to spoken interaction, indicating that online chat may facilitate the noticing of errors in one's own output.

Some research, however, has found that text chat may not engender attention to form. The findings of these studies show that, when compared to oral communication, there are fewer instances of negotiation of meaning and a lower amount of corrective feedback in text-based SCMC (Iwasaki & Oliver, 2003; Rouhshad et al., 2016), indicating fewer

opportunities to attend to linguistic form in online chat. Gurzynski-Weiss and Baralt's (2014) study did not find any significant difference in terms of noticing of L2 feedback between online chat and oral communication. Additionally, Lai and Zhao (2006) observed scant evidence of noticing of recasts in text chat. As mentioned in the previous section, these negative findings might be explained by the relatively rapid nature of online chat interaction, disrupted turn adjacency and the cognitive burden imposed by typing, which may hinder attention to form during chat. Interestingly, Baralt, Gurzynski-Weiss, and Kim (2016), who observed less social, affective and cognitive engagement in L2 text chat than in face-to-face interaction, hypothesize that the reduced cognitive engagement could be caused by low social and affective engagement. In other words, learners may attend less to L2 linguistic form during online chat because text chat interaction is less supportive than face-to-face interaction and online chat might lack a sense of trust and feelings of friendship compared to face-to-face communication. To increase attention to form during text-based SCMC, they suggest that learners' sense of friendship, support and belonging to a chat community should be promoted.

Thus far, an overview of research on attention to form in L2 online chat has been presented. The section that follows will conduct a detailed review of studies carried out to investigate L2 revision during online chat.

3.5 Empirical studies on L2 revision during text-based SCMC

Only a handful of studies have been conducted to explore L2 revision during online chat. Although the present study conceptualizes the changes made during text chat as revisions, previous works investigating these changes have explored them within the narrow scope of self-corrections (Fredriksson, 2015; Golonka, Tare, & Bonilla, 2017; Jepson, 2005; Lai & Zhao, 2006; Lee, 2002; Zeng, 2017) or in terms of self-repairs (Smith, 2008, 2009; Yuan, 2003); the latter are measures commonly used for assessing language adjustments in speech research. One study (Sauro & Smith, 2010), instead of examining corrections or repairs, compared text before and after changes. To date, only two types of data sources have been used in these studies: chat logs and computer screen recordings. In the review below, I will discuss both types of research, beginning with those studies relying only on chat logs

(Fredriksson, 2015; Golonka et al., 2017; Jepson, 2005; Lee, 2002; Yuan, 2003; Zeng, 2017), then move on to those that utilized recordings (Lai & Zhao, 2006; Sauro & Smith, 2010; Smith, 2008, 2009).

Lee (2002) explored self-corrections, both self- and other-initiated, in online chat based on the chat scripts of 34 intermediate learners of Spanish, conceptualizing self-correction as one of the modification devices (i.e. strategies used to adjust incomprehensible messages). The results revealed that self-correction was one of the three devices used most frequently during online chat discussions. While participants tended to ignore their interlocutors' language errors, they frequently corrected their own linguistic mistakes. The researcher hypothesized that the high number of self-corrections observed could be attributed to the increased salience of language output in this medium.

In another study, Yuan (2003) conducted a small-scale study to explore the *self-repairs* of errors (i.e. self-initiated error corrections) of two adult learners of English in online chat. In this study, participants engaged in an online chat session with the researcher every week for 10 weeks and had a face-to-face discussion with the researcher about the linguistic problems in their chat scripts after each session. Examining the participants' chat logs, the researcher found that some of the errors participants corrected during online chat were those discussed in previous face-to-face meetings. These findings suggest that learners do attend to linguistic form while chatting, and that instruction may direct learners' attention towards particular linguistic features. In addition, analyzing the correction rate of each type of error, Yuan found that students tended to correct errors that affected message comprehensibility (e.g. errors relating to word selection and word form) more than those that did not affect meaning (e.g. errors relating to articles and subject-verb agreement). Thus, Yuan concluded that meaning-related errors may be more salient and more prone to being corrected during text chat than errors not relating to meaning.

Although there is evidence showing that students notice and correct their own errors in the above online chat studies, the study conducted by Jepson (2005) yielded different findings. Jepson's research explored other-initiated self-corrections, or what he called

incorporations, and self-initiated corrections as measures of repair moves, operationalized as the negotiation of meaning and negative feedback. Data were gathered from five chat room sessions in which anonymous L2 learners of English participated. Surprisingly, self-initiated corrections were absent from the data. Jepson explained that the online chat environment probably does not encourage noticing of one's own errors and self-initiated corrections. However, in this research, these corrections were examined only from chat scripts, which do not provide information about the errors learners correct while drafting their messages.

In another study by Fredriksson (2015), self-corrections were investigated in the chat scripts of group discussions of learners with different L1 backgrounds and L2 proficiency levels. Participants were 30 students on a German literature course who were asked to attend four online chat seminars guided by a list of open-ended questions relating to German literature. Self-corrections, operationalized as students' modifications of their output, were identified and classified into different types: grammar, lexis, orthography and content. The results showed that corrections of orthographical and grammatical mistakes occurred more frequently than those focusing on lexis and content. Furthermore, in line with Lee's (2002) findings, it was found that language problems were rarely negotiated by the interlocutors, instead, L2 modifications tended to be self-initiated.

Recently, Golonka et al. (2017) examined the chat logs of peer-peer interactions of 25 intermediate learners of Russian for the types of behaviours that might have led them to improve their language skills. The behaviours investigated included self-initiated corrections and corrections of partner's errors. The raw frequencies of different types of corrections were examined (e.g. lexical, grammatical and spelling/ typographical errors). The results revealed that, of all the behaviours examined, self- and partner-corrections and positive affect (i.e. the provision of encouragement) were the behaviours most frequently found in the data. Golonka et al. (2017) hypothesized that the visual nature of chat may make language errors noticeable and promote corrections. In line with Fredriksson's (2015) and Lee's (2002) findings, there were more instances of self-corrections compared to corrections of partner's errors. Most self-corrections were related to spelling and

typographical errors, and grammatical corrections occurred about five times more frequently than lexical ones. Regarding the outcome of corrections, 47 per cent (N = 13) of the self-corrections carried out to rectify grammatical errors (N = 27) and all self-corrections targeting lexical errors (N = 4) were successful.

In another recent study, Zeng (2017) compared self-initiated correction in text chat and face-to-face communication by operationalizing it as a type of LRE. Chat logs and transcripts of a jigsaw and a dictogloss task performed by 32 Chinese learners of English were examined. The results of this study showed that the number of self-corrections in chat logs was more than twice as high as that found in the transcripts of face-to-face interaction; and the percentage of self-corrections in the logs was significantly higher than that in the transcripts. Based on these findings, Zeng suggests that the immediate visibility of L2 chat output could promote language reflection and self-corrections, and the absence of the interlocutor's presence in this mode may allow learners to focus their attention on text and linguistic form.

All the studies reviewed thus far used chat scripts as the only source of data; however, other L2 online chat revision studies have also examined revision by looking at participants' computer screen recordings. One such study is Lai and Zhao's (2006) research, which compared self-corrections in text chat with those in face-to-face communication. In this study, 12 intermediate learners of English carried out two spot-the-difference tasks, one via online chat and the other face-to-face. The performance data gathered from recordings of oral interaction and chat screens were coded for episodes of self-initiated error corrections. The results of this study revealed that there were significantly more self-corrections during online chat than in face-to-face communication. Moreover, as many as 10 out of 12 participants corrected their linguistic mistakes more often in the online mode. Overall, these findings indicate that text-based SCMC might be a better alternative for fostering learners' attention to their own linguistic errors than face-to-face communication.

In another study that utilized screen recording technology, Smith (2008) investigated the nature of self-initiated CMCovert repairs in text chat by examining the recordings of jigsaw

picture-ordering tasks performed by 46 high beginners of German. As mentioned in Section 3.2, CMCovert repair is a type of self-repair unique to text chat. It is overt from the perspective of composers but remains covert from interlocutors, and it can be investigated through screen recordings. Both chat scripts and computer screen recordings were analyzed for self-repairs based on van Hest's (1996) self-repair taxonomy, which has been commonly utilized in speech research. Self-repairs targeting linguistic errors, appropriacy of language and content changes were coded as error, appropriateness and different-information repairs, respectively; the standardized self-repair scores of each participant were calculated by adjusting for output quantity. The results of this study showed that CMCovert repairs occurred frequently, and the number of self-repairs found in the recordings was significantly higher than that found in the chat scripts, indicating the benefit of collecting screen capture data. Comparing the count for each type of repair, Smith found that error repairs occurred most frequently, followed by appropriateness repairs, and the difference between these two types of repairs was significant. Different-information repairs, however, were almost non-existent, with only one instance of this type of repair found in the screen recordings. Thus, the findings seem to suggest that text chat may foster learners' attention to linguistic form at the expense of attention to content. Alternatively, it is possible that online chat allows enough time for learners to conceptualize and formulate messages that can convey their intended meaning in an appropriate manner before/ during transcription; and therefore there is less need for appropriateness and different-information repairs. Furthermore, the study found that, out of all the error corrections, there was a higher percentage of corrections targeting grammar (63 per cent) than those targeting lexis (33 per cent). Even though this difference was not statistically significant, it suggests that learners might attend to grammatical errors more than lexical errors in this medium. Smith explains that grammatical error repairs may occur frequently due to the written nature of the output in chat, which potentially makes minor morphological errors salient to learners.

In another study, Smith (2009) examined the self-initiated CMCovert repairs of 33 high beginners of German during jigsaw tasks, and the correlation between these repairs and the amount of scrolling. Based on Lai and Zhao's (2006) hypothesis, which speculates that the ability to scroll up and down a chat conversation page may encourage noticing, Smith

predicted that the amount of scrolling would correlate positively with the number of self-repairs. As in his other study (Smith, 2008), self-repairs were analyzed from screen recordings based on van Hest's (1996) self-repair taxonomy. In line with the findings of Smith (2008), the results of this study revealed a high number of self-repairs and showed that error repairs occurred most frequently among all types of self-repairs examined (i.e. error, appropriateness and different-information repairs). This indicates that learners usually repair their output in order to rectify their linguistic mistakes. Repairs targeting grammar occurred four times more frequently than those focusing on lexis. In addition, inconsistent with Smith's prediction, the correlation between scrolling and self-repairs was not significant, indicating that these two variables may be independent of each other.

Finally, Sauro and Smith (2010) explored whether students used the online planning time afforded by text chat to attend to the L2 output they produced during a jigsaw picture-ordering task and made their language more complex. Unlike the studies reviewed thus far, which examined self-corrections and self-repairs, this research focused on comparing text before and after changes. The computer screens of 23 learners of German were recorded and the data were coded for *pristine text*, *deleted text*, *post-deleted text*, *post-deleted deleted text* and *pre-deleted text*. Pristine text refers to text which is written from beginning to end without any deletion or correction. Deleted text is text which is deleted and, thus, does not appear in the chat script. Post-deleted text is text that is typed immediately after corrections or deletions. Post-deleted deleted text is post-deleted text which is subsequently deleted. Lastly, pre-deleted text refers to all text written before a correction or a deletion occurs in that turn. Sauro and Smith gauged complexity in terms of lexical diversity (Index of Guiraud) and language complexity (ratio of clauses to c-units and productive use of grammatical gender) and found post-deleted text to be significantly more complex and lexically diverse than both pristine text and deleted text. Their results showed that language in text chat became more sophisticated after revision and learners actually used the opportunity for online planning to monitor their L2 output.

In summary, the findings of previous L2 online chat revision research show that learners' output becomes more complex after revision (Sauro & Smith, 2010) and learners tend to

correct their own errors more than the errors of their interlocutors while chatting (Fredriksson, 2015; Golonka et al., 2017; Lee, 2002). Compared to face-to-face interaction, learners make significantly more self-corrections in this medium (Lai & Zhao, 2006; Zeng, 2017). With regard to the types of error corrections, learners may be more prone to correct errors that affect the comprehensibility of their messages more than those that do not (Yuan, 2003), and they attend more to orthographical, typographical and grammatical errors than to errors associated with lexis or content (Fredriksson, 2015; Golonka et al., 2017; Smith, 2008, 2009). Indeed, the dominance of self-corrections targeting grammar over those targeting lexis found in online chat is the opposite of the pattern observed in most oral communication studies (Fathman, 1980; Lennon, 1984). This indicates that the nature of text chat may increase the salience of grammatical features. As regards different types of self-repairs, previous research by Smith (2008, 2009) found that error repairs were carried out more frequently than appropriateness repairs, and different-information repairs scarcely occurred during online chat. In general, these findings seem to lend support to the claim that text chat could be a useful medium for promoting attention to form.

Despite the generally positive findings, the studies reviewed in this section suffer from several important shortcomings. First, to date, only a few studies have collected data on participants' screen recordings when investigating language revision in text chat (Lai & Zhao, 2006; Sauro & Smith, 2010; Smith, 2008, 2009). Other studies, however, only analyzed learners' language corrections through chat logs, which do not provide any information about learners' drafting processes. As revealed in Smith's (2008) study, the number of repairs found in computer screen recordings was significantly higher than that found in chat scripts. Thus, there is a clear benefit in collecting participants' computer screen data. In his publication, O'Rourke (2008) critiques the use of chat logs as an exclusive data source by pointing out that they only offer a third-person perspective on what is going on while chatting. He proposes various data collection methods that could yield richer information about chat users' experience (i.e. information about the drafting process, interactional tempo, users' focus of attention, their paralinguistic behaviours). The proposed methods include keystroke logging, computer screen capture, eye-tracking and video recording of chat users during chat sessions. Many of these methods may prove useful

if utilized in L2 online chat revision research. Keystroke and screen capture data can provide information about what learners type or delete during the drafting process, and eye-tracking technology can reveal their eye gaze, which indicates what they are paying attention to when making changes to their language (O'Rourke, 2008). However, although a variety of data collection methods can be used, previous L2 online chat revision research has mainly relied on data from chat logs.

The second major limitation of these studies is that self-corrections and self-repairs were analyzed and classified based on the researchers' own judgement. None of the L2 online chat studies investigating self-corrections and self-repairs have taken learners' self-report into account. Regarding her suggestion as to how self-repairs should be classified in speech research, Kormos (1998a, 1998b, 1999a) has pointed out that researchers should not only consider the surface structures of self-repairs, but also explore learners' underlying communicative intentions to ensure that the coding is reliable. I support this suggestion being applied to online chat research and verbal protocols being utilized in L2 online chat studies to ensure the accurate classification of error corrections and self-repairs. Without learners' verbalization, it can be difficult for researchers to gauge what prompts them to revise their text, and the classification of error corrections and self-repairs can be unreliable. For instance, on the surface, the deletion of 's' in 'I knows' might appear to be a correction targeting a grammatical error. However, it may just be a correction of a typographical mistake. The appropriateness repairs explored in Smith's (2008, 2009) studies are also difficult to identify based on surface data because the judgements made on what is or what is not appropriate are subjective. Indeed, verbal protocols such as retrospective stimulated recall and think-alouds have been commonly used in L2 research exploring self-repairs during speech and revisions during writing. Therefore, it is somewhat surprising that no L2 online chat studies to date have utilized any of these protocols to investigate L2 adjustments in this medium.

Third, although previous research findings have revealed the predominance of grammatical and orthographic corrections in L2 text chat, these findings should not be taken as conclusive because only a few L2 online chat revision studies (Fredriksson, 2015; Golonka

et al., 2017; Smith, 2008, 2009; Yuan, 2003) have classified error corrections into different types. Other studies only explored the total number of corrections. Furthermore, most of the research that compared different types of corrections (Fredriksson, 2015; Golonka et al., 2017; Smith, 2008, 2009) merely investigated the frequency or proportion of each type of error correction, rather than the correction rate of each type of error. When different types of corrections are compared using the frequency or proportion of corrections, the results can be influenced by the number of errors made in each category. For instance, when learners make more lexical errors than grammatical ones, it is likely that a higher frequency/ proportion of lexical corrections will be observed in comparison to grammatical corrections. The utilization of error correction rates can overcome this problem. An error correction rate is a rate at which writers or speakers attempt to rectify the errors they made. It can be calculated as the ratio of errors receiving corrections to total errors made. By exploring a correction rate for each type of error, researchers would gain better insights into learners' attention to each type of error during chat sessions. Even though one study has examined error correction rates (Yuan, 2003), it only explored them based on the chat scripts of two participants and errors were classified into highly specific categories (e.g. subject-verb agreement, transition and preposition) rather than more comprehensive ones (e.g. grammar and lexis).

Lastly, the previous research reviewed in this section is limited in the scope of its investigation. Most of these studies merely explored L2 corrections and self-repairs in online chat without investigating how other factors might influence them. The only exception to this is Smith's (2009) study, which examined the relationship between scrolling and self-repairs. It could be that there are other variables that affect corrections and repairs in this medium, such as proficiency and typing skills. Regarding proficiency, studies that compared different types of self-repairs in L2 online chat to date (Smith, 2008, 2009) only recruited participants with low proficiency levels (beginners). This probably explains why these studies found a predominance of error repairs over appropriateness and different-information repairs. Based on the findings of research on self-repairs during speech production (e.g. Kormos, 2002; O'Connor, 1988), Kormos (2006) points out that there could be a shift in attention from lower-level errors (i.e. lexical, grammatical and

phonological errors) to those at the discourse level when the degree of language automatization increases. In line with this, as discussed in the previous chapter, L2 writers with higher proficiency are found to revise more at high levels (such as content, style and discourse), while those of lower L2 proficiency tend to revise language form and typography (Barkaoui, 2016; Choi, 2007; Manchón et al., 2009). Since an increase in proficiency may allow learners to attend more to the global level of L2 (e.g. content and discourse), the occurrences of appropriateness and different-information repairs observed in Smith's (2008, 2009) studies could have been higher if he had examined the data of intermediate or advanced learners. In terms of typing ability, as mentioned in Chapter 2, it has been hypothesized that poor motor skills can increase the cognitive load on the writer's working memory (Kellogg, 1996). Based on this hypothesis, it is possible that chat users with lower typing ability may correct or repair their output less due to the increased cognitive burden imposed by their non-automatized typing process. Although there have been some calls for the exploration of keyboarding and typing ability in L2 online chat research (Lai & Zhao, 2006; Sauro & Smith, 2010), to the best of my knowledge, no L2 chat studies to date have gauged participants' typing ability.

3.6 Summary

Text-based SCMC has been commonly viewed as the integration of oral and written communication because its properties resemble those of speaking and writing. Due to this, learners may benefit from opportunities for L2 interaction as in oral communication and for intensive linguistic processing, e.g. noticing of and reflection on linguistic form, as in writing when engaging in this hybrid mode of communication. In addition to this, the potential of text-based SCMC for SLA also derives from its unique characteristics and affordances. For instance, the reduced paralinguistic cues and the relatively longer time for information processing have been hypothesized to create a less stressful environment for learners compared to face-to-face interaction. The opportunities afforded by this medium to communicate authentically outside class without the constraints of time or space also foster learner autonomy and promote interaction, which could be beneficial for language learning. Moreover, an additional language monitoring stage in chat, which allows learners

to edit their messages before sending them to interlocutors, provides students with time to reflect on their L2 and attend to errors in their output.

In general, the most important claims made regarding the advantages of online chat for SLA are related to attention to form. Many characteristics and affordances of this medium (i.e. its salient and written nature, re-readability and permanent records of chat conversations, and increased online planning time) have been assumed to foster students' attention to language form. Despite this, some doubts have been cast on this potential. In particular, it has been hypothesized that the high cognitive demands during chat, the disruption of turn adjacency and the relatively rapid pace of chat interaction may hinder attention to form. Overall, previous research findings regarding attention to form in this online mode have been mixed.

In terms of research on L2 online chat revision, to date, only a small number of studies have been carried out to explore the changes made during L2 text-based SCMC. These changes have been examined as self-corrections (Fredriksson, 2015; Golonka et al., 2017; Jepson, 2005; Lai & Zhao, 2006; Lee, 2002; Zeng, 2017) or self-repairs (Smith, 2008, 2009; Yuan, 2003), the latter being common measures of language adjustments in speech research. These studies have various limitations concerning data sources, classification methods of corrections and repairs, research measures and the scope of investigation. For this reason, the findings obtained from these studies should not be treated as conclusive, and many questions regarding L2 online chat revision remain to be explored.

In this chapter, I have provided an overview of text-based SCMC. The next chapter will review the literature relating to tasks in SLA.

Chapter 4: Tasks for L2 instruction

Pedagogic tasks have been widely researched in SLA and used in L2 classrooms to provide learners with opportunities for language learning. In this chapter, I will begin by describing the definition of task in SLA and the utilization of tasks in L2 pedagogy. As the current work explores attention to form during online chat tasks through revisions and the influence of post-task anticipation on this attention, the section that follows will discuss the proposals made by researchers in the field of task-based language teaching regarding the importance of attention to form. Moreover, I will briefly outline how this attention can be promoted when teaching L2 through tasks. At the end of the chapter, a review of task implementation methodologies designed to foster attention to form in different task stages will be presented, followed by a review of empirical studies on post-task anticipation.

4.1 Definition of task in SLA

Tasks have been utilized as tools for L2 pedagogy, research and assessment, and various definitions of task have been proposed based on the dimension of task central to each work. These definitions range from the general to the specific. Long (1985), for example, offers a broad definition, defining a task as:

> …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, filling out a form … by "task" is meant the hundred and one things people do in everyday life, at work, at play, and in between. (p. 89)

According to Long, tasks are everyday activities that are not necessarily associated with language. However, a more specific definition was proposed by Skehan (1998), who defines a task as:

> …an activity in which: meaning is primary; there is some communication problem to solve; there is some sort of relationship to comparable real-world activities; task completion has some priority; the assessment of the task is in terms of outcome. (p. 95)

In this definition, Skehan emphasizes the authenticity of tasks and that tasks need to involve some kind of communication and require a focus on meaning. In addition, this definition includes pedagogical aspects, suggesting that the assessment of task should be outcome-based. Other definitions of task include:

> …a piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form. The task should also have a sense of completeness, being able to stand alone as a communicative act in its own right. (Nunan, 1989, p. 10)

> …activities where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome. (Willis, 1996, p. 23)

> …one of a set of differentiated, sequencable, problem-posing activities involving learners and teachers in some joint selection from a range of varied cognitive and communicative procedures applied to existing and new knowledge in the collective exploration and pursuance of foreseen or emergent goals within a social milieu. (Candlin, 1987, p. 10)

> …an activity which required learners to arrive at an outcome from given information through some process of thought and which allowed teachers to control and regulate that process… (Prabhu, 1987, p. 24)

> …a holistic activity which engages language use in order to achieve some non-linguistic outcome while meeting a linguistic challenge, with the overall aim of promoting language learning, through process or product or both. (Samuda & Bygate, 2008, p. 69)

It can be seen that, similar to Skehan (1998), Nunan (1989) also suggests that tasks are meaning-oriented activities, but Nunan further specifies the sense of completeness of tasks. Willis (1996), Candlin (1987), Prabhu (1987) and Samuda and Bygate (2008) all mention

the goal of tasks. While Prabhu (1987) highlights that the outcome of tasks can be reached through information processing, Willis (1996) emphasizes the importance of communication for achieving the task outcome. Similar to Willis (1996), Candlin (1987) stresses "collective exploration" (p. 10) and the pursuance of goals within a social environment, and further emphasizes that tasks must be sequenceable and involve some problems for learners to work through. With respect to Samuda and Bygate's (2008) task definition, they view tasks as holistic activities involving the use of various aspects of language, such as grammar, phonology and vocabulary, and that the ultimate goal of tasks is to foster language learning.

Overall, these examples illustrate that, despite focusing on different aspects of tasks, their definitions can overlap. In relation to this, Ellis (2009) reviewed various proposals of task definition and summarized the criteria for task features. He states that, to be considered a task, a language-teaching activity must satisfy all of the following criteria:

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

According to Ellis (2009), tasks must be meaning-oriented and learners are required to focus on pragmatic and semantic meaning. However, this does not mean that a shift in attention from meaning to form is not possible during task performance. In addition, based on Ellis' criteria, tasks should have certain gaps that learners must fill, for instance through

the exchange of information or inference of meaning. Tasks should also require learners to utilize their own linguistic and non-linguistic resources. Lastly, Ellis (2009) notes that language is only a means to accomplish tasks and the outcome of tasks should go beyond the use of language, such as a decision made based on a discussion or a picture drawn based on instructions.

Thus far, in this section, I have discussed the definition of task in SLA. The next section will focus on the utilization of tasks in L2 pedagogy. It begins by describing the emergence of task-based language teaching, then proceeds to a summary of the different roles of tasks for L2 teaching.

4.2 Tasks in L2 pedagogy

Tasks provide an authentic and meaningful context for language learning and their benefits for L2 learning are supported by SLA theories. As mentioned in Chapter 3, Section 3.3.1, the interaction hypothesis (Long, 1996) and sociocultural theory (Lantolf & Thorne, 2007) assume that interaction is crucial for L2 learning. Based on these perspectives, the opportunities for L2 interaction during tasks can be beneficial for L2 learners. The importance of tasks for L2 pedagogy has been realized by educators since the 1980s, owing to the development of task-based language teaching (TBLT) (Van den Branden, Bygate, & Norris, 2009).

TBLT is a type of instructional method that developed out of concerns over the teacher-centred and form-focused approach that had dominated language education. The teacher-centred approach usually entails a Present-Practice-Produce (PPP) methodology, in which the teacher presents each language unit (e.g. a grammatical structure or a set of lexical items) separately to learners, lets them practise each of the units in isolation, and optionally incorporates practice of more functional language use in the classroom (Van den Branden et al., 2009). It assumes a synthetic approach to language teaching in which learners are expected to be able to acquire the target language by gradually accumulating knowledge of each L2 element (Long, 2015). Synthetic syllabi are form-focused and pre-selected linguistic items are presented to learners in a linear sequence based on, for instance, how

difficult, frequent and communicatively important each item is (Robinson, 2001a). However, this approach has been criticized because it does not take into account learners' developmental readiness. For instance, Long (2015) observes that a predetermined linguistic syllabus requires all learners in the class to learn the same language elements, which some of them may not be ready to acquire yet. In addition, as linguistic elements are presented to learners in a linear fashion, this synthetic approach does not correspond to the process of L2 development, which has been found to be non-linear (Long, 2015). Learners' interlanguage could backslide into earlier stages or students might forget about the new L2 items introduced, leading to a U-shaped learning curve rather than an ascending pattern of development (Kellerman, 1985; Larsen-Freeman, 1991).

Realizing the shortcomings of a synthetic approach, educators proposed communicative language teaching as an alternative. Prabhu (1987) was the first to implement TBLT on a large scale by using what he called a *procedural syllabus* consisting of a series of problem-solving and meaning-oriented pedagogic tasks in his five-year project in India. Task-based language teaching (TBLT) was developed as a learner-based method of teaching, placing communicative tasks at the heart of L2 classrooms (Van den Branden et al., 2009). Learners in TBLT classrooms are required to complete pedagogic tasks that resemble the tasks they may encounter in the real world, and they are encouraged to interact with their peers during task performance. In the strong version of TBLT, tasks form the basis of L2 syllabi, classroom activities and assessment (Long, 1985, 2015; Samuda & Bygate, 2008), and syllabi are created by means of task sequencing without any predetermined language focus (Long & Crookes, 1992). For instance, it has been suggested that task sequencing should consider the cognitive complexity of tasks (Long & Crookes, 1992; Robinson, 2001a, 2010, 2011, 2015). As lessons progress, the increased demands of tasks will approximate to the demands of real-world tasks more and more, thereby preparing learners to handle real-life situations (Long, 2015).

In essence, TBLT adopts an analytic approach which assumes that acquisition results from learners' attention to and analysis of language during meaningful communicative activities (Long & Norris, 2000). It has been claimed that, unlike the synthetic approach, this

approach enables students to learn based on their level of language proficiency and the current stage of their L2 development (Long, 1998, 2015; Skehan, 1998). In addition, the rationale of TBLT corresponds to a philosophy of education that emphasizes the significance of learning by doing and learner-centredness (Long, 2015; Ziegler, 2016). By providing learners with opportunities to engage in pedagogic tasks relevant to the tasks they need to perform in real life, TBLT prepares them for the communicative demands of real-world activities, while learning the target language in the process (Long, 2015). According to Ellis (2009), a task-based approach is suitable for teaching an L2 in an "acquisition-poor" environment (p. 237), where learners lack opportunities to communicate in the target language outside classes, because it provides them with such opportunities.

Apart from TBLT, there are also other approaches to the utilization of tasks for language teaching. Regarding this, Samuda and Bygate (2008) distinguish three approaches: task-referenced, task-supported and task-based language teaching. First, in the task-referenced approach, the aim is to develop learners' abilities to achieve task outcomes. While tasks determine the end goal of task-referenced instruction, they are not necessarily used during lessons. The second approach is task-supported instruction, which uses tasks to complement a form-focused syllabus. In this type of instruction, tasks are usually designed to elicit certain target elements in order to offer learners opportunities to practise them. This approach can be considered a weak version of TBLT – a teaching approach in which tasks are utilized in classrooms but do not form the basis of syllabus design (Skehan, 1996, 2003b). As mentioned earlier, the strong version of TBLT assumes the importance of tasks as the driving force of curriculum design. This type of instruction is the last approach outlined by Samuda and Bygate (2008). Within this approach, tasks are used to form L2 syllabi and classroom activities, and L2 assessment is based on learners' task performance and whether they achieve task outcomes.

In this section, I have discussed the use of tasks in L2 pedagogy. The following section will focus on the assumptions of task researchers regarding the importance of attention to form. In addition, a brief overview of how this attention can be fostered when using tasks for L2 instruction will be presented at the end of the section.

4.3 Attention to form and L2 task pedagogy

As mentioned in the previous section, tasks are used in task-supported and task-based language classrooms. In task-supported classes, they are administered to provide opportunities to practise L2 features that are the target of form-focused syllabi. Even though TBLT syllabi are not based on predetermined language elements, this does not mean that form is ignored by advocates of TBLT. In his response to criticism of TBLT, Ellis (2009) emphasizes that TBLT researchers perceive attention to form as an essential, rather than an optional, element of task-based instruction. Indeed, many researchers have highlighted the importance of attention to form in TBLT (e.g. Ellis et al., 2001; Long, 2015; Skehan, 1996, 1998). In this section, I will discuss two important proposals relating to the significance of attention to form when learning and teaching L2 through tasks.

First, Long (1991, 2015) points out the benefits of a *focus on form* during task-based instruction. Focus on form refers to a reactive and incidental shift of attention from meaning to form during meaning-oriented tasks (Long, 1991). Long believes that the shift of attention from content to form occurring when communication breakdowns arise during tasks, or when learners receive recasts, is crucial for interlanguage development. According to him, a focus on form promotes form-meaning mappings by providing learners with opportunities to attend to problematic L2 elements in a meaning-oriented context, and it allows them to learn based on their internal syllabus. Lending support to Long, Ellis et al. (2001) note that a focus on form is "psycholinguistically plausible" (p. 410) because it triggers the type of attention to form that occurs during natural language learning; it addresses the language problems that learners actually have; and it promotes noticing, which may facilitate acquisition.

Second, in his work on TBLT, Skehan (1996, 1998) suggests that during tasks, attention should not be devoted to meaning alone, but attention to form is important for achieving balanced development of different L2 performance goals – language complexity, accuracy and fluency. His hypothesis is based on VanPatten's (1994) theory, which assumes that meaning and form compete for learners' attention and that meaning is primary during input

processing. According to Skehan (1996, 1998), there is a trade-off between three different aspects of task performance – fluency, complexity and accuracy – and learners simply tend to focus on fluency (i.e. meaning) rather than try to produce accurate or complex L2 output. As tasks can sometimes be accomplished without accurate or complex use of language, learners may simply resort to using communication strategies to achieve task goals (Skehan, 1996). Eventually, these coping strategies may be proceduralized to the point that they become inflexible to change and impede L2 learning (Skehan, 1996). Based on these assumptions, Skehan (1996, 1998) recommends that a focus on form must be engineered in task-based instruction to draw learners' attention from meaning to form, and promote balanced development of accuracy, complexity and fluency.

To date, there have been various suggestions as to how attention to form can be fostered when using tasks for L2 instruction. For instance, it has been assumed that task selection and instructional methods influence how learners allocate their attention during task performance. With respect to task selection, Robinson (2001a, 2001b, 2005, 2011) proposes that the increased demands of tasks in some areas may lead to increased attention to form. He makes a distinction between tasks that are complex along the *resource-directing* dimension and the *resource-dispersing* dimension, proposing that the first will direct learners' attention to specific linguistic codes while the latter will disperse their attention across various non-linguistic aspects. For instance, it is speculated that tasks that are made complex by requiring learners to refer to events occurring in distant time and space could direct their attention to developmentally more advanced tense and aspect markings; meanwhile, complex tasks with limited planning time disperse students' attention across many non-linguistic and linguistic aspects of task, instead of promoting the noticing of certain L2 linguistic forms (Robinson, 2011).

Another proposal relating to the influence of task selection on attention to form was put forth by Skehan (1996, 1998). Skehan (1996, 1998) hypothesizes that learners may have fewer attentional resources available to attend to form when performing tasks that are cognitively more demanding, requiring increased attention to communicate meaning, such as those requiring complex thinking processes. Furthermore, in his more recent work,

Skehan (2014) suggests that, in order to maximize the opportunities for attention to form, educators can manipulate learners' attention by considering the influences of tasks on learners' psycholinguistic processes, namely conceptualization, lemma retrieval and syntactic encoding. Skehan (2009) proposes four types of task influences: *complexifying*, *pressuring*, *easing* and *focusing*. The complexifying influence is related to increased cognitive demands on the conceptualizer; the pressuring effect involves increased pressure on the formulator and it may hinder parallel processing; an easing influence occurs when a task does not place a high cognitive load on the conceptualizer or the formulator, and therefore it supports parallel processing; and the focusing effect is associated with how a task may motivate learners to focus on language accuracy. Although Skehan (2014) does not explicitly discuss how task influences can be manipulated to engineer attention to form, it can be assumed that the use of tasks that exert a focusing influence, e.g. tasks for which an accurate language use is important, could be useful for promoting attention to form.

In addition to Robinson's (2001a, 2001b, 2005, 2011) and Skehan's (1996, 1998, 2009, 2014) proposals, Loschky and Bley-Vroman (1993) suggest that the development of and attention to certain grammatical structures in task-based instruction can be enhanced by selecting tasks based on the degree to which those structures are involved in each task. They point out that a structure can be associated with a task to varying extents, and describe three kinds of relationships between grammatical structure and task: *task-naturalness*, *task-utility* and *task-essentialness*. The first refers to when a certain structure may naturally occur in a certain task, although the task can be accomplished easily without it. Task utility describes a situation in which the use of a certain structure makes it easier to complete the task. Task essentialness is a relationship in which the use of a structure is unavoidable when performing a task, and the task cannot be performed successfully without it. Loschky and Bley-Vroman (1993) recommend that grammatical structures be made essential to tasks because, when a language structure is essential for task performance, students' attention will be directed to that structure and this heightened attention can promote hypothesis formation and restructuring.

Apart from task selection, researchers have also highlighted how instructional methods potentially promote attention to form. These methods include, for instance, the provision of recasts during tasks and the implementation of different task stages. In terms of recasts, it has been claimed that recasts can foster a focus on form as the juxtaposition of learners' own erroneous output and the feedback they receive may facilitate noticing the mismatch between the two (Doughty, 2001). Furthermore, during recasts, learners already know the meaning they want to convey. Therefore, the cognitive load is reduced, allowing them to devote their attention to the form which is the target of feedback (Long, 2015). In the section that follows, I will discuss the potential of another instructional method – the implementation of different task stages – for raising learners' attention to form.

4.4 Attention to form during different task stages

To date, there have been many proposals for the incorporation of multiple task stages in L2 lessons (e.g. Lee, 2000; Nunan, 2004; Prabhu, 1987; Skehan, 1996, 1998; Willis, 1996). These proposals usually involve a pre-task, a main-task and a post-task stage (Ellis, 2006). In this section, I will review two widely-cited designs of task implementation methodology (Skehan, 1996, 1998; Willis, 1996) and discuss how attention to form may occur during different task stages.

In 1996, Willis proposed a detailed TBLT framework consisting of three task components: 1) pre-task, 2) task cycle and 3) language focus. This framework is holistic in the sense that it focuses on both meaning and form. According to Willis, the pre-task stage aims to prepare learners for the main task by introducing them to the task topic and lexical items related to it. The next stage, task cycle, comprises three parts: task, planning and report. The cycle begins with task performance, which, according to Willis, gives learners the opportunity to boost their L2 fluency and confidence in using the target language. As learners are not required to perform the task publicly, accuracy is not the primary concern at this stage (Willis, 1996). While the task focuses on meaning and fluency, subsequent parts – planning and report – are implemented to draw learners' attention to form. Planning involves the preparation of a public report, e.g. an oral report made in front of the class or a written report exchanged between groups of students. During planning, learners can experiment

with their L2, check on their use of grammar and seek assistance from the teacher when they have any linguistic problems (Willis, 1996). Because the report is public, learners may be concerned about language form during planning and reporting. According to Willis (1996), the report stage is a natural stimulus that pushes learners to improve their language, and this motivates them to think about accuracy as well as fluency and to use their complex version – or "prestige version" (p. 64) – of the L2. Furthermore, learners have an additional opportunity to receive feedback from the teacher following the report (Willis, 1996), which can trigger noticing of their own linguistic errors. Finally, the last stage within Willis' (1996) framework is language focus, which aims to increase learners' knowledge of the linguistic form that occurs naturally during the task cycle. Language focus consists of two components: analysis and practice. During this stage, learners perform consciousness-raising activities that allow them to analyze L2, for instance in a sample text or a transcript, and they make discoveries related to the target features by themselves. In addition, form-focused practice might be included to ensure the consolidation of L2 knowledge and acquisition. Essentially, within Willis' framework, attention to form may occur in various stages – the planning, the report and the language focus phase – and learners are allowed to work by themselves rather than always being led by the teacher to attend to a particular form.

Apart from Willis (1996), Skehan (1996, 1998) also put forth his design for task methodology by assuming that attention to form in task-based classrooms can be engineered though task selection, task sequencing and task implementation. With respect to task implementation, following Willis' (1996) framework, Skehan (1996, 1998) proposed a three-stage task implementation methodology, which comprises a pre-task, a task and a post-task stage. According to Skehan (1998), a pre-task activity may be in the form of teaching, a consciousness-raising activity or pre-task planning. Through teaching, teachers can introduce learners to new L2 elements and increase the chance of L2 restructuring (Skehan, 1998). Consciousness-raising activities (e.g. identifying certain L2 aspects in a text, reading materials with some L2 elements highlighted, pre-task brainstorming) involve increasing learners' awareness of the language structure or information relevant to the main task. By utilizing this type of activity, teachers can

promote the restructuring of L2 (Skehan, 1998). In addition, consciousness-raising activities can familiarize students with the main task, leading to a diminished cognitive load during the main task stage (Skehan, 1998). Implicated in this hypothesis is the notion that, when this happens, more attentional resources will be available to devote to form during the main task performance. Finally, the last main pre-task activity outlined by Skehan (1998) is planning. Drawing on previous research findings (Foster & Skehan, 1996; Mehnert, 1998), he posits that pre-task planning might lead to better task performance. It is speculated, based on the findings of Foster and Skehan (1996), that unguided planning (i.e. planning without guidance from the teacher) and guided planning can both promote attention to form, but along different dimensions. Unguided planning potentially leads to greater accuracy, because learners may focus on form and think about the language they are going to use when left to themselves (Skehan, 1998). On the other hand, planning that is guided by the teacher's suggestions about content and language may result in an increase in complexity (Skehan, 1998).

For the main task stage, Skehan (1996, 1998) emphasizes the importance of task selection and the conditions in which tasks are implemented. He believes that task demands can be manipulated along three dimensions that determine task difficulty: *code complexity* (the language required), *cognitive complexity* (the thinking required) and *communicative stress* (performance conditions). Tasks that demand high language processing (e.g. syntactically and lexically complex language) and high cognitive processing (e.g. processing of unstructured or unclear input information) are difficult along the code complexity and the cognitive complexity dimension, respectively. It is assumed that difficult tasks that require learners to process meaning excessively are detrimental to attention to form (Skehan, 1996, 1998). Regarding task condition, Skehan (1996, 1998) hypothesizes that the decisions made on different task variables (e.g. time constraint, number of task participants, task modality) can influence learners' allocation of attention. For instance, increasing the time pressure during a task may make the task more difficult, and learners might focus on accomplishing the task in time (i.e. meaning) rather than paying attention to form.

Lastly, as regards the post-task stage, Skehan (1996, 1998) describes two potential benefits of post-tasks relating to learners' attention to form. First, he hypothesizes that post-tasks facilitate reflection on L2 and the consolidation of L2 knowledge. Pointing out Willis' (1996) language focus phase as an example, Skehan (1998) suggests that post-task activities can encourage form-meaning mappings and pattern identification, which may be transitory during meaning-oriented main tasks. In addition, in this stage, the main task performance can be utilized as task input to promote learners' noticing of gaps in their interlanguage and L2 restructuring (Skehan, 1998). Second, Skehan (1996, 1998) states that another benefit of the post-task stage is associated with how the incorporation of a post-task activity can alter the allocation of learners' attention during the main task. In particular, he proposes that a subtle way to promote attention to form during the main task without direct interference from the teacher is to create learners' awareness of the post-task. This post-task, he suggests, should be an activity for which form is important, such as a language test, a public performance task or a task involving the analysis of language recorded from the main task performance (Skehan, 1998). It is hypothesized that learners' attention will be redirected from meaning to form during the main task when attention to form is beneficial for the upcoming post-task. In Skehan's (1998) own words:

> …foreknowledge of which [i.e. the post-task activities for which form is important] will lead the learners to allocate slightly more attention to form during the task, if such attention is thought to be useful as preparation for the task to come (the post-task activity). In this case, it is likely that if attention can be rechannelled towards form, the principal effect would be on accuracy, since standards of subsequent performance would be less likely to make more complex language prominent than to make correct use of whatever level of language was mobilized during the task. (p. 148)

Importantly, Skehan points out that the main effect of such post-task awareness is on accuracy rather than complexity. Noting the positive effects of post-task anticipation on accuracy, Foster and Skehan (2013) state that:

> We can also portray these effects by examining the connection between working memory and L2 performance. What is clear is that learners, under the right set of

> conditions, are able to mobilize attentional resources through central executive use in such a way that form is advantaged. Greater monitoring of performance may occur, or more intensive accessing of long-term memory material. (p. 264)

In this statement, Foster and Skehan (2013) posit that post-task anticipation can create a condition in which learners mobilize their cognitive resources to monitor their performance more closely and access the L2 knowledge in their long-term memory more intensively, thereby resulting in an increase in accuracy.

Overall, Skehan's (1996, 1998) proposal describes how attention to form can occur in every task stage. This is different from Willis' proposal (1996), which recommends delaying attention to form until after the main task. However, a common assumption shared by these frameworks is that post-tasks are useful for promoting reflection on the L2 and the consolidation of L2 knowledge. An interesting speculation by Skehan is that anticipation of some post-task activities can alter learners' attentional allocation during the main task and enhance attention to form. Although this hypothesis is a plausible one, to date, very few studies have been conducted to test it. In the section that follows, I will review the research in this area in detail.

4.5 Empirical studies on post-task anticipation

To the best of my knowledge, to date, only two studies have been carried out to investigate post-task anticipation in SLA (Foster & Skehan, 2013; Skehan & Foster, 1997). They were both conducted by the same researchers to explore the effects of post-task anticipation on L2 complexity, accuracy and fluency.

In the first study, Skehan and Foster (1997) examined how the three aspects of language performance are influenced by planning and post-task anticipation. Post-task anticipation was operationalized as the anticipation of being selected to perform the main task publicly later. Forty pre-intermediate learners of English were assigned to either a control (- post-task anticipation) or an experimental (+ post-task anticipation) group. Participants in the experimental condition were told prior to the main task that, once the main task finished,

two pairs would be selected to perform the task again in front of the whole class. The main tasks administered in this study included a decision-making, a narrative and a personal information exchange task. Recordings of the main task performance were analyzed for complexity, accuracy and fluency, which were assessed by the number of clauses per c-unit, the percentage of error-free clauses and the frequency of pauses, respectively. It was hypothesized that awareness of the post-task would encourage learners to monitor their speech more extensively during the main task performance, resulting in greater accuracy, while complexity and fluency would remain unaffected.

The findings of this research confirmed the hypothesis pertaining to complexity. Post-task anticipation did not significantly affect language complexity in any of the three tasks. The difference across post-task anticipation conditions in terms of fluency was only significant in the narrative task, with significantly lower fluency observed in the experimental group than in the control group. As far as accuracy is concerned, the effect of post-task anticipation was unclear. The results showed that participants in the experimental condition produced significantly more accurate language than students in the control group during the decision-making task. However, although there was a similar trend towards increased accuracy in the experimental group in other task types, the differences across groups in those tasks were not significant. Therefore, the influence of post-task anticipation on accuracy is inconclusive based on these findings.

Skehan and Foster (1997) point out that there could be three explanations for the findings concerning accuracy. First, it might be the case that post-task anticipation does not influence accuracy. Second, the effect of post-task anticipation on accuracy may be weak and a larger sample size is needed to detect the effect more clearly. Third, the operationalization of post-task anticipation in this study was probably not robust enough to raise participants' consciousness of their L2 accuracy. Regarding this, Foster and Skehan (2013) outline three shortcomings of the operationalization of post-task anticipation in this study (Skehan & Foster, 1997) which possibly led to a relatively weak effect observed on

accuracy. These limitations concern the degree of "threat"[2] (Foster & Skehan, 2013, p. 254) posed to learners and the nature of the post-task. First, the authors hypothesize that the threat of the public performance task may seem remote to learners because the chance of them being selected to perform it was low (i.e. only two pairs were selected from the whole class). Second, it is speculated that learners might not see the public performance task as a threat. Third, the public performance task may not emphasize the importance of language correctness and error avoidance.

In an attempt to address these shortcomings, Foster and Skehan (2013) conducted another study by operationalizing post-task anticipation differently. In this study, post-task anticipation was operationalized as the anticipation of post-task transcription. Foster and Skehan (2013) reasoned that this task reduces the sense of remoteness to the threat of the post-task compared to the task utilized in their previous study (Skehan & Foster, 1997), because it demands that all learners confront their own L2 output. Unlike the previous study, in which only two pairs of students were selected to perform the post-task, this study required every learner to individually transcribe the discussion they had with their peer during the main task as homework. In addition, they assumed that the transcription activity would require learners to pay attention to language form and correctness (Foster & Skehan, 2013). In the post-task, each participant was asked to transcribe their performance from a tape recording as accurately as possible and identify the mistakes they made without making corrections. In this study, 45 pre-intermediate learners of English were randomly allocated to either a control (- post-task anticipation) or an experimental (+ post-task anticipation) condition. The main tasks utilized were a decision-making and a narrative task. Recordings of the main task performance were made during normal classroom sessions and these were coded for L2 complexity, accuracy and fluency. Complexity was measured by the ratio of clauses to AS units. Fluency was assessed in terms of breakdown (mid-clause pauses per 100 words) and repair fluency (reformulations per 100 words). Accuracy was gauged by a main measure (proportion of error-free clauses) and an additional measure (number of words in the longest clause satisfying a 70 per cent accuracy

[2] In this sense, "threat" refers to an anxiety-inducing situation or the possibility that it will happen. Following Foster and Skehan (2013), I will use the terms "threat" and "threatening" in this thesis. From this point on, these terms will be strictly associated with anxiety and by no means with the notion of harm.

criterion). To fulfil the 70 per cent criterion, at least 70 per cent of the clauses of that length must be accurate. Following Skehan and Foster (1997), Foster and Skehan (2013) predicted that post-task anticipation would not affect fluency and complexity, but would rather promote accuracy in all main tasks. They hypothesized that anticipation of the post-task transcription might have a stronger effect on students' L2 accuracy compared to anticipation of the public performance task.

Confirming the authors' prediction, the results of this study did not reveal any significant effect of post-task anticipation on fluency in any task. However, inconsistent with their prediction, some influence of post-task anticipation on complexity was detected. It was found that post-task anticipation led to an increase in complexity in both main tasks (i.e. decision-making and narrative tasks), and the difference in L2 complexity across post-task anticipation conditions was significant in the decision-making task. In terms of accuracy, anticipation of the post-task resulted in a significant increase in accuracy, as assessed by the additional measure, in all tasks. While the data for the main accuracy measure showed that the experimental group produced language more accurately than the control group during both main tasks, the difference was only significant in the decision-making task. This suggests that post-task anticipation may have greater influence on accuracy during the decision-making task compared to the narrative task. Based on their findings, which revealed positive effects of post-task anticipation on accuracy, Foster and Skehan (2013) speculated that the anticipation of post-task transcription potentially enhances learners' control over their L2 repertoires by encouraging them to direct their attentional resources towards language monitoring and accessing linguistic knowledge in their long-term memory.

Interestingly, although the overall findings of this study indicate that post-task anticipation is associated with both increased complexity and accuracy, Foster and Skehan (2013) observed that participants tended to prioritize only one of these aspects during task performance. Hence, they concluded that post-task anticipation may promote attention to form, but this attention could be manifested in increased complexity or accuracy depending on learners' personal orientation.

A summary of the findings of Skehan and Foster (1997) and Foster and Skehan (2013) is presented in Table 2. Based on these findings, it seems that awareness of a post-task may influence accuracy more than other performance aspects – complexity and fluency. Accuracy is the only performance aspect that was found to be significantly influenced by post-task anticipation in both studies.

Table 2: Summary of the effects of post-task anticipation found in previous research

			Narrative task	Decision-making task	Personal information exchange task
Skehan and Foster (1997)	Fluency	Pause fluency	- Fluency	No	No
	Complexity	Clauses per c-unit	No	No	No
	Accuracy	Error-free clauses	No	+ Accuracy	No
Foster and Skehan (2013)	Fluency	Mid-clause pauses	No	No	
		Reformulation	No	No	
	Complexity	Clauses per AS unit	No	+ Complexity	
	Accuracy	Error-free clauses	No	+ Accuracy	
		70% criterion	+ Accuracy	+ Accuracy	

Note. + = significant increase; - = significant decrease; No = no significant effect

Compared to Skehan and Foster's (1997) study, Foster and Skehan's (2013) research detected a clearer influence of post-task anticipation on accuracy. While the first study observed a significant positive effect of post-task anticipation on accuracy only in the decision-making task, the latter found this effect in both the decision-making and the narrative tasks when assessing accuracy with the 70 per cent accuracy criterion. It could be that, as Foster and Skehan (2013) note, the 70 per cent accuracy criterion is more sensitive to the differences in accuracy than the measure utilized in Skehan and Foster's (1997) study – error-free clauses. Moreover, the new way of operationalizing post-task anticipation may be more effective for enhancing accuracy (Foster & Skehan, 2013). Despite this, a common finding of both these studies is that post-task anticipation has a greater influence on accuracy in the decision-making task than in other task types. Foster and Skehan (2013) explain that this may be due to the nature of the decision-making task, which can promote accuracy. According to them, the interactive nature of this task allows learners to plan their

speech while their interlocutor is talking and to receive assistance from the interlocutor, which helps them produce more accurate output.

4.6 Summary

Tasks have been widely utilized for L2 instruction following the emergence of TBLT in the 1980s. To date, there have been numerous approaches to the utilization of tasks for language teaching. In the task-supported approach, tasks are administered to give students opportunities to practise linguistic elements that are the target of a form-focused syllabus. Despite the fact that TBLT syllabi are not form-focused, most advocates of TBLT have acknowledged the importance of attention to form. For instance, Skehan (1996, 1998) hypothesizes that this attention promotes balanced development of different L2 performance goals, and Long (1991, 2015) highlights the benefits of focus on form in task-based instruction.

To date, there have been various proposals as to how to raise students' attention to form when teaching L2 through tasks, and some of them concern the implementation of different task stages. Regarding this, two designs of task implementation methodology (Skehan, 1996, 1998; Willis, 1996) demonstrate that attention to form can be promoted during various task stages. Interestingly, Skehan (1996, 1998) speculates that the anticipation of some post-task activities can alter learners' attentional allocation during the main task and foster attention to form. Although Skehan proposed his hypothesis over two decades ago, to the best of my knowledge, only two SLA studies have investigated post-task anticipation. Taken together, the findings of these studies indicate that awareness of a post-task may have a positive effect on accuracy during the main task, and accuracy may be more affected by this awareness than fluency or complexity. In addition, the findings suggest that the effect of post-task anticipation might depend on the type of the main task and how the post-task is operationalized.

Due to the lack of research in this area, the influence of post-task anticipation has been explored within a limited scope. All the main tasks utilized in previous research have been oral production tasks, and only two types of post-tasks have been administered – a public

performance task and a transcription task. Research that extends this scope of investigation by employing other types of main tasks or post-tasks could be useful for deepening the understanding of the effect of post-task anticipation.

Thus far, I have conducted a review of the literature relevant to my study in the last three chapters. The next chapter will discuss the methodology of the present study.

Chapter 5: Research methodology

This chapter is divided into eleven sections. The first two sections outline my research questions, and the rationale of the study. Then, the operationalization of revision and post-task anticipation, and my hypotheses, are discussed. In the latter sections, I explain the research design and ethics, participants, research instruments and the data collection procedure. Finally, the chapter ends with a description of how the data were coded and analysed.

5.1 Research questions

RQ1 What are the L2 revision behaviours of Thai learners of English with A2 to B2 proficiency levels during text chat?

RQ2 What are the effects of post-task anticipation (± post-task anticipation) and type of post-task anticipation (anticipation of an individual vs a collaborative language correction post-task) on revision behaviours, speed fluency and accuracy?

RQ3 How are L2 revision behaviours related to text accuracy in chat scripts?

RQ4 How is learners' typing ability related to L2 revision in text chat?

5.2 Rationale of the study

As CMC has become popular and integral to our daily lives, the main purposes of this study are to investigate the potential of L2 learning through computer-based text chat, and how tasks could be implemented to facilitate learning in this medium. In this section, I will outline the rationale for each of my research questions.

As discussed in Chapter 3, to date, the central claims associated with the benefits of online chat for L2 acquisition concern its ability to promote learners' attention to form, owing to the unique characteristics and affordances of this medium. With increased attention to form, it is possible that learners will attend to their own output and revise and correct their errors

extensively in an online context. Indeed, L1 researchers have hypothesized that the composing medium and transcribing technology have an impact on writing and revision processes (Hayes, 1996, 2012). Due to the scarcity of research exploring revisions in L2 chat, the current research aims to explore L2 revision behaviours in **RQ1**. By uncovering how learners revise during chat, we can gain a better understanding of how L2 learners pay attention to form in this medium. The findings of this research will complement previous research where findings have been mixed with some studies detecting increased attention to form in online chat while others do not show such effects.

Furthermore, as discussed in Chapter 3, revisions in previous studies of L2 chat have been explored as either self-corrections or self-repairs. Considering revisions in chat as self-corrections is too narrow in scope because it only focuses on changes made targeting erroneous features. Self-repair, on the other hand, is a concept commonly associated with speaking. To date, language monitoring in text chat has been exclusively compared to speech monitoring. However, because dialogue boxes allow chat users to edit their output privately during drafting, monitoring of this type of communication is not the same as that in face-to-face interaction, in which all adjustments made to the produced output, i.e. utterances, are noticeable to the listener. I suggest that, due to the text-based nature of chat, changes made in this medium would be better conceptualized as revisions. As chat is text-based, involves typing and allows more time to reflect on output compared to speaking, the psycholinguistic processes involved in L2 production in chat might be similar to those in writing. Therefore, in the current study I take the initiative to explore linguistic changes made during L2 chat within the framework of L2 writing revision.

Another main rationale for the investigation of L2 revision behaviours during text chat (RQ1) is that there are several serious shortcomings in the studies in this domain (as outlined in Chapter 3, Section 3.5) that need to be addressed. In this study, I aim to address issues relating to data sources, the classification of changes and measures of L2 corrections. First, regarding data sources, while many L2 chat revision studies have relied on chat logs as the only data source, this study will, following O'Rourke's (2008) suggestion, adopt various data collection methods to obtain richer information about chat users' experience.

These include computer screen recording and keystroke logging, which provide information on learners' drafting process. Since, to the best of my knowledge, no previous studies on L2 online chat revision have utilized keystroke logging software, the use of such technology in the current study will be a methodological contribution to this field. Keystroke logging software tracks mouse movements, keyboard activities and temporal data, such as pauses, total active writing time and time stamps of each action, allowing researchers to obtain important information about the drafting process.

Second, in relation to the classification of changes, although verbal protocols have been commonly adopted in research exploring language adjustments made during both L2 speech and writing (i.e. self-repair and revision), no previous studies of online-chat self-corrections and self-repairs have taken learners' self-reporting into account. In previous research, all changes were analyzed and classified based on researchers' own judgement. In the domain of L2 speech research, Kormos (1998a, 1998b, 1999a) recommends that researchers examine learners' underlying communicative intentions, in addition to surface structures, to ensure reliability in the coding of self-repairs. Drawing on this suggestion, I propose that verbal protocols should be adopted in L2 online chat studies to gauge learners' underlying communicative intention when revising and to ensure the accurate classification of types of error corrections and changes made during chat. To this end, stimulated recall interviews were conducted in this study.

Finally, with respect to the problem pertaining to the measuring of corrections, most studies examining L2 corrections during chat have only explored the total number of error corrections, or compared the frequency or proportion of different types of corrections. To date, only one study has examined error correction rates (Yuan, 2003), which could yield better insights into how much attention learners pay to each type of error during chat sessions, as discussed in Chapter 3, Section 3.5. However, because in that study error corrections were explored from the chat scripts of just two participants, and the types of errors were not classified into comprehensive categories, the present study fills these gaps by exploring lexical and grammatical error correction rates through both the chat log and drafting-process data of 84 participants.

Concerning **RQ2**, the investigation of the effects of post-task anticipation is motivated by Skehan (1998), who hypothesizes that L2 learners' knowledge of a post-task may redirect their focus from fluency to accuracy during the main task. This knowledge might create a condition that encourages learners to monitor their language more closely, resulting in greater control over their L2 repertoires (Foster & Skehan, 2013). Since Skehan and Foster pointed to the lack of research on post-task anticipation in 1997, the issue has remained underresearched until now with, to the best of my knowledge, only two studies (Foster & Skehan, 2013; Skehan & Foster, 1997) having addressed the topic. With such scant research and the positive findings of these two studies which reveal the potential of post-task anticipation for promoting attention to form and accuracy, it is warranted to explore the topic further. While both of these studies investigated the effects of post-task anticipation on L2 complexity, accuracy and fluency during oral production tasks, this study extends the scope of investigation to online writing (i.e. text chat), focusing on its effect on revision, as well as on accuracy and fluency. Although no previous studies have examined the effect of post-task anticipation on revision, it could be that attention to form and L2 accuracy fostered through an awareness of the post-task may lead to more revisions. Moreover, the post-tasks administered in previous studies were either a public performance or a transcription task. This research employs a new type of post-task – a language correction post-task. By extending the scope of research, I aim to broaden the understanding of the effects of post-task anticipation. As discussed in Chapter 4, Section 4.6, previous research findings suggest that the influence of post-task anticipation on the main task performance might be different depending on the type of main task utilized and how the post-task is operationalized. Essentially, the investigation of post-task anticipation falls within the research domain of task implementation conditions, which, according to Skehan (2016), can be easily applied to pedagogic settings and used with a whole range of tasks without cost implications.

In terms of **RQ3**, the investigation of the relationship between L2 revision behaviours and accuracy in online chat is motivated by Révész, Michel, and Lee's (2017) L2 writing study, which explored this relationship, and the great scarcity of writing research in this area. As

reviewed in Chapter 2, Section 2.2.6, previous research exploring the relationship between L2 revision and text quality found no significant link between total revision frequency and text quality, or between subcategories of revisions and most text quality measures. It may be that, rather than the frequency of total revision or of different types of revision, the success of revisions is a more important factor determining learners' text quality. Although it is more than a decade ago since Stevenson et al. (2006) suggested that the outcome of L2 revisions might be important for the improvement of text quality, to the best of my knowledge, there has been no attempt to test this hypothesis empirically. Thus, this study aims to fill this research gap by including the rate of successful error corrections as a measure of revision, and examining how it is related to text quality, which is operationalized as final text accuracy. In addition to this, another important rationale for RQ3 is that, while Sauro and Smith (2010) compared L2 complexity in revised, unrevised and deleted online chat text, to the best of my knowledge, no research to date has explored the relationship between L2 revisions carried out during chat and accuracy.

Finally, the justification for **RQ4** is that, despite the possible influence of typing ability on writing processes (Kellogg, 1996), very few studies have investigated how typing ability is related to L2 writing (Barkaoui, 2014, 2016; Barkaoui & Knouzi, 2018). Among these studies, only one (Barkaoui, 2016) explored the effect of typing skills on L2 revision. To address this lack of research, RQ4 explores how learners' typing ability is related to L2 revision in text chat. While Barkaoui's (2016) research was done in an individual computer-assisted writing context, the current study is the first to explore the link between L2 typing ability and L2 revision in online chat. Although there have been some calls for the exploration of keyboarding and typing skills in L2 text-chat research (Lai & Zhao, 2006; Sauro & Smith, 2010), no studies on L2 text chat to date have gauged participants' typing ability.

Thus far, I have outlined my research questions and the rationale of my study. In the sections that follow, I will explain how revision and post-task anticipation are operationalized in this research.

5.3 Operationalization of 'revision'

Two important categorizations of revisions in the literature have been drawn up by Stevenson et al. (2006) and Lindgren and Sullivan (2006), who propose similar classifications of revisions based on location, categorizing revisions into internal and external revisions. *Internal revisions* are those occurring in the writer's mind before transcription. As opposed to internal revisions, *external revisions*, which are the focus of this study, are visible during text production. This study operationalizes revision as any changes made to the text in a text-based CMC context. It examines 1) revisions that occur during drafting, i.e. *CMCovert repairs* (Smith, 2008), which are overt from the perspective of composers but remain covert from (i.e. cannot be discerned by) interlocutors and 2) adjustments made to the already-sent text. The former is shown through the actions such as deleting, adding, inserting, splitting and merging any parts of the text while drafting chat messages, while the latter involves sending a new message to repair the previous one. Although repairs can be both other- and self-initiated, all revisions examined in this study were self-initiated.

5.4 Operationalization of 'post-task anticipation'

Drawing on Foster and Skehan's (2013) hypothesis, which assumes that the effect of post-task anticipation on accuracy is based on 1) the emphasis of the post-task on language correctness and 2) the degree of threat posed to learners, this study operationalizes post-task anticipation as the awareness of a language correction post-task. This post-task entails learners' confronting their own errors. Hence, it may predispose them to increased attention to accuracy during the main task. Following the procedure of Foster and Skehan's (2013) study, this study requires all participants in the experimental condition to perform the post-tasks, therefore the sense of remoteness to the threat of post-tasks is reduced. In addition, the operationalization of post-task anticipation of this study is in line with Skehan's (1998) suggestion, which proposes that the awareness of a language analysis post-task may influence L2 performance more strongly than that of the public performance task utilized by Skehan and Foster (1997). In his own words:

> ...instead of public performance, it may be more effective to require learners to have their task-based performances analysed either by themselves or by others. For example, performances can be recorded on video, and then learners, in their own groups or in larger groups, can analyse their performance for form or communicative effectiveness ... the need to analyse one's own performance (or the anxiety that others may be going to do so) may be more galvanizing, and have a more direct effect on attention manipulation. (p. 148)

The current study utilizes two types of post-tasks: an individual and a collaborative language correction task. While, in the individual task, learners were asked to perform the post-task alone and only to correct their own errors, in the collaborative version they had to discuss all errors in chat scripts with their partner. The awareness that their language errors may be seen by the partner and discussed during the collaborative post-task may be especially face-threatening. Thus, the collaborative post-task may have even greater impact on the main task performance, compared to the individual language correction task.

5.5 Hypotheses

In this section, I will state my hypotheses relating to each research question. Regarding **RQ1**, my hypotheses are based on the notion that revisions in L2 chat may share some characteristics with those in individual computer-assisted L2 writing. Drawing on the common findings of research investigating revision during individual computer-assisted L2 writing (Barkaoui, 2016; New, 1999) and L2 online chat (Golonka et al., 2017), my first hypothesis is that there will be a prevalence of typographical revisions. With respect to the linguistic unit of revisions, I hypothesize that L2 learners will revise lower-level linguistic units (e.g. below-word, word and phrasal levels) more than higher-level ones (e.g. clausal and sentential levels) during chat sessions. This hypothesis is based on previous research findings (Barkaoui, 2016; Choi, 2007; Li, 2006; New, 1999; Révész, Michel, & Lee, 2017; Stevenson et al., 2006) that show that revisions occur more frequently to lower-level than higher-level linguistic units during individual computer-assisted L2 writing. It may be that learners tend to make local changes when writing on a computer. While previous L2 chat research has found that learners seem to attend more closely to errors relating to grammar

than lexis (Fredriksson, 2015; Golonka et al., 2017; Smith, 2008, 2009), carry out more repairs targeting form-related errors than language appropriateness or content (Smith, 2008, 2009) and successfully correct a high proportion of errors (Golonka et al., 2017), their shortcomings, outlined in Chapter 3, Section 3.5, render the findings of these studies tentative. Hence, it is difficult to form hypotheses regarding the target of revisions and the outcome of error corrections based on these findings. It may be that form-related revisions occur frequently during L2 chat and learners often correct their errors successfully because text chat potentially promotes attention to and reflection on form.

As far as **RQ2** is concerned, I speculate that post-task anticipation will have no significant effect on fluency, but it will promote increased L2 accuracy, revisions and error corrections during the main task. This prediction draws on a hypothesis that assumes the potential of post-task anticipation to redirect learners' attention from meaning to form, and motivate learners to monitor their L2 more extensively (Foster & Skehan, 2013; Skehan, 1998). Additionally, it is based on the findings of Foster and Skehan (2013), who observed a clear significant effect of post-task anticipation on accuracy, but no significant effect on fluency. While the main tasks utilized in their study were oral production tasks, I hypothesize that the positive effect of post-task anticipation may carry over to the main tasks performed in a text chat environment. In terms of types of post-task anticipation, because the collaborative language correction post-task is likely to be more threatening to learners, it may have greater impact on the main task performance in comparison to the individual post-task.

In the case of **RQ3**, my hypothesis is based on the results of previous writing research exploring the relationship between L2 revision and accuracy (Révész, Michel, & Lee, 2017). The results of the study indicate no significant relationship between accuracy and any of the revision measures (frequencies of revisions made at different linguistic levels, ratio of words and ratio of characters in the final text to those actually typed). Therefore, I predict that the relationship between accuracy and most measures of revisions (i.e. the quantity, linguistic unit and target of revisions) might be non-significant in my study. Yet, drawing on Stevenson et al.'s (2006) speculation about the possible link between the

outcome of revisions and text quality, I predict that there could be a significant relationship between accuracy and the rate of successful error corrections.

Lastly, no hypothesis is formulated in relation to the relationship between learners' typing ability and L2 revision during online chat (**RQ4**). This is because only one study (Barkaoui, 2016) has investigated the impact of typing ability on L2 revision behaviours, and it was done on individual computer-assisted writing. No L2 chat studies to date have gauged how competent participants were at typing. It could be that both L2 revision during online chat and during individual computer-assisted composition have a similar relationship with typing ability. Based on this perspective it is possible that, consistent with the findings of Barkaoui's (2016) study, revision behaviours such as the quantity and target of revisions are not associated with L2 typing ability in online chat. However, with the scant research on L2 typing ability, the connection between typing ability and L2 revision remains unclear. Based on the hypothesis vis-à-vis L1 writing (Kellogg, 1996), poor motor skills (e.g. typing) may increase L2 learners' cognitive load, thereby negatively affecting their writing and revision processes during chat.

5.6 Research design

This study adopts a mixed design manipulating both between-participant and within-participant factors. The former involves allocating students to either a control group or an experimental group. While the control group performed both main tasks without any post-task anticipation, the experimental group was informed about a post-task before each of the main tasks. The within-participant factor of this study is type of post-task anticipation: individual vs collaborative post-task anticipation. All participants in the experimental condition were told to anticipate an individual language correction post-task before one main task, and a collaborative language correction post-task before another. Therefore, the experimental group anticipated both types of post-tasks.

A combination of different data collection methods was utilized in the attempt to gain rich research data and fill the gaps in previous research. Performance data were gathered from

participants' screen recordings, keystroke logs and chat scripts. Stimulated recall interviews cued by replaying screen recordings of their performance were also conducted to explore participants' thoughts during revision. This type of data was collected in an attempt to ensure the reliability of the coding of revisions. To gauge typing ability, a typing test was administered. Finally, an exit questionnaire provided additional insights into the participants' perspectives regarding tasks, L2 revision in online chat and post-task anticipation.

5.7 Research ethics and participant recruitment

Prior to participant recruitment, this research was reviewed and approved by Lancaster University's research ethics committee. In accordance with the university's ethical guidelines, students' participation in this study was entirely voluntary and they could decide to opt out of the experiment at any time during the experiment if they felt uncomfortable with continuing without having to provide any reason. Each student was awarded 300 baht, approximately six pounds sterling, for taking part in the study, and they would not forfeit this even if they decided to withdraw from the experiment. In addition, all collected data were securely stored.

Regarding the recruitment procedure, a paper-based Oxford Placement Test (Allan, 2004) was distributed in seven English classrooms at three universities to assess students' English proficiency. Students whose proficiency levels were within the range of A2 to B2 were contacted, and sent an information sheet (Appendix A) if they agreed to take part in the study. The information sheet described the details of the research project, such as its aims and the data collection procedure, and this was sent to participants before the date of the experiment. To observe natural revision behaviours and performance, participants were not informed about the focus of this research, simply that it explored L2 learning through online chat. On the day of the experiment, before it began, they were given a paper-based consent form (Appendix B) and had the opportunity to read and ask questions before signing it. Both the information sheet and consent form were written in the participants' native language.

5.8 Participants

Participants were 84 Thai learners of English (Female = 71, Male = 13) who were undergraduate students of language studies, psychology, dentistry and communication arts courses at three universities in Thailand. Participants in each pair came from the same English class. Therefore, all of them knew their chat partner before taking part in the experiment. Participants' proficiency levels, as measured by the Oxford Placement Test, were within the range of A2 to B2 in the Common European Framework of References for Languages (Council of Europe, 2001) (*M* = 60.15, *SD* = 5.89, Score range = 50 to 73 out of 100) and their ages ranged from 17 to 22 (*M* = 19.13) years. All had been learning English for a period of six to 20 years. Based on the responses to the background questionnaire, out of 84 participants, 83 spent more than one hour per week on a computer, with more than half (N = 46) reporting spending more than six hours a week. A great majority (N = 78) spent more than one hour a week on instant messaging, and 80 participants had used English for this activity prior to the experiment. In terms of typing ability, their typing test scores calculated from the number of correct words typed per minute ranged from 24 to 66, with the mean score being 42.98 (*SD* = 9.64). Participants' background information is summarized in Tables 3 and 4.

Table 3: Background data (1)

Proficiency (Oxford Placement Test scores)	Mean	60.15
	Range	50 – 73
Age (years)	Mean	19.13
	Range	17 – 22
English learning experience (years)	Mean	14.19
	Range	6 – 20
Typing test scores (correct words typed per min.)	Mean	42.98
	Range	24 – 66

Table 4: Background data (2)

Gender	Male	13
	Female	71
Native language	Thai	84
Time spent on a computer per week	Half an hour – 1 hour	1
	1 – 2 hours	6
	3 – 4 hours	16
	5 – 6 hours	15
	More than 6 hours	46
Time spent on instant messaging per week	Less than half an hour	1
	Half an hour – 1 hour	5
	1 – 2 hours	5
	3 – 4 hours	9
	5 – 6 hours	15
	7 – 8 hours	11
	More than 8 hours	38
Frequency of using English for instant messaging	Always	14
	Often	13
	Sometimes	22
	Rarely	31
	Never	4

Participants were randomly allocated to either a control (N = 28) or an experimental (N = 56) condition. Although the mean proficiency and typing test scores of the control group (Proficiency: $M = 60.96$, $SD = 5.80$, Typing: $M = 45.52$, $SD = 10.30$) were higher than those of the experimental group (Proficiency: $M = 59.75$, $SD = 5.97$, Typing: $M = 41.71$, $SD = 9.18$), the results of independent sample t-tests revealed no significant difference in proficiency scores ($t(82) = 0.89$, $p = .378$, BCa 95% CI [-1.398, 3.860]) or typing test scores ($t(82) = 1.72$, $p = .089$, BCa 95% CI [-0.828, 8.308]) across groups. Therefore, group comparability in terms of English proficiency and typing ability was confirmed. In addition, the two groups were also comparable in terms of the amount of time they spent per week on a computer (Fisher's exact test: $p = .614$, $V = .191$) and on instant messaging (Fisher's exact test: $p = .085$, $V = .364$) and the frequency of using English for instant messaging ($U = 886.50$, $z = 1.01$, $p = .311$, $r = .11$).

5.9 Research instruments

This section provides details pertaining to all instruments used in this research. They include two main tasks, two post-tasks, proficiency and typing tests, questionnaires and keystroke logging and screen capture tools.

5.9.1 Main tasks

This research utilized two main tasks: *The Evolution of Televisions* and *The Evolution of Cameras* (Appendix C). Both were information-gap tasks that involved picture description and decision-making. In each of these tasks, participants assumed the role of classmates in a Product Design course. They were told that they had not been able to follow a teacher's lecture about the evolution of televisions/ cameras very well. While the lecture was on four early designs of televisions/ cameras, in each task each participant was given two pictures showing the designs of two products only. Each participant had two different pictures. They were instructed to exchange information with their partner in order to get complete information about all the designs to prepare for an upcoming test. They were asked to describe the pictures they had to their partner in as much detail as possible by including descriptions of the shape, colour, material and other external features of the products, and to keep in mind that the task outcome might have an effect on the test. In addition, participants had to agree on which two out of the four designs depicted the earliest and the latest designs. Because the cameras and the televisions shown in the pictures were very different from those used nowadays, some parts of the pictures were labelled in Thai, the participants' L1, to indicate specific components and functions of the products that might not be familiar to them. A list of specific English vocabulary (e.g. *loudspeaker* and *viewfinder*) was also given to aid task transaction. All participants had 25 minutes to complete each task.

My rationale for adopting information-gap tasks is that these involve a two-way exchange of information, which is natural in online chat. Moreover, the main tasks utilized in this study were specially designed to promote revision. They involve describing the photos of

very old models of cameras and televisions to a partner who might not have seen them before. Therefore, they require close attention to accuracy of expression and detailed descriptions. Furthermore, because successful descriptions are important (i.e. the partner needs to use the information to prepare for an upcoming test), participants might be motivated to describe the photos more accurately. All of these factors may promote revision and language correction during the main tasks.

5.9.2 Post-tasks

The two post-tasks utilized in this study were an individual and a collaborative language correction task. In both tasks, participants were asked to read the log of a chat conversation they had had with their partner during the main task in a word processor, *MS Word*. They were asked to identify language errors and any instances where they thought the language should be improved, and make corrections to the *Word* document. In the collaborative post-task, participants were required to work in pairs, with each pair reading the chat log of their joint chat conversation, discussing all of their errors orally (face-to-face), and finding ways to correct them together in the *Word* document. The individual post-task was performed alone, with each participant reading the chat log by him/herself and correcting only his/her own errors in the word processor. All participants were given 15 minutes to complete each post-task.

5.9.3 Proficiency and typing tests

Participants' proficiency was measured by the Oxford Placement Test (Allan, 2004), consisting of 100 multiple-choice items targeting grammar, vocabulary and reading skills. Following the test guidelines (Allan, 2004), each of the participants was given 50 minutes to complete the test.

To measure the participants' typing skill, a typing test was administered. Participants had two minutes to retype the sentence "*All you need is love.*" in *MS Word* as many times and as accurately as they could. This task assesses participants' typing skills without the effect

of text production or language processing abilities, as with most other tasks used to gauge typing ability in previous L1 and L2 writing research, such as text copying and typing a memorized nursery rhyme (Barkaoui, 2014, 2016; Leijten & Van Waes, 2013; Leijten, Van Waes, Galbraith, & Torrance, 2011). In this study, the sentence the participants were required to type was easy to remember, not imposing a high cognitive load on their working memory, which might otherwise confound the test results.

5.9.4 Questionnaires

Two paper-based questionnaires were created and administered in this study: a background questionnaire (Appendix D) and an exit questionnaire (Appendix E). A 14-item background questionnaire enquired about participants' biographic data (such as gender, age, L1 and English language learning background), computer usage, participation in online communication activities and their relationship with their chat partner in the study. This questionnaire included additional questions targeting variables that have been hypothesized to affect writing revision and L2 monitoring (Barkaoui, 2007; Kormos, 1999b; Krashen, 1978; Seliger, 1980): personal orientation towards certain performance aspects and self-perceived language ability. Participants were asked to rate how important they perceive each performance aspect (complexity, fluency and accuracy) to be when communicating in English and online chat. Because text chat has been found to resemble both speaking and writing (e.g. Kern, 1995; Smith, 2005; Tudini, 2003; Yates, 1996; Yuksel & Inan, 2014), participants were asked to rate both their speaking and writing skills by locating themselves on the self-assessment grid of the Common European Framework of Reference for Languages (Council of Europe, 2001), which outlines language abilities of A1 to C2 learners. In addition to this, another question asked participants to rate how competent they were at using English for online chat communication.

A 12-item exit questionnaire consisting of both closed- and open-ended questions was created to elicit participants' responses pertaining to the usefulness of online chat for language learning, L2 revision in this medium, task perception and, in the case of the experimental group, their experience with post-task anticipation. Regarding L2 revision,

participants were asked to list the aspects of language they tended to revise during chat, and identify the factors they thought might affect their online chat revision. Two questions targeted the noticing of errors and language corrections by asking whether they had noticed any language errors while performing online chat tasks during the experiment, and whether they thought they usually noticed and corrected their language errors more during speaking or chatting. An additional question explored why participants chose not to revise their erroneous messages during the experiment, even when they noticed their errors. In terms of participants' task perception, participants were required to rate the extent to which they agreed with the given statements. These statements related to the perceived usefulness, difficulty and stressfulness of each main task and post-task, and how interesting each of them was. Lastly, with respect to post-task anticipation, participants in the experimental group were asked to rate how strongly each type of post-task anticipation influenced them to use English more accurately during the main task, and to give reasons to justify their answers.

5.9.5 Keystroke logging and screen capture tools

During the main tasks, participants' computer screens and keystrokes were recorded by *Flashback Express 5* (2017) and *Inputlog 7* (Leijten & Van Waes, 2013), respectively. Flashback Express, free screen-capturing software, records everything shown on the computer screen and can create video playbacks of recordings. Inputlog logs the data of all keystrokes, mouse movements and clicks, and temporal data such as pauses, total active writing time and time stamps of each action. The benefit of keystroke logging is that it provides direct evidence of participants' writing processes with relatively low obtrusiveness (Leijten & Van Waes, 2013; Van Waes, Leijten, Wengelin, & Lindgren, 2012). It does not involve direct elicitation from participants, as is the case with some other methods, such as think-aloud protocols. Hence, it reduces the risk of jeopardizing the ecological validity of the research (Van Waes, Leijten, & Van Weijen, 2009). To date, a number of L2 writing studies have adopted this logging technology to investigate revisions (e.g. Barkaoui, 2016; Choi, 2007; Lindgren & Sullivan, 2006; Spelman Miller et al., 2008; Stevenson et al., 2006), because it provides important insights into all revisions made during writing

processes. According to Lindgren and Sullivan (2006), keystroke logging "allows revisions undertaken during a writing session to be viewed in the context of their occurrence, and analysed according to their content and position in the developing text ... [and] affords the possibility to analyse the deleted revisions" (p. 158). Due to these benefits, this technology could be a powerful tool to investigate L2 revision in text chat.

5.10 Data collection procedure

This section explains the pilot study and the main data collection procedures. The first part is a brief description of the procedure of my pilot study, the latter part outlines the procedure of the main data collection.

5.10.1 Pilot study

Prior to carrying out the main data collection with 84 participants, I conducted a pilot study with the same design as the main experiment (see Section 5.6) with 10 L2 English speakers in the United Kingdom, whose L1s included Urdu, Russian, Korean, Malay, Chinese and Italian. Their proficiency levels (as measured by the Oxford Placement Test) were within the range of B1 to C1 in the Common European Framework of References for Languages (Council of Europe, 2001) (M = 71.60, SD = 7.87, Score range = 54 to 87 out of 100). All participants reported spending more than one hour per day on a computer, and on average their typing score was 53.90 (SD = 12.17). As the main aims of the pilot were to test the data collection procedure and explore the suitability of all research instruments, including post-tasks, as many as eight out of ten participants were assigned to the experimental group, and the rest to the control group. All the main tasks, post-tasks, proficiency and typing tests, questionnaires and digital tools utilized in the main study were tested during the pilot. The time allotted to complete each main task (25 minutes) and post-task (15 minutes) during the main study was based on the pilot data, which found that participants took approximately these amounts of time to complete these tasks. The output of Wilcoxon signed-rank tests (Appendix I) performed on the pilot data showed that the differences across the two main tasks (camera and television) in terms of revision behaviours and other performance measures were not significant, thus confirming the comparability of the two

main tasks. In addition, the difference across tasks in terms of participants' perceived task difficulty was also non-significant. Perceived task difficulty scores were obtained from the exit questionnaire. As the data collection procedure and research instruments (main tasks, post-tasks, proficiency and typing tests, digital tools and questionnaires) were found to be suitable for the study, only small changes were made to improve the clarity of the task instructions and some questions in the questionnaires before the main data collection.

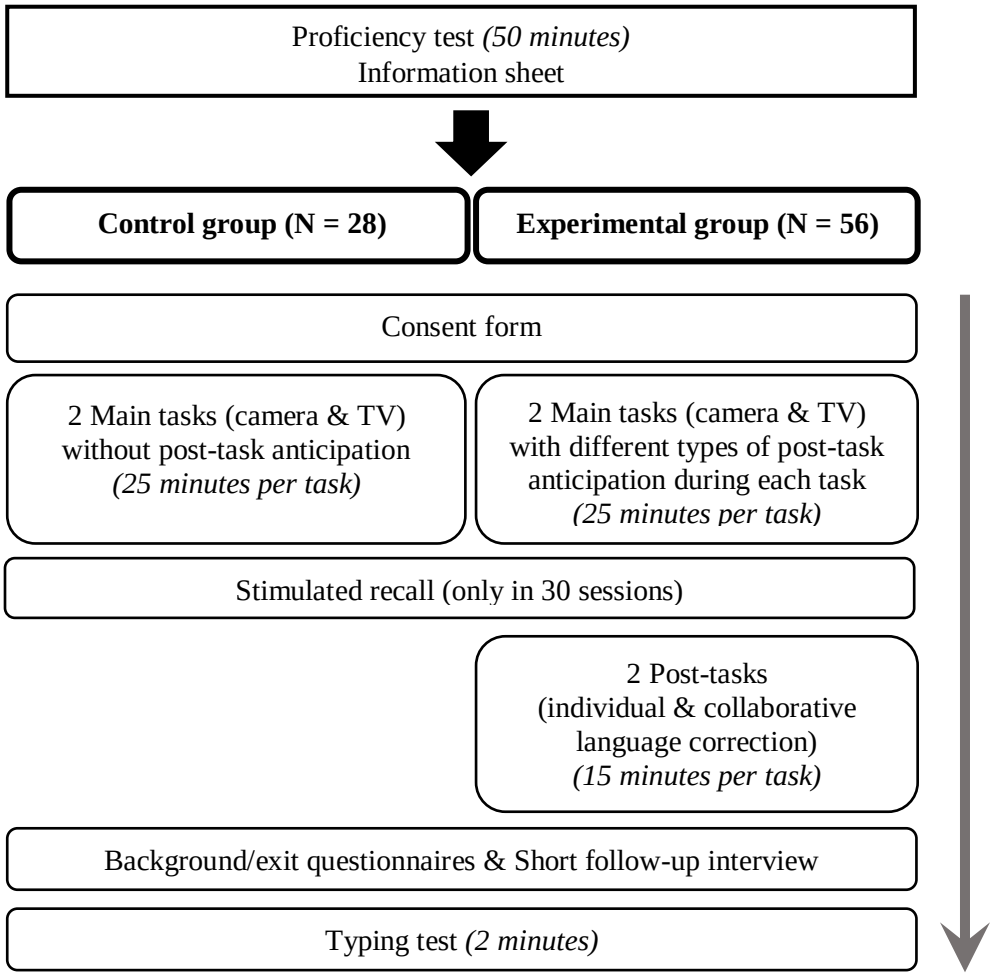
5.10.2 Main data collection

The data collection procedure of the main study is illustrated in Figure 9. As mentioned in Section 5.7, a paper-based Oxford Placement Test was distributed prior to the experiment in the participants' English classrooms to determine their proficiency levels and ensure their levels were relatively comparable. In addition, they were all sent an information sheet before the date of the experiment.

In each data collection session, two participants were invited to come on the same day and time outside of their normal class time. Once a consent form was signed, the experiment began with the two main tasks. In the experimental condition, participants performed two main tasks with different types of post-task anticipation during each task. They were told to anticipate their respective type of post-task (i.e. an individual or a collaborative language correction post-task) before each main task. The order of the main tasks and types of anticipation were counterbalanced over the participants.

In order to ensure the ecological validity of chat conversations, participants were placed in two separate rooms, so that they were not in the presence of each other during the main tasks. Chat interactions were performed in the written chat mode of *Skype*. Specially for the purpose of this research, Skype accounts were created and pseudonyms were used as the users' names. The software features that allow automatic correction and identification of spelling mistakes were disabled during the main tasks.

Figure 9: Illustration of the research procedure



Participants were given 25 minutes to complete each main task via text chat after reading the task instructions. They were told to communicate only in English, to always remain in the Skype window and not to use any referencing tools, such as an online dictionary. During the tasks, their computer screen and keystrokes were recorded. Immediately after the second task, a stimulated recall interview was conducted at random in 30, out of 42, sessions. Participants watched a replay of their screen recording and were asked to pause the video whenever they found instances where they revised their messages or corrected their language, and to verbalize what they were thinking at the time in their L1, Thai. Whenever participants failed to identify these instances, I paused the video to elicit their thoughts. Stimulated recall sessions were recorded by a video camera which captured the video being played back on the computer screen and the participant's voice during the

interview. Due to time constraints and the presence of only one researcher, recall was only done on the second task performance; and in each session, only one participant was randomly selected for recall. In line with Gass and Mackey's (2000) recommendations, I avoided pushing participants for further comments when they could not remember their thoughts, giving concrete reactions or extended responses to their recall comments, and interrupting them. In addition, all interviews were carried out immediately after the second task to ensure that there was as little memory decay as possible.

After the stimulated recall, the experimental group performed an individual and a collaborative post-task. The type of post-task they had been told to anticipate before the first main task was carried out first, followed by the type of post-task they had been told about before the second main task began. All participants were given 15 minutes to complete each post-task and were told not to use any proofing tool (e.g. grammar and spell-checkers) to facilitate task transaction. Logs of the chat conversations were copied into *Word* documents before the post-tasks began. During these tasks, all changes made to chat logs were tracked by the Track Changes feature in *MS Word*. While participants carried out the individual post-task alone in separate rooms, they performed the collaborative post-task together, face-to-face. During the collaborative post-task, participants were allowed to discuss corrections in Thai and their discussions were recorded by a voice recorder.

Finally, all participants completed paper-based background and exit questionnaires, followed by a short interview to clarify their answers and a two-minute typing test. All interview sessions were recorded by a voice recorder. Before the typing test began, participants were given an instruction sheet with the model sentence '*All you need is love.*' They were told to strictly follow the orthographic pattern provided (e.g. to capitalize the letter '*A*' and to always type '.' at the end of the sentence). They were also asked to memorize the sentence, and to recite it aloud immediately before the test started to ensure successful memorization. In all, each data collection session took approximately two hours for the control group, and two hours and a half for the experimental group.

5.11 Data coding and analysis

This section discusses how the data were coded and analyzed. It begins with a description of the coding procedures, followed by the different research measures adopted in this study, and how the data were analyzed to address each research question.

5.11.1 Data coding

The coding was done on keystroke logs, responses to questions in the exit questionnaire, and transcripts of the follow-up interviews. Keystroke logs were coded in *Atlas.ti 8* (2017), data analysis software that allows the systematic coding of text. Log information regarding the focus of revisions was coded twice, first based on surface structures and later on stimulated recall comments. The purpose of this was to ensure the reliability of the initial coding. Keystroke logs of task performances not recalled by the participants were coded only once based on surface structures.

5.11.1.1 Coding of keystroke data

Coding was done on the linear logs generated by Inputlog. To prepare the data for Atlas.ti, linear keystroke-logging files were copied into a word processor and saved as text documents. An example of log data is shown below:

> {45131}this•{2418}camera•{25319}have•two•stands•and•one•{4899}g[LEFTClick]rip•[BACK][BACK][BACK][RIGHT]rip•so•r[BACK]that•it•can•be•stand•on•floor{2528}•easily.[RETURN]

Linear logs reveal the following information:

1) pause length in milliseconds (e.g. *{45131}*)
2) characters typed
3) special keys pressed on the keyboard (e.g. spacebar, backspace, return, left shift and right arrow keys, as indicated by the symbols •, *[BACK]*, *[RETURN]*, *[LSHIFT]* and *[RIGHT]*, respectively)

4) mouse actions (e.g. left click, as indicated by *[LEFTClick]*)

The logs were examined for evidence of revisions during drafting (e.g. whenever participants pressed the backspace or delete key to delete text, used an arrow key or mouse clicks to edit earlier parts of messages, or used the right-click or control-key command for copying and pasting) and adjustments made to already-sent text (i.e. when participants sent a new message to repair a previous one). In cases where it was difficult to code based on keystroke logs alone (such as when revisions were lengthy and complex, or when participants used the mouse for editing), recordings of the participants' computer screens were examined to better discern what was happening at the moment of revision. As the investigations in previous research on L2 online chat revisions were limited to the target and frequency of error corrections/ repairs, this research aimed to explore revisions more extensively by using a detailed taxonomy. The coding criteria for revision (Table 5) were adapted from two revision taxonomies utilized in previous writing research: Van Waes and Schellens' (2003) taxonomy of revision behaviour and Stevenson et al.'s (2006) multi-dimensional revision taxonomy (Appendix F).

Table 5: Coding criteria for keystroke data

Revision	
Focus of revision	(Content, grammar, lexis, reorganization, rephrasing and unclear)
Linguistic unit of revision	(Clause-and-above, phrase, word, letter and punctuation)
Trigger of revision	(Error-triggered, non-error-triggered and unclear)
Outcome of error revision	(Successful and unsuccessful)
Type of revision	(Covert and overt)
Errors not corrected	
Grammatical errors	
Lexical errors	
Errors not in final text	

Like Van Waes and Schellens (2003) and Stevenson et al. (2006), this research explored the linguistic unit and focus of revisions. Following Stevenson et al. (2006), the triggers of revisions and the outcomes of error revisions were also examined. Regarding the focus of revisions, in the initial stage, coding was based on surface structures. That is, stimulated recall data were not yet examined. This dimension was subdivided into different categories:

content, grammar, lexis, reorganization, rephrasing and unclear. Content revisions involve any changes made to the meaning of the text. They include, for instance, corrections aimed at making the description more specific and message abandonment. Grammatical revisions include changes relating to grammar, such as verb forms, tenses, sentence structure, articles, parts of speech and grammatical number of nouns. Lexical revisions are changes associated with lexical items, such as replacing *fridge* with *refrigerator* and *on the middle* with *in the middle*. Revisions were coded as reorganization when any part of the message was rearranged, for instance by changing the clause or sentence order. Rephrasing revisions were defined as any changes which involve expressing an idea in an alternative way, while the meaning of the text remains unaltered. In this study, rephrasing revisions only concern changes made to non-erroneous original text. Participants might rephrase because they were unsure whether what they were typing was correct or not, or they could simply want to find a better way to convey their message. Lastly, revisions with an unclear focus were those which could be classified into more than one category based on keystroke and screen recording data. In this study, typographical revisions were excluded from coding as they are associated with errors in motor functions rather than language use. Moreover, because it was often difficult to tease apart typographical revisions to revisions targeting spelling errors, the decision was made to also exclude spelling revisions.

The second dimension, the linguistic unit of revisions, was divided into five levels: punctuation, letter, word, phrase and clause-and-above. This dimension concerns the boundary of changes made as opposed to the boundary of actions. For instance, when a participant changed *this* to *that* by replacing the letters '*is*' with '*at*', the change was coded as a word-level, rather than a letter-level, revision.

In addition to the above, revision triggers and outcomes of error revisions were coded. All revisions except content revisions were identified as to whether they were triggered by an error. Any revisions with an obscure trigger were coded as *unclear*. The justification for not coding triggers for content revisions was that it was difficult to discern whether these revisions were made to rectify content-related problems or not. Revisions might be triggered by meaning-related errors or the writer might simply have wanted to change the

content of the message. As it was often not apparent, any judgement made on the trigger for this type of revision could be very subjective. Once the trigger was identified, any revisions initiated by an error were further coded as successful or unsuccessful based on their outcome.

While the dimensions described thus far were taken from the taxonomies used in writing research, a new element compatible with text chat data was also added: type of revision. Because online chat allows users to revise both covertly (during drafting) and overtly (by repairing text already sent), changes made to the text were coded either as covert or overt revisions. Overt revisions are evident in chat scripts, whereas covert revisions are not. Importantly, it should be noted that the term *covert* used here, and henceforth, refers to CMCovert repairs (i.e. revisions which are overt from the perspective of the composer but remain covert from the interlocutor). The term should, by no means, be associated with the changes made in one's mind before actual output is produced.

As shown in Table 5, keystroke data were coded not only for revisions, but also for grammatical and lexical errors not corrected. Any errors produced during writing but not corrected were identified. Here, the focus was on whether there was an attempt at error correction, not on its success. Lexical errors include erroneous choices of lexical items, such as *in the left* and *time is going out*. Grammatical errors include those related to grammar, e.g. noun-verb agreement, verb forms, sentence structure, tenses, articles, grammatical number of nouns and parts of speech. During the coding process, it was found that participants frequently omitted an apostrophe when typing (e.g. *its on the lens*). This is in line with some researchers (Gong & Ooi, 2008; Herring, 2001, 2012) who have observed that English language in computer chat does not usually conform to traditional orthographic conventions. Chat users often use non-standard features deliberately as a strategy to economize on their typing effort (Herring, 2001). They shorten words and phrases by using abbreviations or acronyms, and leaving out redundant characters not necessary for comprehension (e.g. *wat* for *what*) to overcome their slow typing speed and time constraints imposed by the technology (Gong & Ooi, 2008). Hence, the decision was made not to code any errors caused by the omission of apostrophes.

As the present study examines final text accuracy, I also coded errors that did not appear in the final text. These include errors (i.e. uncorrected grammatical and lexical errors and outcomes of unsuccessful error revisions) not shown in chat logs because they were subsequently deleted during content revisions. An example of this is when a participant deleted *i has picture A. It is very[...]* and replaced it with *ok* to respond to the partner's new message. Here, the uncorrected grammatical error *has* was deleted during content revision, and thus cannot be seen in the participant's log.

All the initial coding was done by me. An English language teacher who is a native speaker of English then coded approximately 10 per cent of the data (i.e. the data of eight participants) again for different dimensions of revisions and for uncorrected errors. The purpose of this was to test the reliability of the initial coding. Inter-coder reliability (Cohen's kappa) for focus, linguistic unit, trigger and type of revisions was 0.79, 0.88, 0.81 and 1.00, respectively. The figure was 0.77 for the outcome of error revisions, and 0.74 for uncorrected errors. Based on the guidelines from Landis and Koch (1977), these figures signify a substantial degree of agreement, thereby confirming the reliability of the initial coding.

5.11.1.2 Coding of keystroke data based on stimulated recall

As previously discussed, keystroke data were initially coded based on surface structures. In the next stage, keystroke logs of task performances recalled by participants were coded again. Each instance of revision mentioned in stimulated recall was coded again for the focus of revisions based on recall comments. The aim of this was to ensure the reliability of the initial coding through data triangulation. New codes, which included types of focus explored previously (i.e. content, grammar, lexis, reorganization and rephrasing), were added to existing ones. Instances where participants failed to recall and reported typographical or spelling revisions were coded as *No recall* and *Typo/spelling revision*, respectively.

5.11.1.3 Coding of exit questionnaire and follow-up interview responses

Open-ended questionnaire responses and follow-up interview transcripts were coded manually for themes, and answers to exit questionnaire items related to stress during the main tasks were coded with numbers. Regarding coding for themes, questionnaire and interview responses relating to two open-ended exit questionnaire items were examined. These items were associated with the perceived effect of post-task anticipation on accuracy. They required participants to give reasons why they thought individual and collaborative post-task anticipation did or did not influence them to produce language accurately during the main task performance. First, reasons were grouped based on whether they were in support of the effectiveness of post-task anticipation in encouraging L2 accuracy. Then, the responses were scrutinized a second time for the main themes. Once the coding was complete, all codes were checked again to ensure their reliability.

With respect to the exit questionnaire items concerning stress, participants' responses to two five-point Likert items were coded. These items asked participants to rate whether they *strongly agree*, *agree*, *neither agree nor disagree*, *disagree* or *strongly disagree* that each of the main tasks was stressful. Answers were coded with the numbers 1 to 5, with lower values indicating higher stress levels. In other words, the answer *strongly agree* was coded as 1, and *strongly disagree* was coded as 5.

5.11.2 Data analysis

This section presents the different research measures used in this study and describes how the data were analyzed to answer each research question.

5.11.2.1 Measures of revision

Revision behaviours were explored in six different dimensions as summarized in Table 6. The first dimension, quantity of revision, is divided into three measures: frequency of covert revisions, frequency of all (both covert and overt) revisions and ratio of characters, including spaces, in the final text to those actually typed. Frequencies of covert revisions

and total revisions were calculated per 100 words in the final text. This standardized revision frequency has been utilized in many L2 writing studies (e.g. Barkaoui, 2016; Choi, 2007; Hall, 1990; New, 1999; Stevenson et al., 2006). Following Révész, Kourtali, and Mazgutova's (2017) research, which investigated the effect of task complexity on L2 revision, revision quantity was also explored by dividing the number of characters in the final text by the number of characters typed. This ratio compares the quantity of text produced during the whole writing process to that of the final text, i.e. the revised text. The total numbers of words and of characters including spaces in the final text (chat logs) were counted using the word-count function in *MS Word*, and process data were taken from Inputlog output.

Apart from quantity, this study also examined revisions in terms of their linguistic unit, focus and trigger. As with overt and total revisions, the frequency of each type of linguistic unit (clause and above, phrase, word, letter, punctuation), focus (content, grammar, lexis, rephrasing, reorganization, unclear) and trigger (error-triggered, non-error-triggered, unclear) of revisions was calculated per 100 words in the final text. To gauge the outcome of error revisions, the percentage of successful error revisions was calculated by dividing the number of successful revisions by the number of error-triggered revisions, and multiplying the result by 100.

The last measures of revision were error correction rates, which show proportions of errors participants attempted to correct out of the total errors made. This study explored grammatical, lexical and total error correction rates. To determine grammatical and lexical error correction rates, the total number of errors made in each of these categories was first calculated by adding the number of error-triggered grammatical and lexical revisions to the number of errors left uncorrected in each category, respectively. Then, error correction rates were computed by dividing the number of error-triggered grammatical and lexical revisions by the total errors made in each respective category, and multiplying the results by 100. The total error correction rate was obtained following the same procedure, but without making a distinction between the two types of errors and revisions. The formula for calculating error correction rates is given in Table 7.

Table 6: Measures of revision

1. Quantity of revision
1.1 Covert revisions (per 100 words in the final text)
1.2 All revisions (per 100 words in the final text)
1.3 Characters including spaces in the final text divided by characters including spaces typed
2. Linguistic unit of revision (revisions per 100 words in the final text)
2.1 Clause and above
2.2 Phrase
2.3 Word
2.4 Letter
2.5 Punctuation
3. Focus of revision (revisions per 100 words in the final text)
3.1 Content
3.2 Grammar
3.3 Lexis
3.4 Rephrasing
3.5 Reorganization
3.6 Unclear
4. Trigger of revision (revisions per 100 words in the final text)
4.1 Error-triggered
4.2 Non-error-triggered
4.3 Unclear
5. Outcome of error-triggered revision (%)
5.1 Successful error revision
6. Error correction rate (%)
6.1 Grammatical error correction rate
6.2 Lexical error correction rate
6.3 Total (grammatical + lexical) error correction rate

Table 7: Formula for calculating error correction rates

Error correction rate
= (Number of error-triggered revisions/ Number of errors) x 100
= [Number of error-triggered revisions/ (Number of error-triggered revisions + Number of uncorrected errors)] x 100

5.11.2.2 Measures of accuracy

Two types of accuracy were investigated in the current study: accuracy during writing and final text accuracy. Table 8 shows the measures of both types of accuracy.

Table 8: Measures of accuracy

Accuracy during writing
1. Lexical errors (per 100 words typed)
2. Grammatical errors (per 100 words typed)
3. Total (grammatical + lexical) errors (per 100 words typed)
Final text accuracy
1. Total (grammatical + lexical) errors (per 100 words in the final text)

L2 accuracy during writing was gauged through lexical, grammatical and total errors made per 100 words typed. The ratio of errors per 100 words has been used in many SLA studies (e.g. Lee, Joo, Moon, & Hong, 2007; Mehnert, 1998). In a comparative study conducted by Inoue (2016), it was found that this ratio measure was the most valid compared to other oft-utilized accuracy measures, percentage of error-free clauses and errors per AS unit, because it aligned best with raters' judgements of accuracy. The three process-based accuracy measures can be investigated in this study due to the utilization of key logging software. To the best of my knowledge, previous research on L2 chat has only assessed accuracy through chat logs and has never explored process-based accuracy. For the purposes of this study, this type of accuracy was investigated because it reveals whether post-task anticipation affects the accuracy of learners' language before revision, i.e. whether they try to be more accurate during transcription.

To calculate the ratios of grammatical and lexical errors per 100 words typed, the total numbers of grammatical and lexical errors were divided by the total number of words typed by participants, and multiplied by 100. The ratio of total errors was computed in the same manner but by taking both types of errors into consideration. As explained in the previous section, the total numbers of grammatical and lexical errors made were obtained by adding the number of error-triggered grammatical and lexical revisions to the number of errors left

uncorrected in each respective category. The number of typed words was generated by Inputlog. The formula for calculating errors per 100 words typed is summarized in Table 9.

Table 9: Formula for calculating errors per 100 words typed

Errors per 100 words typed
= (Number of errors/ Number of words typed) x 100
= [(Number of error-triggered revisions + Number of uncorrected errors)/ Number of words typed] x 100

Apart from accuracy during writing, this study also employed a measure of final text accuracy: total errors per 100 words in the final text. Unlike the previous measures, this measure only concerns errors which can be seen in the final text. To date, there has not been any accuracy measure established specially for analyzing L2 text chat data. Although some researchers have used error-free clauses (Liao & Fu, 2014; Ziegler & Mackey, 2014) and errors per AS unit (Adams & Nik, 2014; Adams et al., 2015; Nik, 2010) to gauge the accuracy of L2 chat logs, determining the boundaries of clauses and AS units in chat logs can be problematic. Regarding clauses, chat messages are usually short, sometimes containing only a word or a short phrase. It is not uncommon to find messages below the clausal level. Therefore, coding for clauses may not be suitable for text chat data. In terms of AS units, Foster, Tonkyn, and Wigglesworth (2000) point out that an AS unit is the speaker's unit of planning and the coding of this unit should be done by taking two indicators of planning into account: pauses and intonation. However, it is impossible to examine pauses in chat scripts and intonation is not applicable to chat. Despite the attempts in previous L2 chat research (Adams & Nik, 2014; Adams et al., 2015; Nik, 2010) to address this problem by considering punctuation as "a stand-in" (Adams et al., 2015, p. 70) for pausing and intonation, there is no robust evidence confirming that punctuation is an indicator of planning in text chat. In addition, sometimes one turn in text chat is split into different messages. For instance, a chat user may simply respond to the interlocutor's suggestion with '*Maybe yes*' before adding '*but I think it looks old*' in a subsequent message. This split makes the boundary of the user's thinking unclear and, in turn, causing difficulties

when coding for AS units. For these reasons, the decision was made to examine final text accuracy in terms of errors per 100 words in the final text, as this measure might be more suitable for determining language accuracy in chat logs. The formula for calculating this frequency is shown in Table 10. Total number of errors in the final text was calculated by adding the number of errors left uncorrected during writing to the number of unsuccessful error revisions, and subtracting the number of errors not appeared in the final text from the result. Then, to calculate the frequency of errors per 100 words in the final text, this number was divided by the total number of words in the final text, and multiplied by 100. As mentioned earlier, the number of words in the final text was counted using the word-count function in *MS Word*.

Table 10: Formula for calculating errors per 100 words in the final text

Errors per 100 words in the final text
= (Number of errors in the final text/ Number of words in the final text) x 100
= {[(Number of uncorrected errors + Number of unsuccessful error revisions) - Number of errors not appeared in the final text]/ Number of words in the final text} x 100

5.11.2.3 Measures of speed fluency

Speed fluency was measured by two real-time measures: minutes per character (including spaces) typed and mean characters between pauses. These are process-based measures that indicate fluency at the time of writing. As with accuracy, to date there has been no standard measure of fluency in L2 chat research. To the best of my knowledge, previous studies have exclusively examined chat logs for fluency by, for instance, focusing on the number of words, turns and words per turn (Adams & Nik, 2014; Nik, 2010), and words per minute spent on task (Liao & Fu, 2014). The problem with analysing chat logs for fluency is that it only allows researchers to gauge product-oriented fluency (e.g. final text quantity, quantity of the final text controlled for time on task, average length of a specified linguistic unit in the final text). Abdel Latif (2013), in his review of writing fluency assessment, questions the validity of product-oriented measures, suggesting that they disregard the

influence of task variables on text quantity and composing rates. For instance, the length of the final text may depend on learners' familiarity with the task topic. Learners who are not familiar with the topic may write less than those with a higher degree of task familiarity. Moreover, product-based measures such as average length of t-units and linguistic features of rhetorical functions may indicate language complexity and/or accuracy of language, rather than fluency. For these reasons, the author proposes, based on previous empirical evidence (Abdel Latif, 2009; Spelman Miller, 2000; van Bruggen, 1946), that the length of translating episodes, which is a process-based measure, is a more valid indicator of fluency than product-based measures. In writing research, the length of translating episodes is often operationalized as pause burst (or *P-burst*), which is defined as "the string of actions delimited by an initial and end pause exceeding the defined pause threshold. The length of a P-burst could be defined in ms [milliseconds] or in characters" (Van Waes & Leijten, 2015, p. 86).

In line with Abdel Latif's proposal, the current research gauged fluency through mean characters between pauses and another real-time fluency measure, minutes per character (including spaces) typed. As opposed to examining the final text, these measures assess writing fluency in real time. Therefore, they are better alternatives than the product-based measures (Abdel Latif, 2009, 2013) used in previous L2 chat research. Minutes per character (including spaces) typed was calculated by dividing the total writing time in minutes excluding pauses by the total number of characters typed. Writing time, the number of characters including spaces, and mean characters between pauses were obtained from Inputlog. The pause threshold was set at 2,000 milliseconds in conformity with common practice in writing research (e.g. Révész, Kourtali, & Mazgutova, 2017; Spelman Miller et al., 2008; Sullivan & Lindgren, 2002).

Apart from product-based measures, I also excluded two other types of writing fluency measures: revision and pausing behaviours (e.g. mean pause length and number of pauses) (Van Waes & Leijten, 2015). Revision might reflect individual writing strategies rather than fluency, and pausing could indicate planning and monitoring, which facilitate writing rather than reflect dysfluency (Abdel Latif, 2013). In addition, pausing behaviours are not

suitable measures for assessing fluency during online chat because, in this medium, pauses are influenced by the time spent reading chat messages, waiting for responses and scrolling. Participants of this study also paused frequently to re-read the task sheet and note down some information.

5.11.2.4 Measure of typing skill

The current study determined participants' typing skills by their typing speed (in words per minute) adjusted for accuracy. This measure was utilized in previous writing research investigating the relationship between keyboard skills and L2 writing (Barkaoui, 2014, 2016; Barkaoui & Knouzi, 2018). Typing test output was examined for the number of correct words typed per minute. Because all participants were instructed to strictly follow the orthographic pattern of the model sentence provided on the instruction sheet (see Section 5.10.2), words containing any deviation from this pattern were regarded as incorrect.

5.11.2.5 Additional measures: Text quantity and stress levels during the main tasks

Apart from the main measures mentioned above, additional measures relating to text quantity and stress were also adopted. Text quantity was investigated in terms of words typed and words in the final text. The number of typed words was obtained from Inputlog. Stress levels during the main tasks were investigated through participants' responses to two five-point Likert items on the exit questionnaire. Ordinal data obtained from the questionnaire were used for analysis.

5.11.2.6 Data analysis procedures

There are two main components of the analysis: the analysis of performance and participants' self-reported data. Both types of data were used in conjunction with each other to answer the research questions.

For the self-reported data, participants' answers to two closed-ended questionnaire items relating to the effect of post-task anticipation on accuracy were counted. They asked participants to rate the degree of influence they thought each type of post-task anticipation had on L2 accuracy during the main task. The results were used to address RQ2. Additionally, the justifications participants wrote on the questionnaire and gave during the follow-up interviews for their answers to these items were also examined for themes, as mentioned in Section 5.11.1.

In terms of performance data, a number of statistical analyses were carried out in SPSS version 23 for Windows. To ensure the validity of the initial coding of the focus of revisions, the codes added based on stimulated recall comments were compared to the initial codes. I first identified instances of revisions where participants could recall why they revised their text during stimulated recall interviews, then examined how each of them had been classified in the initial coding process when only surface structures were analyzed. Of all the grammatical revisions coded the second time, 89 per cent of them were classified in the same category again. The percentage was higher for other categories (content, lexis, reorganization and rephrasing) (Appendix G). This suggests that a high proportion of revisions were correctly coded into each category during the initial coding, confirming its reliability. Hence, I proceeded to use the initial codes in the main analyses.

Prior to the main statistical analyses, correlations between different measures of 1) quantity of revision, 2) error correction rate, 3) accuracy and 4) speed fluency were checked to ensure that each measure investigated different aspects of these constructs. The data were only approximately, not completely, normally distributed.[3] Therefore, Spearman's rank order correlation analyses were used to test these correlations. The results (Appendix H) revealed a very strong and significant correlation between two revision quantity and two accuracy measures (Revision quantity: frequency of covert and all revisions, $r_s = .997$, $p < .001$; Accuracy: total errors made per 100 words typed and final text accuracy, $r_s = .957$, p

[3] Normal distribution was assessed based on skewness and kurtosis z-scores (i.e. skewness and kurtosis scores divided by their standard error), Q-Q plots and histograms. The indicators of non-normal distribution were 1) skewness and kurtosis z-scores outside the range of -2 to 2, 2) deviation of values away from the diagonal line in Q-Q plots and 3) histograms without a bell-shaped curve.

< .001). As the results indicated a high correlation between frequency of covert revisions and total frequency of revisions, further significance and correlation analyses focusing on revision quantity only adopted one of these measures. Total frequency of revisions was chosen over frequency of covert revisions due to its more encompassing nature. Unlike revision quantity, both accuracy measures were retained in all analyses. The reason for this is that one of them (errors made per 100 words typed) is a process-based measure, while the other (errors per 100 words in the final text) is a product-based measure. As they are different in nature, it is warranted to investigate both these measures.

In addition to assessing correlations between different measures of the same construct, I also generated boxplots to identify any outliers in the data for each measure. None of the cases in the data set were extreme, therefore all were included in the main analyses.

To investigate L2 revision behaviours during text chat (**RQ1**), descriptive statistics for revision behaviours were examined. Furthermore, the linguistic unit, trigger and outcome of each type of revision were explored. Follow-up paired sample t-tests were carried out to compare the mean frequency of grammatical to lexical revisions, and the mean correction rate of grammatical to lexical errors in each task (camera and television). The data satisfied the assumptions of the t-tests, with scores containing no significant outliers and being approximately normally distributed. As the results showed a very low mean frequency of revisions made at the punctuation level (M = 0.07, SD = 0.22), rephrasing revisions (M = 0.12, SD = 0.30) and reorganization revisions (M = 0.02, SD = 0.11) per 100 words in the final text, and only a few participants engaged in these types of revisions, these three measures were excluded from further statistical analyses.

Regarding **RQ2**, the effects of post-task anticipation and type of post-task anticipation were gauged using different statistical analyses. First, the effects of post-task anticipation (± post-task anticipation), i.e. group (control vs experimental), were explored through a series of independent-sample t-tests and two MANOVAs; then the effects of type of post-task anticipation (individual vs collaborative post-task anticipation) were investigated using a series of paired-sample t-tests and two repeated-measure MANOVAs.

These MANOVAs were run to check the effects of post-task anticipation and different types of post-task anticipation on revision quantity and fluency. For the analyses relating to revision quantity, two measures of revision quantity (frequency of all revisions and ratio of characters including spaces in the final text to those actually typed) were computed as dependent variables. To explore the effect on fluency, two measures of fluency (minutes per character including spaces typed and mean characters between pauses) were used.

Of the three measures of revision quantity, the frequency of all revisions and the product/ process ratio of characters including spaces were chosen because they correlated with each other when examining both the overall data (r_s = -.761, p < .001) and the data for the experimental group (r_s = -.787, p < .001), and the correlations did not exceed .8 which would otherwise raise a concern over multicollinearity. Frequency of covert revisions was omitted because it correlated too strongly with frequency of all revisions (Overall data: r_s = .997, p < .001; Experimental group data: r_s = .978, p < .001), thereby violating the assumption of no multicollinearity between dependent variables. The two fluency measures correlated with each other with the coefficients being -.776 (p < .001) and -.786 (p < .001) when examining all data and the data for the experimental group, respectively. Overall, the data for these four measures of revision quantity and fluency satisfied all assumptions of MANOVAs.[4]

Although the constructs of error correction rate and accuracy during writing were also assessed by more than one measure, they were excluded from MANOVAs because the data relating to these measures did not meet some assumptions of MANOVAs. In terms of error correction rate, while MANOVAs are suitable when there is approximately a moderate degree of correlation between different measures of the same construct, the rates of

[4] 1) The data of all four dependent variables were approximately normally distributed for each group of the independent variable; 2) there were no multivariate outliers, no multicollinearity between dependent variables, and no significant univariate outliers in each group of the independent variable; 3) in general, there was a linear relationship between each pair of dependent variables for each group of the independent variable; 4) in the case of the MANOVAs testing differences across groups, there was homogeneity of covariance matrices and variance.

grammatical and lexical correction did not correlate with each other (Overall data: r_s = -.056, p = .468; Experimental group data: r_s = -.043, p = .651). As far as accuracy during writing is concerned, when the rates of lexical and grammatical errors were set as dependent variables, three Mahalanobis distance scores calculated from the whole data set and one from the data of the experimental group were found to exceed the critical value of 13.82.[5] This indicated that the MANOVA assumption of no multivariate outliers could be violated. In addition, scatterplots showed that there was not a linear relationship between these accuracy measures for all groups of each predictor variable, group and type of post-task anticipation. Thus, the MANOVA assumption of linearity was not satisfied.

In addition to MANOVAs, a series of t-tests were carried out to investigate the effects of post-task anticipation and type of post-task anticipation on accuracy and all revision measures, apart from revision quantity. As mentioned above, independent-sample t-tests and paired-sample t-tests were used to explore differences across groups (control vs experimental, i.e. ± post-task anticipation) and types of post-task anticipation (individual vs collaborative), respectively. The assumptions of these t-tests were met with the data containing no significant outliers and being approximately normally distributed. Whenever the assumption of homogeneity of variances assumed by independent-sample t-tests was violated, I will report the outcomes, labelled *Equal variances not assumed* by SPSS, which take into account the Satterthwaite adjustment for the degrees of freedom. Further to this, additional analyses were performed as follow-up tests to explore the effects of post-task anticipation and its type on stress during the main tasks. Stress scores were compared across groups using a Mann-Whitney U test and across types of post-task anticipation using a Wilcoxon signed-rank test. These non-parametric tests were utilized because the data were ordinal.

To address **RQ3** and **RQ4**, a series of Spearman's correlation analyses were performed to gauge the relationships between different measures of revision and 1) total errors per 100 words in the final text (RQ3) and 2) typing test scores (RQ4). Following the main analyses

[5] 13.82 is the critical chi-square value for two degrees of freedom at a critical alpha of .001. When there are two predictors in the model, Mahalanobis distance scores larger than this value indicate the presence of multivariate outliers.

addressing RQ3, Spearman's correlation analyses were performed to investigate the relationships between participants' proficiency scores and 1) final text accuracy and 2) the revision measures found to be significantly correlated with final text accuracy. As a follow-up to RQ4, additional Spearman's correlation analyses were carried out to investigate the relationships between typing test scores and accuracy, speed fluency and text quantity. In addition, because L2 proficiency might be confounded with L2 typing ability, a Spearman's correlation analysis was performed to check the correlation between typing test scores and proficiency scores. Because the result revealed a weak and non-significant correlation between typing ability and proficiency ($r_s = .116$, $p = .294$), the influence of proficiency on the results of the analyses relevant to RQ4 can be dismissed. A rank-order Spearman's correlation analysis was chosen to address RQ3 and RQ4 because the data did not satisfy all the assumptions[6] of Pearson's correlation analysis.

In this chapter, I have discussed the methodology of my study. The next chapter will present the results of the research.

[6] 1) There were no outliers in either variable; 2) variables were normally distributed; 3) there was a linear relationship between the variables.

Chapter 6: Results

In this chapter, the results obtained from statistical tests, the exit questionnaire and the follow-up interviews addressing each of the research questions are presented in order. A summary of the results is provided at the end of the chapter. In accordance with Plonsky and Oswald's (2014) benchmarks for the interpretation of *r* and *d* in L2 research, absolute values of the correlation coefficients of .25, .40 and .60 will be interpreted as representing weak, moderate and strong correlation, respectively; absolute values of *d* will be judged differently based on whether the test explored differences between groups or within a group. For between-group comparisons, absolute values of 0.40, 0.70 and 1.00 show a small, medium and large effect size, respectively, while these values are 0.60, 1.00 and 1.40 for within-group comparisons. Eta-squared values (η^2) will be interpreted based on Cohen's (1988) benchmarks, with .01, .06 and .14 indicating a small, moderate and large effect, respectively.

6.1 Research Question 1: L2 revision behaviours during text chat

Table 11 shows descriptive statistics for revision behaviours. The linguistic unit, the trigger and the outcome of each category of revision are presented in Table 12.

Concerning the quantity of revision, descriptive statistics revealed that the mean rate of revisions made per 100 words in the final text was 13.93 (*SD* = 6.08), and a great majority of revisions (*M* = 13.77, *SD* = 6.11) were made covertly. Overt revisions constituted only a very small proportion. The mean ratio of product/ process characters including spaces was 0.77 (*SD* = 0.08).

Of all the types of linguistic units examined, word-level revisions occurred most often, with a mean frequency of 7.42 times per 100 words in the final text (*SD* = 3.73). This was followed by phrasal revisions, which were found to occur less than twice as often as word revisions (*M* = 3.50, *SD* = 2.07). Revisions to letter- and clause-and-above units were made

at mean rates of 1.76 (*SD* = 1.20) and 1.18 (*SD* = 1.05) times, respectively. Punctuation was found to be the unit that participants revised the least (*M* = 0.07, *SD* = 0.22).

In terms of the focus of revisions, a great majority of revisions were content-related (*M* = 9.10 times per 100 words in the final text, *SD* = 4.50). Frequency of grammatical revisions was much lower (*M* = 2.04, *SD* = 1.35), followed by lexical revisions (*M* = 1.53, *SD* = 1.22). The results of follow-up t-tests comparing the mean frequency of grammatical revisions to lexical revisions showed that, in both tasks, there was a significantly higher mean frequency of grammatical compared to lexical revisions (Camera: $t(83) = 4.03$, $p = .001$, BCa 95% CI [0.344, 1.016]; Television: $t(83) = 1.78$, $p = .047$, BCa 95% CI [0.031, 0.719]). Revisions associated with rephrasing (*M* = 0.12, *SD* = 0.30) and reorganization (*M* = 0.02, *SD* = 0.11) were rarely found in the data. For every 100 words in the final text, there were about 1.11 revisions (*SD* = 1) that could not be classified into any category based on their surface structures.

With regard to triggers, the results showed that revisions tended to be error-triggered (*M* = 2.03, *SD* = 1.33) as opposed to non-error-triggered (*M* = 1.52, *SD* = 1.21). When revisions were initiated by an error, the outcome was likely to be successful, as indicated by the high percentage of successful error revisions (*M* = 81.34, *SD* = 23.26). The mean rate of revisions with an unclear trigger was 1.27 times per 100 words in the final text (*SD* = 1.06).

When both grammatical and lexical errors were taken into consideration, the mean rate of error corrections was 15.05 per cent (*SD* = 8.99). When teasing these two types of errors apart, there was a higher mean rate for grammatical error corrections (*M* = 16.14, *SD* = 12.23) than lexical error corrections (*M* = 14.44, *SD* = 15.66). The results of follow-up t-tests comparing the mean correction rate of grammatical to lexical errors revealed that the rates of grammatical and lexical error corrections were not significantly different in either the camera task or the television task (Camera: $t(83) = 0.77$, $p = .451$, BCa 95% CI [-2.437, 5.570]; Television: $t(83) = 0.80$, $p = .438$, BCa 95% CI [-2.404, 5.914]).

Table 11: Descriptive statistics for revision behaviours

	Mean	SD
Quantity		
Covert revisions (per 100 words in the final text)	**13.77**	6.11
All revisions (per 100 words in the final text)	**13.93**	6.08
Characters including spaces in the final text/ typed	0.77	0.08
Linguistic unit (revisions per 100 words in the final text)		
Clause and above	1.18	1.05
Phrase	**3.50**	2.07
Word	**7.42**	3.73
Letter	1.76	1.20
Punctuation	0.07	0.22
Focus (revisions per 100 words in the final text)		
Content	**9.10**	4.50
Grammar	**2.04**	1.35
Lexis	1.53	1.22
Rephrasing	0.12	0.30
Reorganization	0.02	0.11
Unclear	1.11	1.00
Trigger (revisions per 100 words in the final text)		
Error-triggered	**2.03**	1.33
Non-error-triggered	1.52	1.21
Unclear	1.27	1.06
Outcome of error revision (%)		
Successful error revision	**81.34**	23.26
Error correction rate (%)		
Grammatical	**16.14**	12.23
Lexical	**14.44**	15.66
Total (grammatical + lexical)	**15.05**	8.99

Note. Noteworthy figures, e.g. those indicating high proportions or frequencies, are marked in bold.

From Table 12, it can be seen that changes made at the word level were predominant in content, grammatical and lexical revisions, accounting for 47.05, 73.68 and 88.14 per cent of these types of revisions, respectively. For content revisions, the second most common

unit of revisions was phrase, which accounted for 34.18 per cent of total content-related revisions. Smaller proportions of these revisions were made at the clause-and-above level (11.74 per cent), the letter level (6.64 per cent) and the punctuation level (0.38 per cent). In terms of grammatical revisions, letter-level changes were found to account for 20.56 per cent of total revisions, followed by phrasal and clause-and-above changes, which were observed at much lower rates of 4.98 and 0.78 per cent, respectively. Regarding lexical revisions, 10.02 per cent of this type of revision was observed at the phrase level. A minimal number of lexical revisions were made at the letter (1.43 per cent) and the clause-and-above (0.41 per cent) levels.

It was found that for rephrasing and reorganization, participants only made changes to either clause-and-above or phrasal units. The majority (75 per cent) of rephrasing revisions concerned phrasal changes, while most revisions associated with reorganization involved revising clauses or larger units (57.14 per cent). While the difference between the proportion of revisions made to the clause-and-above unit and to the phrasal unit was relatively large for rephrasing (50 per cent difference), the gap found for revisions involving reorganization was much smaller at 14.28 per cent.

Letter revisions seemed to be the most problematic when coding based on surface structures. More than half (64.29 per cent) of the unclear revisions were revisions made at the letter level, and about a quarter (25.71 per cent) were made to the word unit. Smaller percentages were detected for unclear revisions at the phrasal (5.71 per cent), punctuation (3.71 per cent) and clause-and-above (0.57 per cent) levels.

As far as triggers are concerned, it was found that most grammatical revisions were triggered by an error. About 63 per cent of this type of revision were error-triggered, while only 29.13 per cent were non-error-triggered. On the other hand, lexical revisions were mostly non-error-triggered. There was a slightly higher rate of non-error-triggered (48.26 per cent) compared to error-triggered (46.22 per cent) lexical revisions. As, in this study, rephrasing revisions only concern changes made to non-erroneous original text, none of the revisions related to rephrasing were caused by an error. The same pattern was also found

for textual reorganization. A great majority (94.54 per cent) of revisions with an unclear focus could not be classified as to whether they were triggered by an error or not.

Finally, when it comes to the outcomes of error revisions, relatively high proportions of successful revisions were found for both grammatical and lexical error revisions. Of all the revisions targeting grammatical errors, 84.35 per cent yielded a successful outcome. The rate was slightly lower for the success of lexical error revisions at 73.89 per cent.

Table 12: Linguistic unit, trigger and outcome for each type of revision

	Content	Grammar	Lexis	Rephrasing	Reorganization	Unclear
Linguistic unit						
Clause and above	336 (11.74%)	5 (0.78%)	2 (0.41%)	9 (25%)	**4 (57.14%)**	2 (0.57%)
Phrase	**978 (34.18%)**	32 (4.98%)	49 (10.02%)	**27 (75%)**	**3 (42.86%)**	20 (5.71%)
Word	**1346 (47.05%)**	**473 (73.68%)**	**431 (88.14%)**	0	0	90 (25.71%)
Letter	190 (6.64%)	132 (20.56%)	7 (1.43%)	0	0	**225 (64.29%)**
Punctuation	11 (0.38%)	0	0	0	0	13 (3.71%)
Trigger						
Error-triggered		**409 (63.71%)**	**226 (46.22%)**	0	0	9 (2.59%)
Non-error-triggered		187 (29.13%)	**236 (48.26%)**	**36 (100%)**	**7 (100%)**	10 (2.87%)
Unclear		46 (7.17%)	27 (5.52%)	0	0	329 (94.54%)
Outcome of error revision						
Successful		**345 (84.35%)**	**167 (73.89%)**	0	0	8 (88.89%)
Unsuccessful		64 (15.65%)	59 (26.11%)	0	0	1 (11.11%)

Note. Noteworthy figures indicating high proportions or frequencies are marked in bold.

6.2 Research Question 2: Effects of post-task anticipation and type of post-task anticipation on revision behaviours, speed fluency and accuracy

Research Question 2 focuses on the effects of post-task anticipation and type of post-task anticipation on various aspects: revision behaviours, speed fluency and accuracy. The first three parts of this section present the results of statistical tests relating to each of these dimensions. The subsequent part reports the results of follow-up statistical analyses testing the effects of post-task anticipation and its type on stress during the main tasks. The last part of this section discusses the results obtained from the exit questionnaire and follow-up interviews concerning this research question. Descriptive statistics for all the dependent variables arranged by group (control and experimental) and type of post-task anticipation (individual and collaborative) can be found in Appendix J.

6.2.1 Effects of post-task anticipation and type of post-task anticipation on revision behaviours

This section begins by reporting the results concerning revision quantity, then proceeds to those relating to other revision measures. Regarding revision quantity, the results of the MANOVAs testing differences in this construct across groups (F (2, 165) = 1.029, p (Λ) = .359, η^2 = .012, Observed power = 0.228) and types of post-task anticipation (F (2, 54) = 0.023, p (Λ) = .978, η^2 = .001, Observed power = 0.053) were not significant and the effect sizes were small. These indicate no significant difference in terms of revision quantity between the control group and the experimental group or between the individual and the collaborative post-task conditions. Due to the low observed power of these MANOVAs[7], follow-up ANOVAs were performed to test the effects on each measure of revision quantity separately. The results for the comparisons across groups and types of post-task anticipation are reported in Tables 13 and 14, respectively. Again, when each measure of revision quantity was tested separately in ANOVAs, no significant effect of group or type of post-task anticipation was detected and effect sizes were small. Therefore, neither group nor type of post-task anticipation significantly influenced revision quantity.

[7] However, it should be noted that observed power may not represent the true power of the study (Yuan & Maxwell, 2005).

Table 13: Results of ANOVAs testing differences in the quantity of revision across groups

	F	*p*	η^2
All revisions (per 100 words in the final text)	0.220	.640	.001
Characters including spaces in the final text/ typed	0.308	.580	.002

Table 14: Results of ANOVAs testing differences in the quantity of revision across types of post-task anticipation

	F	*p*	η^2
All revisions (per 100 words in the final text)	0.026	.872	.001
Characters including spaces in the final text/ typed	0.046	.831	.001

Turning to the results concerning other revision behaviours, the t-test results in Table 15 illustrate that there were some significant differences across groups. Frequencies of grammatical, lexical and non-error-triggered revisions were significantly higher in the experimental group (Grammatical: $M = 2.27$, $SD = 1.34$; Lexical: $M = 1.66$, $SD = 1.26$; Non-error-triggered: $M = 1.65$, $SD = 1.30$) than in the control group (Grammatical: $M = 1.67$, $SD = 1.25$; Lexical: $M = 1.23$, $SD = 1.11$; Non-error-triggered: $M = 1.20$, $SD = 0.96$). However, the effect sizes were small. In addition, when several statistical tests are carried out using one data set, as was the case with these t-tests, the probability of having Type I errors increases. This means that there is a higher chance of obtaining false positive results, leading to a false assumption of significance differences. Hence, to address this problem, I applied a Bonferroni correction to correct the critical p-value (.05). The p-value of .05 was divided by the number of tests performed (13), yielding a new critical p-value of .004. When observing the data from this new perspective, the outcomes of all tests were non-significant, indicating no effect of group on the linguistic unit, the target, triggers of revisions or the outcomes of error revisions and error correction rates.

Table 15: Results of t-tests testing differences in revision behaviours across groups

	p	*d*	*t*	*df*	BCa 95% CI	
					Lower	Upper
Linguistic unit (revisions per 100 words in the final text)						
Clause and above	.352	0.15	-0.93	163	-0.500	0.175
Phrase	.932	0.01	-0.08	163	-0.667	0.663
Word	.431	0.13	-0.79	163	-1.604	0.790
Letter	.554	0.09	0.59	163	-0.273	0.507
Focus (revisions per 100 words in the final text)						
Content	.445	0.12	0.77	163	-0.881	2.061
Grammar	.006*	0.47	-2.80	163	-1.003	-0.188
Lexis	.029*	0.37	-2.20	163	-0.791	-0.096
Trigger (revisions per 100 words in the final text)						
Error-triggered	.052	0.36	-2.25	163	-0.907	0.038
Non-error-triggered	.012*	0.40	-2.53	142.59	-0.778	-0.167
Outcome of error revision (%)						
Successful error revision	.080	0.30	1.76	163	-0.343	14.547
Error correction rate (%)						
Grammatical	.941	0.01	0.07	163	-4.308	4.868
Lexical	.527	0.10	-0.63	163	-6.663	3.808
Total (grammatical + lexical)	.496	0.12	-0.68	86.25	-4.038	2.252

Note. * = $p < .05$

Table 16 reports the results from the t-tests investigating the effects of type of post-task anticipation on the linguistic unit, the target, triggers of revisions and the outcomes of error revisions and error correction rates. No significant difference was found in any measure and the effect sizes were small, suggesting that learners' revision behaviours under the individual and the collaborative post-task anticipation conditions were comparable.

Table 16: Results of t-tests testing differences in revision behaviours across types of post-task anticipation

	p	*d*	*t*	*df*	BCa 95% CI	
					Lower	Upper
Linguistic unit (revisions per 100 words in the final text)						
Clause and above	.466	0.13	-0.73	52	-0.491	0.280
Phrase	.982	0.00	-0.02	52	-0.678	0.640
Word	.563	0.08	0.58	52	-0.683	1.232
Letter	.431	0.13	-0.79	52	-0.513	0.216
Focus (revisions per 100 words in the final text)						
Content	.554	0.08	-0.60	52	-1.476	0.788
Grammar	.576	0.09	0.56	52	-0.312	0.562
Lexis	.325	0.15	0.99	52	-0.159	0.572
Trigger (revisions per 100 words in the final text)						
Error-triggered	.932	0.02	0.09	52	-0.367	0.433
Non-error-triggered	.159	0.21	1.43	52	-0.104	0.717
Outcome of error revision (%)						
Successful error revision	.271	0.18	1.11	52	-3.024	13.186
Error correction rate (%)						
Grammatical	.459	0.11	0.75	52	-2.040	4.410
Lexical	.559	0.11	-0.59	52	-6.391	3.012
Total (grammatical + lexical)	.614	0.09	0.51	52	-1.992	3.189

6.2.2 Effects of post-task anticipation and type of post-task anticipation on speed fluency

The MANOVA testing the effect of group on speed fluency found a non-significant effect and a small effect size (F (2, 165) = 2.344, p (Λ) = .099, η^2 = .028, Observed power = 0.470). Likewise, a non-significant finding and a small effect size were observed in the outcomes of the MANOVA testing differences in speed fluency across the two types of post-task anticipation (F (2, 54) = 1.003, p (Λ) = .374, η^2 = .036, Observed power = 0.216). Taken together, these suggest that the control group and the experimental group were comparably fluent, and fluency did not differ significantly across post-task anticipation conditions. Because the observed power of the MANOVAs was low, follow-up ANOVAs were run to investigate the effects of group and type of post-task anticipation on each measure of speed fluency separately. The outcomes for the comparisons across groups and types of post-task anticipation are reported in Tables 17 and 18, respectively. All ANOVAs

yielded a non-significant finding and a small effect size, indicating that neither group nor type of post-task anticipation had an impact on any speed fluency measure. Hence, the non-significant effects of group and type of post-task anticipation found from the MANOVAs were confirmed.

Table 17: Results of ANOVAs testing differences in speed fluency across groups

	F	*p*	η^2
Minutes/character (including spaces) typed	1.001	.318	.006
Mean characters between pauses	4.300	.060	.025

Table 18: Results of ANOVAs testing differences in speed fluency across types of post-task anticipation

	F	*p*	η^2
Minutes/character (including spaces) typed	1.571	.215	.028
Mean characters between pauses	0.000	.987	.000

6.2.3 Effects of post-task anticipation and type of post-task anticipation on accuracy

Tables 19 and 20 report the results of the t-tests investigating the effects of group and type of post-task anticipation on different accuracy measures, including the measures for both accuracy during writing and final text accuracy. It can be seen that none of the findings were significant and all effect sizes were small. Thus, there was no significant difference across groups (control vs experimental) or types of post-task anticipation (individual vs collaborative) regarding accuracy.

Table 19: Results of t-tests testing differences in accuracy across groups

	p	*d*	*t*	*df*	BCa 95% CI	
					Lower	Upper
Accuracy during writing (errors per 100 words typed)						
Lexical errors	.078	0.29	-1.77	166	-1.234	0.048
Grammatical errors	.053	0.31	-1.97	166	-2.446	0.200
Total (grammatical + lexical) errors	.051	0.35	-2.23	166	-2.505	0.039
Final text accuracy						
Total (grammatical + lexical) errors (per 100 words in the final text)	.070	0.29	-1.82	166	-3.387	0.283

Table 20: Results of t-tests testing differences in accuracy across types of post-task anticipation

	p	*d*	*t*	*df*	BCa 95% CI	
					Lower	Upper
Accuracy during writing (errors per 100 words typed)						
Lexical errors	.662	0.07	-0.44	55	-0.770	0.488
Grammatical errors	.083	0.16	-1.76	55	-1.157	0.026
Total (grammatical + lexical) errors	.099	0.15	-1.68	55	-1.476	0.119
Final text accuracy						
Total (grammatical + lexical) errors (per 100 words in the final text)	.063	0.20	-2.18	55	-1.956	0.090

6.2.4 Follow-up analyses: Effects of post-task anticipation and type of post-task anticipation on stress during the main tasks

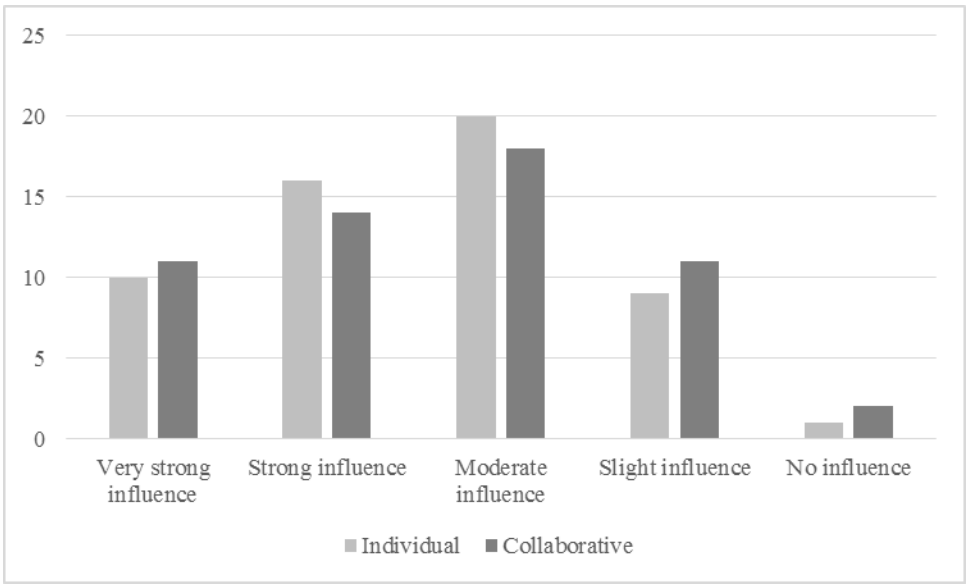
The Mann-Whitney U test comparing stress levels during the main tasks of the control group to those of the experimental group yielded a non-significant result and a small effect size (U = 3619.00, p = .079, z = 1.76, r = .14). These findings were also observed in the outcomes of the Wilcoxon signed-rank test exploring the effect of type of post-task anticipation on stress (p = .317, z = -1.00, r = -.13). Therefore, neither group nor type of post-task anticipation significantly affected stress levels during the main tasks.

6.2.5 Results of exit questionnaire and follow-up interviews: Perceived effect of post-task anticipation on accuracy

This section reports the results obtained from the exit questionnaire responses for the items related to the perceived effect of post-task anticipation on accuracy during the main task, and from the interviews conducted to elaborate and clarify participants' answers to these items.

Figure 10 shows the questionnaire results. The perceived effect of both types of post-task anticipation (i.e. individual and collaborative) on accuracy followed similar trends. For both types of anticipation, the most frequently reported perceived effects were moderate and strong, respectively. Only a small number of participants thought these post-tasks had no impact on them.

Figure 10: Questionnaire responses on the effect of post-task anticipation on accuracy during the main tasks



Out of 56 participants in the experimental group, 20 reported that the anticipation of an individual language correction post-task moderately influenced them to try to use language more accurately during the main task. Sixteen thought the influence was strong, and 10 and nine participants reported a very strong and a slight degree of influence, respectively. Only

one person thought that the anticipation of an individual post-task had no impact on language accuracy.

Regarding the effect of anticipation of a collaborative post-task on accuracy, 18 participants in the experimental group perceived a moderate influence, and 14 a strong influence. Both very strong and slight levels of impact were reported by 11 participants, and only two thought there was no effect of collaborative post-task anticipation.

The reasons participants gave on the questionnaire and during the follow-up interviews to justify their answers regarding the effect of individual and collaborative post-task anticipation on accuracy are summarized in Table 21. The reasons are grouped based on whether they are in support of the effectiveness of post-task anticipation in encouraging L2 accuracy during the main task performance.

Similar reasons were reported for both individual and collaborative post-task anticipation. Thirty-one responses reported that the knowledge of an individual (N = 14) and a collaborative (N = 17) post-task encouraged more accurate use of English during the main task, because there would be fewer corrections to make in the post-task stage if the logs were already accurate. Another reason given to support the effectiveness of post-task anticipation is associated with avoiding encountering one's own errors. In eight responses, participants reported that they would feel incompetent if they saw their L2 errors while performing the individual (N = 4) and collaborative (N = 4) language correction post-tasks. Therefore, they tried to be as accurate as possible during the main tasks. Apart from these reasons, six responses concerned the provision of assistance to peers. They revealed that some participants avoided making errors during the main tasks because they wanted their partner to understand their messages in the chat logs when completing the individual (N = 2) and collaborative (N = 4) post-tasks. Five participants thought that the anticipation of a collaborative post-task encouraged accuracy because they would feel embarrassed if their partner saw their errors during the post-task. Only one participant, however, referred to embarrassment in the response pertaining to the effectiveness of individual post-task anticipation. Besides the reasons mentioned thus far, 12 participants also reported that the

lack of peer assistance during the individual language correction post-task promoted L2 accuracy during the main task. When they knew they would have to perform the language correction post-task alone, with no support from their partner, they tried to use more accurate language while chatting.

Regarding the reasons not supporting the effectiveness of post-task anticipation, participants reported that the anticipation of individual (N = 7) and collaborative (N = 6) post-tasks might not foster L2 accuracy because they had the predisposition to be accurate, regardless of whether they were anticipating a post-task. The opportunity for error corrections during the individual (N = 7) and collaborative (N = 5) post-tasks was also reported to have a negative effect on accuracy. Responses revealed that some participants were not much concerned about accuracy during the main tasks because they could correct their language mistakes during the post-tasks.

In addition to this, 11 responses referred to the nature of chat language. Participants reported that they thought chat was an informal form of communication in which informal and inaccurate language was tolerated. For them, message comprehensibility was more important than accuracy when chatting. Hence, even with post-task anticipation, they were not very concerned about L2 accuracy. Eight and three responses, respectively, mentioned the lack of threat from the post-tasks and the time constraints of the main tasks. As far as threat is concerned, some participants thought that the individual (N = 4) and collaborative (N = 4) language correction post-tasks were not threatening enough to them, and that they did not worry about or pay attention to the post-tasks while performing the main tasks. With regard to the time constraints, participants reported that they were so preoccupied with finishing the main task in time that they did not think about the post-task.

Table 21: Justifications for questionnaire answers relating to the perceived effects of individual and collaborative post-task anticipation on accuracy during the main tasks

Reasons	N (Ind)	N (Col)	Samples
Effective			
Less work during post-task	14	17	*I tried to make as few errors as possible during the main task so that both of us had fewer errors to correct in the post-task.*
Lack of peer assistance during post-task	12	-	*I tried to be accurate when knowing that I would have to perform an individual language correction post-task afterwards, because I knew I would not get any help from my friend during the post-task. In this post-task, I had to use my own English knowledge. I think it motivated me to be more accurate than the collaborative one.*
Avoiding encountering one's own errors	4	4	*I wanted to use English accurately during the main task, because I didn't want to encounter my own errors during the post-task. It made me feel bad about myself.*
Assisting peers' post-task transactions	2	4	*I tried to use language as accurately as I could and describe everything in as much detail as possible so that my chat partner could understand me during the main task, and could make sense of my language when making corrections during the post-task.*
Embarrassment	1	5	*Because I knew about the post-task, I tried to be more accurate [during the main task]. It would be embarrassing if my friend saw my English errors during the post-task.*

Note. Ind = anticipation of an individual post-task; Col = anticipation of a collaborative post-task

(Continued)

Table 21: Continued

Reasons	N (Ind)	N (Col)	Samples
Not effective			
Predisposition to be accurate despite post-task anticipation	7	6	*I always try to use language accurately when I chat online. It doesn't matter whether I knew about the post-task or not. So, it didn't really affect me.*
Opportunity for error corrections during post-task	7	5	*I didn't worry too much about not being accurate [during the main task], because I knew I would have the opportunity to correct my mistakes again in the post-task.*
Attitudes towards chat language	5	6	*I didn't try to be more accurate. I think chatting is all about getting your messages across. I don't think accuracy is very important during chatting as long as your partner can understand you.*
No threat from post-task	4	4	*A post-task is not like a test. I was not stressed or worried about it. So, it did not affect me that much. I was not so concerned about having to type everything accurately.*
Time constraints	3	-	*Because I rushed to complete the main task, I did not have time to think about the post-task while performing the main task.*

Note. Ind = anticipation of an individual post-task; Col = anticipation of a collaborative post-task

Interestingly, during the follow-up interviews, eight participants mentioned that the relationship between interlocutors might affect the effectiveness of collaborative post-task anticipation in terms of encouraging L2 accuracy. They all reported that if they had not been familiar with their partner, they would have been more careful not to make any errors during the main task. The reason for this is because they thought it was embarrassing and awkward having to discuss their errors with a stranger or someone they did not know well during the collaborative post-task. Moreover, one participant reported that if she had not known her partner well enough to know how proficient he/she was in English, she might have tried to use English more accurately during the main task. When she was not familiar with her partner's English skills, she was not certain how much the person could help her during the collaborative language correction post-task. Thus, she thought it would be better to make sure that the language in the logs was as accurate as possible from the start.

6.3 Research Question 3: Relationship between L2 revision behaviours and text accuracy in chat scripts

Table 22 provides Spearman's correlation coefficients showing the relationship between L2 revision and accuracy of the final text, as measured by total grammatical and lexical errors per 100 words in the final text. Descriptive statistics for L2 revision and final text accuracy can be found in Table 11, Section 6.1, and Appendix J, respectively.

The results presented in Table 22 revealed that all measures of quantity, linguistic unit, focus, triggers of revisions and outcomes of error revisions had a weak correlation with final text accuracy ($-.4 < r_s < .4$). Of these measures, significant correlations were detected for the relationships between final text accuracy and four measures: frequency of revisions ($r_s = .170$, $p = .027$), frequency of word-level revisions ($r_s = .176$, $p = .023$), frequency of error-triggered revisions ($r_s = .172$, $p = .027$) and percentage of successful error revisions ($r_s = -.248$, $p = .001$). Overall, total errors per 100 words in the final text positively and weakly correlated with frequency of revisions, and frequency of all types of linguistic units, foci and revision triggers. However, total errors per 100 words in the final text correlated negatively and weakly with the ratio of product/ process characters, including spaces, and the percentage of successful error revisions. Taken together, these results show that

decreased accuracy in the final text could be linked to higher quantity of revisions in general and to less success in error revisions, although the correlations were weak.

Table 22: Results of Spearman's correlation analyses investigating the relationship between revision and final text accuracy (N = 168)

Quantity	Correlation coefficient	*p* (2-tailed)
All revisions (per 100 words in the final text)	.170*	.027
Characters including spaces in the final text/typed	-.097	.210
Linguistic unit (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Clause and above	.082	.292
Phrase	.063	.417
Word	.176*	.023
Letter	.131	.090
Focus (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Content	.097	.209
Grammar	.140	.070
Lexis	.150	.052
Trigger (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Error-triggered	.172*	.027
Non-error-triggered	.131	.093
Outcome of error revision (%)	Correlation coefficient	*p* (2-tailed)
Successful error revision	-.248**	.001
Error correction rate (%)	Correlation coefficient	*p* (2-tailed)
Grammatical	-.446**	.000
Lexical	-.015	.850
Total (grammatical + lexical)	-.435**	.000

Note. * Correlation is significant at the .05 level (2-tailed).
** Correlation is significant at the .01 level (2-tailed).

Regarding error correction rates, while final text accuracy correlated weakly with the rate of lexical error corrections (r_s = -.015), the correlations were moderate and significant where grammatical (r_s = -.446, p < .001) and total (r_s = -.435, p < .001) error correction rates were concerned. A negative relationship was detected between the measure of final

text accuracy (frequency of errors in the final text) and all types of error correction rates, suggesting a connection between reduced final text accuracy and lower rates of error corrections.

Table 23 provides the results of follow-up Spearman's correlation analyses investigating the relationships between proficiency scores and 1) final text accuracy and 2) the revision measures previously found to be significantly correlated with final text accuracy. The mean proficiency score was 60.15 (SD = 5.89).

Table 23: Results of Spearman's correlation analyses investigating the relationships between proficiency scores and 1) final text accuracy and 2) revision measures found to be significantly correlated with final text accuracy (N = 168)

Final text accuracy	Correlation coefficient	*p* (2-tailed)
Total (grammatical + lexical) errors (per 100 words in the final text)	-.363**	.000
Quantity (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
All revision	.024	.764
Linguistic unit (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Word	-.008	.920
Trigger (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Error-triggered	-.029	.710
Outcome of error revision (%)	Correlation coefficient	*p* (2-tailed)
Successful error revision	.216**	.005
Error correction rate (%)	Correlation coefficient	*p* (2-tailed)
Grammatical	.202**	.009
Total (grammatical + lexical)	.192*	.013

Note. * Correlation is significant at the .05 level (2-tailed).
** Correlation is significant at the .01 level (2-tailed).

The results show that proficiency scores were significantly and negatively correlated with the rate of errors in the final text, indicating that higher proficiency was related to increased final text accuracy (r_s = -.363, p < .001). This relationship was weak, as indicated by the

value of the coefficient. Regarding the revision measures, significant correlations were observed between proficiency and 1) the outcome of error revisions (r_s = .216, p = .005) and 2) grammatical (r_s = .202, p = .009) and total (r_s = .192, p = .013) error correction rates. The correlation coefficients showed positive relationships, suggesting that higher proficiency was linked to greater success in error revisions and higher grammatical and total error correction rates. However, these correlations were weak. As far as revision frequencies were concerned, frequencies of total (r_s = .024), word-level (r_s = -.008) and error-triggered (r_s = -.029) revisions were only weakly and not significantly related to proficiency levels.

6.4 Research Question 4: Relationship between learners' typing ability and L2 revision in text chat

Spearman's correlation coefficients showing the relationships between different measures of L2 revision and typing test scores (i.e. correct words typed per minute during the typing test) are reported in Table 24. Descriptive statistics for revision can be found in Table 11 in Section 6.1. With regard to typing test scores, the mean score was 42.98 (SD = 9.64).

In Table 24 it can be seen that all coefficient values were between -.4 and .4, suggesting a weak correlation between typing test ability and all measures of L2 revision. Among these revision measures, typing test scores correlated significantly with the frequency of lexical revisions (r_s = .163, p = .034) and grammatical (r_s = .252, p = .001), lexical (r_s = .170, p = .029) and total (r_s = .272, p < .001) error correction rates. These coefficient values were positive, indicating that higher typing ability was weakly linked to higher frequency of lexical revisions and higher rates of grammatical, lexical and total error corrections.

Table 24: Results of Spearman's correlation analyses investigating the relationship between revision and typing test scores (N = 168)

Quantity	Correlation coefficient	*p* (2-tailed)
All revisions (per 100 words in the final text)	.091	.238
Characters including spaces in the final text/typed	-.097	.210
Linguistic unit (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Clause and above	-.056	.467
Phrase	.142	.067
Word	.097	.213
Letter	-.068	.381
Focus (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Content	.086	.270
Grammar	.090	.247
Lexis	.163*	.034
Trigger (revisions per 100 words in the final text)	Correlation coefficient	*p* (2-tailed)
Error-triggered	.086	.272
Non-error-triggered	.086	.272
Outcome of error revision (%)	Correlation coefficient	*p* (2-tailed)
Successful error revision	-.021	.789
Error correction rate (%)	Correlation coefficient	*p* (2-tailed)
Grammatical	.252**	.001
Lexical	.170*	.029
Total (grammatical + lexical)	.272**	.000

Note. * Correlation is significant at the .05 level (2-tailed).
** Correlation is significant at the .01 level (2-tailed).

As a follow-up to the above statistical tests, additional Spearman's correlation analyses were carried out to investigate the relationships between typing test scores and accuracy, speed fluency and text quantity. The results are presented in Table 25 below. Descriptive statistics for the data concerning accuracy, speed fluency and text quantity are shown in Appendix J.

Table 25: Results of Spearman's correlation analyses investigating the relationships between 1) accuracy, 2) speed fluency, 3) text quantity and typing test scores (N = 168)

Accuracy during writing (errors per 100 words typed)	Correlation coefficient	*p* (2-tailed)
Lexical errors	-.061	.431
Grammatical errors	-.316**	.000
Total (grammatical + lexical) errors	-.283**	.000
Final text accuracy	Correlation coefficient	*p* (2-tailed)
Total (grammatical + lexical) errors (per 100 words in the final text)	-.301**	.000
Speed fluency	Correlation coefficient	*p* (2-tailed)
Minutes/character (including spaces) typed	-.495**	.000
Mean characters between pauses	.460**	.000
Text quantity	Correlation coefficient	*p* (2-tailed)
Total words typed	.477**	.000
Total words in the final text	.437**	.000

Note. ** Correlation is significant at the .01 level (2-tailed).

The results revealed moderate and significant correlations between typing test scores and 1) speed fluency and 2) text quantity. For speed fluency, typing test scores had a negative correlation with the number of minutes per character (r_s = -.495, p < .001) and a positive correlation with the number of characters between pauses (r_s = .460, p < .001). These results indicated that higher typing ability was moderately associated with increased speed fluency. As far as text quantity is concerned, there was a positive relationship between typing test scores and both the number of typed words (r_s = .477, p < .001) and the number of words in the final text (r_s = .437, p < .001). Therefore, better typing skills were moderately associated with greater quantity of text being produced and text in chat logs.

Regarding accuracy, there was a negative relationship between typing test scores and all accuracy measures, i.e. frequencies of errors in chat logs (r_s = -.301, p < .001) and lexical (r_s = -.061), grammatical (r_s = -.316, p < .001) and total errors (r_s = -.283, p < .001) made, indicating a link between better typing ability and increased accuracy, both in terms of final text accuracy and accuracy as typed. However, these correlations were weak. Of all

measures of accuracy, the frequency of lexical errors made was the only measure that did not correlate significantly with typing test scores.

6.5 Summary of results

RQ1: L2 revision behaviours during text chat

This study found that most revisions in the data were made covertly during drafting. Of all the linguistic units examined, word-level revisions occurred most often, followed by revisions to phrase, letter, clause-and-above and punctuation units, respectively. Phrasal revisions were found to occur less than twice as often as word revisions. In terms of the focus of revisions, a great majority of revisions were content-related. Grammatical revisions, the second most common type of revision, occurred at a much lower rate, followed by lexical revisions. There was a significantly higher mean frequency of grammatical revisions compared to lexical revisions in both tasks. Revisions associated with rephrasing and reorganization were rarely found in the data. With respect to the triggers of revisions, this study explored revision triggers by disregarding content revisions and found that revisions tended to be error-triggered. Error revisions were usually carried out successfully, with a high success rate of approximately 81 per cent. The mean rate of grammatical error corrections was higher than that for lexical error corrections, although this difference was not statistically significant in either of the tasks.

When examining each type of revision separately, it was found that word-level changes were predominant in content, grammatical and lexical revisions. The majority of revisions associated with rephrasing and reorganization concerned changes to phrasal and clause-and-above units, respectively. While most grammatical revisions were error-triggered, there was a slightly higher rate of non-error-triggered compared to error-triggered lexical revisions. Finally, relatively high proportions of grammatical and lexical error revisions were carried out successfully, with the success rates being 84.35 and 73.89 per cent, respectively.

RQ2: Effects of post-task anticipation and type of post-task anticipation on revision behaviours, speed fluency and accuracy

The results of statistical tests showed that revision behaviours, speed fluency and accuracy did not differ significantly across groups (control vs experimental) or types of post-task anticipation (individual vs collaborative). Similarly, the follow-up analyses did not reveal any significant effect of group or type of post-task anticipation on stress levels during the main tasks.

The questionnaire results regarding the perceived effect of post-task anticipation on accuracy showed similar trends for both types of anticipation (individual and collaborative), with the two most frequently reported effects being a moderate and a strong effect, respectively. Only a small number of participants thought these post-tasks had no impact on them. The most common reason given to support the effectiveness of post-task anticipation in fostering accuracy concerns participants' attempts to minimize their workload during the post-task stage. Other reasons relate to participants' avoiding encountering their own errors in the post-tasks, their attempts to provide assistance to their peers and their fear of embarrassment. For the individual language correction post-task, the lack of peer assistance was also reported to promote L2 accuracy during the main task. Regarding the reasons not supporting the effectiveness of post-task anticipation, some participants reported their predisposition to be accurate despite post-task anticipation. Other reasons concern the opportunity for error corrections during the post-tasks, learners' attitudes towards chat, the lack of threat from the post-tasks and the time constraints of the main tasks. Interestingly, some participants reported during the follow-up interviews that the relationship between chat partners might affect the effectiveness of collaborative post-task anticipation in encouraging L2 accuracy.

RQ3: Relationship between L2 revision behaviours and text accuracy in chat scripts

All measures of the quantity, linguistic unit, focus and triggers of revisions, and outcomes of error revisions, were found to have a weak correlation with final text accuracy. Weak significant relationships were observed between decreased final text accuracy and 1) higher

frequencies of total, word-level and error-triggered revisions, and 2) a lower percentage of successful error revisions. While a weak and non-significant correlation was found between final text accuracy and the rate of lexical error corrections, correlations were moderate and significant where grammatical and total error correction rates were concerned. Overall, the results suggest that higher frequency of errors in the final text was associated with lower error correction rates.

Follow-up analyses investigating the relationships between proficiency and revision measures previously found to be significantly correlated with final text accuracy revealed weak significant correlations between higher proficiency and 1) greater success in error revisions and 2) higher grammatical and total error correction rates. Frequencies of total, word-level and error-triggered revisions were only weakly and not significantly related to proficiency scores. Finally, higher proficiency was significantly and weakly related to increased final text accuracy.

RQ4: Relationship between learners' typing ability and L2 revision in text chat

Correlation analyses revealed a weak correlation between typing test ability and all measures of L2 revision. Higher typing ability correlated weakly and significantly with higher frequency of lexical revisions and higher rates of all types of error corrections.

Follow-up correlation analyses investigating the relationships between typing test scores and accuracy, speed fluency and text quantity revealed moderate and significant correlations between typing test scores and 1) speed fluency and 2) text quantity. Higher typing ability was found to be moderately associated with increased speed fluency and quantity of text. With respect to accuracy, a weak significant correlation was observed between typing ability and 1) frequencies of grammatical and total errors made and 2) final text accuracy. Overall, there was evidence that better typing ability was linked to more accurate language production.

Chapter 7: Discussion

The previous chapter reported the results of the current study. In this chapter, the findings related to each research question are discussed. Each section begins with a brief description of the data collection and analysis procedures related to each research question, and then proceeds to a discussion of the findings. The chapter ends with a summary of the discussion.

7.1 Research Question 1: L2 revision behaviours during text chat

The first research question of the study sought to explore the L2 revision behaviours of Thai learners of English with A2 to B2 proficiency levels during text chat. This research question was investigated through the analysis of descriptive statistics of revision behaviours. The linguistic unit, trigger and outcome of each type of revision were also explored. Follow-up paired sample t-tests were carried out to compare the mean frequency of grammatical to lexical revisions, and the mean correction rate of grammatical to lexical errors in each task (camera and television). Data were collected by means of two online chat tasks involving picture description and decision-making. These tasks were performed via Skype. During the tasks, participants' computer screens and keystrokes were recorded by *Flashback Express 5* (2017) and *Inputlog 7* (Leijten & Van Waes, 2013), respectively. Keystroke logs were coded in Atlas.ti 8 (2017) data analysis software, and statistical analyses were carried out in SPSS version 23 for Windows. To ensure the reliability of the initial coding of the focus of revisions, initial codes were compared to codes added based on stimulated recall comments.

The analysis yielded two main positive findings that suggest the potential of text chat for L2 learning. First, the results revealed a high mean percentage of successful error revisions (81.34 per cent). Participants of this study were not allowed to use any online resource. Thus, the high successful error correction rate indicates that text chat might allow sufficient time for learners with upper-elementary to upper-intermediate levels to reflect on their erroneous output and access their L2 knowledge when making corrections. Based on Hayes et al.'s (1987) cognitive model of L1 revision processes, this finding could mean that during

online chat interaction, these learners are likely to successfully diagnose problems detected in the text they write and choose appropriate ways to rectify them. Indeed, compared to the rate of successful language error corrections found by Stevenson et al. (2006), who explored revisions in individual computer-assisted writing (62 per cent), the percentage found in this study is much higher. This is surprising, because participants in both studies were of comparable proficiency levels and individual computer-assisted writing is not interactive in nature. The non-interactivity of individual computer-assisted writing does not place as much time pressure on learners as text chat interactions, and in turn potentially allows more time for reflection on L2 output. It could be that as the participants in Stevenson et al.'s (2006) study were required to think aloud during writing, verbalization might have interfered with their writing and cognitive processes, leading to a reduced rate of successful error corrections. Overall, the results of my study show that, despite the time pressure imposed by chat, when learners engage in form-related revisions in online chat, they can successfully correct errors in their output.

The second encouraging finding was the high total revision frequency. On average, participants revised 13.93 times per 100 words in the final text, or approximately once for every 10 words in the chat scripts. Considering that typographical and spelling revisions were excluded, this frequency is relatively high, indicating that learners pay close attention to their language use during chatting. The mean frequency of changes detected in this study is more than double that found by Smith (2008) in his L2 chat study, i.e. 6.1 times per 100 words in the final text. The disparity between Smith's study and mine could be due to the different types of tasks utilized in the research. While this study applied tasks that combined picture description with decision-making, Smith (2008) used picture-ordering tasks. In that study, most tasks involved one learner watching a video clip and another learner viewing eight stills from it. As both learners could see the same scenes in their task materials, there was no need for accurate and detailed description. As a result, this may have led to fewer revisions compared to my study in which each participant received different pictures and was required to describe them to their partner. It is possible that picture description tasks encourage L2 revisions, as learners might have difficulty in selecting the right words and language structures to describe pictures to their partner, especially when their partner

cannot see those pictures. Many participants reported such difficulty when answering the questionnaire questions about task perception. For instance, a participant reported that "I doubt that my friend could imagine what the televisions look like. They were difficult to describe, and I haven't even seen them myself." Another reported "The tasks were challenging. I tried to be accurate and very detailed when describing the photos and had to edit my messages frequently to make sure that I included everything." Hence, decisions made regarding task choice could be important in both research and pedagogical contexts when the goal is to foster learners' L2 revisions and attention to their own output during chat. This corresponds to the cognitive models of L1 writing processes (Chenoweth & Hayes, 2001; Hayes, 1996, 2012; Hayes & Flower, 1980), which assume that writing and revision are influenced by task environment, such as task materials and writing topics. Information about the task environment guides writing goals and writing plans, which in turn may determine how text production is carried out (Hayes & Flower, 1980).

Comparing the total frequency of changes found in this study with that of L2 speech research, the mean frequency is much higher in this study than reported in L2 speech production contexts. In Kormos' (2000, 2003) L2 speech repair research, the mean frequency of changes was less than four times per 100 words, which is less than a third of the frequency detected in this study. Taking this together with the results of previous research that found significantly more L2 self-corrections in chat than in speaking (Lai & Zhao, 2006; Zeng, 2017), this suggests greater potential of text chat to encourage attention to language form compared to speaking. Yet, my study found that the overall rate of corrected errors of participants, who were upper-elementary to upper-intermediate L2 learners, was 15.05 per cent. This is not higher than the figures found in L2 speech repair research (Kormos, 2000, 2003), i.e. 20.39 and 21.05 per cent for pre-intermediate and 16.21 and 11.76 per cent for upper-intermediate students. Therefore, it is possible that repairs occur more frequently in text chat than oral communication not because of the potential of chat to encourage attention to form, but because this type of communication does not involve the same turn-taking mechanism as oral communication. During oral communication each interlocutor typically waits until their partner finishes their utterance before speaking. In contrast, in online chat, all participants can compose their text

simultaneously and privately in their own dialogue boxes. While learners are typing, they might receive a new message or question from their partner which can trigger a change in the content of the message being typed. For instance:

1.) {7255}[RSHIFT]A•has•many•legs{35256}[BACK][RSHIFT]D•also•did•not

2.) and•on•the•top[BACK]{7071}well•yer[BACK]s•it•is[RETURN]

In both examples, while participants were composing a text to describe their photos (1. *A has many legs* and 2. *and on the top*), they were interrupted by a new message/ question from their partner (1. *but b didn't have it* and 2. *so it isn't a TV?*) and decided to delete all the text they had typed to respond to that message/ question (with 1. *D also did not* and 2. *well yes it is*). It was evident from the data that content-related changes carried out due to this reason occurred very frequently and, therefore, they could be the cause of the higher total frequency of changes found in this study in comparison with speech repair research. While the current study found that participants made content-related revisions as often as 9.10 times per 100 words, Kormos' (2000, 2003) speech repair studies found different-information repairs to occur only at 2.02 and 0.47 times per 100 words for pre-intermediate learners and 0.41 and 0.48 times per 100 words for upper-intermediate learners.

When comparing the frequency of revisions observed in this study to that detected in Choi's (2007) computer-assisted L2 writing research, the figure in this study is much lower. In Choi's study, it was found that, on average, learners made changes, excluding those related to typography and orthography such as spelling and space corrections, a total of over 29 times per 100 words in the final output when composing argumentative essays alone. Even though this might indicate fewer L2 revisions in text chat than in individual computer-assisted writing, the mean product/ process ratio of characters found in the present study is comparable to other individual computer-assisted L2 writing research. In my study, the ratio is 0.77, similar to 0.74 found in Révész, Michel, and Lee's (2017) study, and to 0.74

and 0.73 found in the simple and complex task conditions, respectively, in Révész, Kourtali, and Mazgutova's (2017) research. Hence, it is still inconclusive whether the quantity of L2 revisions in text chat is greater or lower than that of individual computer-assisted writing. While Révész and colleagues' studies were conducted with advanced learners, Choi's research included learners at both higher- and lower-proficiency levels. Based on the findings of research that detected the effects of L2 proficiency on revision behaviours (Barkaoui, 2016; Choi, 2007; Manchón et al., 2009), it may be that the influence of proficiency on revision contributes to the inconsistency discussed above.

Apart from the two main positive findings discussed above, the results regarding frequencies and triggers of grammatical and lexical revisions, error correction rates and outcomes of error revisions reveal that text chat may be more suitable for promoting the development of grammatical structures than lexical development.

First, with respect to frequencies of grammatical and lexical revisions, in both the camera and television tasks there was a significantly higher mean frequency of grammatical revisions compared to lexical revisions. These results indicate that students might pay more attention to grammatical structures than lexical items while performing these tasks via chat. The findings are in line with Smith's (2008, 2009) results, which showed that participants made more grammatical than lexical repairs in L2 online chat.

Second, regarding the triggers of grammatical and lexical revisions and error correction rates, the findings provide evidence that text chat may be more suitable for encouraging learners' noticing and corrections of grammatical mistakes than lexical ones. Although the rates of grammatical and lexical error corrections were not significantly different in either the camera task or the television task, the mean rate of grammatical error corrections (16.14 per cent) was higher than that of lexical error corrections (14.44 per cent). In addition, more error-triggered grammatical revisions (N = 409) were found compared to lexical ones (N = 226). While 63.71 per cent of grammatical revisions were error-triggered, this percentage was lower at 46.22 for lexical revisions. Interestingly, the findings of the current study differ from the common findings of speaking research, which indicate more lexical than

grammatical error repairs (Fathman, 1980; Lennon, 1984) and a higher lexical than grammatical error correction rate (Kormos, 2000). It can be hypothesized that learners tend to pay more attention to their grammatical errors than lexical errors during chat, while attending more to mistakes related to lexis than grammar in speech. Alternatively, it could be that the increased time pressure during oral communication and the nature of the tasks utilized contribute to these differences. First, the time pressure during the oral mode is higher than that during online chat. Due to this time pressure, learners may choose not to repair grammatical errors when their output is still understandable in speech. Furthermore, the time pressure during oral communication may hinder lexical retrieval, which leads to erroneous lexical items being frequently retrieved and, hence, repairs of lexical errors are frequent. With respect to the nature of the tasks, it is possible that the role-playing (Kormos, 2000), interview (Fathman, 1980) and storytelling (Fathman, 1980; Lennon, 1984) tasks administered in previous speech research predisposed learners to focus on lexical items rather than on grammatical structures, while the tasks employed in this study had the opposite effect.

Third, with regard to the outcome of error revisions, although the success rates for both grammatical (84.35 per cent) and lexical (73.89 per cent) error corrections were high, the rate of successful grammatical error revisions was higher than that of lexical ones. This reveals that learners may be able to correct their grammatical mistakes more successfully than lexical mistakes during chat. The results contradict the findings of Golonka et al. (2017), who found a higher success rate for self-initiated lexical error corrections (100 per cent) than for grammatical error corrections (47 per cent) when observing a series of information-gap, reasoning and opinion-giving L2 tasks performed via text chat. However, because in Golenka et al.'s (2017) study corrections were only examined based on chat logs, the findings of my study may yield more accurate insights into error correction processes in text chat.

In terms of the focus of revisions, the main findings reveal that a great majority of revisions were content-related (M = 9.10 times per 100 words in the final text) and the frequencies of other types of revisions were much lower (Grammatical, M = 2.04; Lexical, M = 1.53;

Rephrasing, M = 0.12; Reorganization, M = 0.02). Although typographical and spelling revisions were excluded from the analysis, it was obvious when examining the data that there was a prevalence of typographical/ spelling revisions. As the results of my study reveal that these revisions and revisions related to content can frequently occur at the same time, they seem to contradict the hypothesis of Whalen and Ménard (1995). Whalen and Ménard (1995) compared L1 and L2 revision and found more revisions relating to linguistic form during L2 writing and more pragmatic and textual revisions during L1 writing. Based on these findings, they assumed that L2 linguistic processing might hinder processing at more global levels of discourse, e.g. ideas generation and content organization. It may be that in my study typographical/ spelling revisions did not interfere substantially with high-level processing, and that the L2 learners in my research were able to allocate their attention to both typographical/ spelling mistakes and content at the same time during online chat.

The high occurrence of typographical revisions is in line with the common findings of individual computer-assisted L2 writing research (Barkaoui, 2016; New, 1999). According to Barkaoui (2016), there are three possible explanations for the high occurrence of typographical revisions. First, learners, specifically those with higher typing skills, may type fast when writing on a computer under time pressure, and as a result make many typographical errors which necessitate revision. Because the tasks in Barkaoui's (2016), New's (1999) and this study were timed, this explanation is plausible. Second, L2 learners who have low typing skills may make many typographical errors and, therefore, engage in many typographical revisions. Third, drawing on Stevenson et al.'s (2006) hypothesis, Barkaoui points out that L2 orthographic processing may be non-automatic and slow, which increases the likelihood of making typographical errors and, consequently, typographical revisions. The last two reasons may be especially true for Thai learners who might not be familiar with English keyboards and have non-automatic English orthographic processing.

As far as content-related revisions are concerned, during stimulated recall, there were many instances when participants reported that they opted to change the content completely because they did not know the lexical item that describes the product component they were writing about. Sometimes, they changed their content because they were not sure whether

the grammatical structures in the text they were producing were correct or not. This may explain the high frequency of content-related revisions observed in this study. For instance, one participant, being uncertain whether the verb form she had used was correct, changed the word *keep* in *The lense can be keep* to *kept*. Later, as she was still unsure about the accuracy of the new verb form, she decided to change the content of her text completely, i.e. from *The lense can be kept inside but if* to *you can pull the lense out for usage*.

> The•lense•can•be•{7348}keep[BACK][BACK][BACK][BACK]kept•inside•but•if•{2309}[BACK]you•can•pull•the•lense•out•for•usage.

During the interview, she recalled that "I changed [from *keep* to *kept*] because I remembered that 'be' should be followed by a past participle, but then I deleted the whole thing because what I remembered could be wrong. I was not so sure about it." The example above illustrates the complexity of L2 chat revisions and how the utilization of verbal protocols can reveal this complexity. In my research, learners' verbalizations show that meaning-related revisions can indeed be triggered by learners' attention to form.

Although, as discussed above, online chat might have the potential to promote the development of L2 grammatical structures and some meaning-related changes might be triggered by learners' attention to form, the high proportion of content revisions compared to form-related revisions found in this research shows that learners' attention may primarily be drawn to meaning during chat. The prevalence of content revisions detected in my study is in contrast with much higher rates of language-related over content-related revisions observed in individual computer-assisted L2 writing studies (Barkaoui, 2016; Choi, 2007; Stevenson et al., 2006). This difference might be attributed to the highly interactive nature of chat, which may encourage increased attention to meaning compared to individual computer-assisted writing. Unlike traditional writing, interaction in text chat involves real-time communication, therefore revisions in this medium can be affected by communicative factors. For instance, as discussed earlier, chat users may decide to change the content of their text based on new information obtained from their interlocutor while they are typing. Similarly, they may delete the whole message they are composing in the dialogue box in

order to reply to a new question just received from their interlocutor. Revisions made for these reasons were very common when coding the data. In addition, the lower frequency of form-related revisions compared to content revisions observed might be explained by the relatively rapid speed of chat communication, which may predispose learners to attend to meaning rather than form. An additional factor that could account for the results relates to how language use is influenced by the interlocutor's status and age in Thailand. When Thai interlocutors have the same status (here students) and are of approximately the same age, as was the case with the participants in this study, communication tends to be casual and require less precision. This could, therefore, lead to reduced concern over the accuracy of linguistic form during peer-peer online chat interaction. Lastly, the reduced concern over form may be attributed to learners' attitudes towards online chat language. Many participants reported that they thought chat was an informal type of communication in which reduced language accuracy was acceptable. For instance, when asked about the importance of accuracy in online chat communication, one participant reported that "I usually chat with my friends. I don't think my language needs to be very accurate because chatting is supposed to be casual…I'm more concerned about typing fast [than being accurate], because I don't want my friend to wait long for my responses." Similarly, another participant said "Only my friend and I can see our chat log, so I don't think it's important to use language accurately. It's enough just to make my messages understandable."

In terms of the nature of revisions, the findings show that of all the revisions examined in this study, about 99 per cent (i.e. 13.77 times per 100 words in the final text) were made covertly. Only a very small proportion of revisions were made overtly. The high rate of covert revisions is consistent with Smith (2008), who found significantly more revisions in screen-recording data than in chat logs. Therefore, the findings lend support to Smith's (2008) call for the use of screen capture technology to investigate learners' drafting processes during L2 chat and his call for researchers to abandon their reliance on chat logs when analyzing L2 chat interactions.

Finally, as regards the linguistic unit of revisions, this study found that most revisions occurred at the word level, followed by phrase, letter and clause-and-above units,

respectively. Revisions made to punctuation were scarce. Similar to the overall findings, when examining each revision category separately, most types of revisions were mainly made at lower-level linguistic units. In other words, most content, grammatical and lexical revisions were made at the word level, and the majority of rephrasing revisions concerned phrasal changes. Interestingly, although content changes usually entail extensive revisions, the results revealed that only 11.74 per cent of them were made at the clause level and above. Most content revisions were word-level revisions, such as changing '*it*' to '*picture*' to make the meaning more specific. Reorganization of text was the only type of revision that was mostly done to higher-level linguistic units, i.e. clauses or larger units. These results are in line with my hypothesis and the common findings of individual computer-assisted L2 writing studies which have found more revisions made to lower linguistic units (e.g. below-word, word and phrasal levels) than to higher ones (e.g. clausal, sentential and multi-sentential levels) (Barkaoui, 2016; Choi, 2007; Li, 2006; New, 1999; Révész, Michel, & Lee, 2017; Stevenson et al., 2006). The predominance of revisions made to lower-level linguistic units observed in this study indicates that, when chatting, learners' attention may be limited to the part of text they are typing. The results suggest a limitation of text chat in fostering learners' attention to their overall output. While chatting, the focus of students' attention can be restricted to the letter, word or phrase being typed, instead of the whole clause or sentence. As the findings of this research resemble those of individual computer-assisted L2 writing studies, it could be that the nature of computer-assisted writing promotes local revisions. Other plausible explanations are that the relatively high speed of chat interaction limits the time learners have available to revise at the level of larger linguistic units; and the prevalence of revisions to lower-level linguistic units was found because chat messages are usually short, sometimes containing only a single word or phrase.

Thus far, I have discussed the findings related to RQ1. The following sections will discuss the results pertaining to RQs 2–4.

7.2 Research Question 2: Effects of post-task anticipation and type of post-task anticipation on revision behaviours, speed fluency and accuracy

The second research question enquired into the effects of post-task anticipation and type of post-task anticipation on revision behaviours, speed fluency and accuracy. These effects were gauged using the t-tests and MANOVAs that investigated whether the scores for these dependent variables differed significantly across groups (control vs experimental) or types of post-task anticipation (individual vs collaborative). Two additional analyses, Mann-Whitney U and Wilcoxon signed-rank tests, were also carried out as follow-ups to explore the effects of group and type of post-task anticipation on stress during the main tasks. Apart from the performance data, the results of the questionnaire and interviews relating to the perceived effect of post-task anticipation on accuracy were also used to address RQ2.

The results revealed no significant effect of group or type of post-task anticipation on fluency. These findings indicate that post-task anticipation and type of post-task anticipation did not influence fluency. This is consistent with the findings of both Skehan and Foster (1997) and Foster and Skehan (2013), who did not find a clear influence of post-task anticipation on speaking fluency. Even though this research explored post-task anticipation in a different medium to these previous studies, it yielded the same result, suggesting that post-task anticipation might not be related to fluency regardless of the task medium.

As with speed fluency, neither group nor type of post-task anticipation had an impact on learners' revision behaviours as investigated through the linguistic unit, target, triggers and quantity of revisions, the outcome of error revisions and error correction rates. Furthermore, these two variables did not significantly affect accuracy during writing or final text accuracy. The non-significant effects of group suggest that awareness of an upcoming language correction post-task might not encourage learners to be accurate or to revise differently compared to when they are not aware of the post-task. Therefore, they do not offer support for Skehan's (1998) and Foster and Skehan's (2013) hypotheses that assume that L2 learners' awareness of a post-task may redirect their focus from fluency to accuracy during the main task and create a condition in which learners monitor their performance more closely.

The findings of this study relating to accuracy are in line with Skehan and Foster's (1997) findings, which did not show a clear effect of post-task anticipation on accuracy. However, they contradict the results of Foster and Skehan (2013), who found that post-task anticipation led to a significant increase in accuracy in all tasks examined. This inconsistency is surprising because, as in Foster and Skehan's (2013) research, the tasks used in this study were carefully designed by considering the degree of threat posed to learners and the emphasis of the post-tasks on linguistic accuracy (see Chapter 5, Section 5.4). The inconsistencies between the findings of my study and Foster and Skehan's (2013) research could be attributed to the research setting. While Foster and Skehan's (2013) data collection took place during regular classroom sessions, my study was carried out in an experimental setting. It may be that an experimental research setting places greater pressure on participants than a more naturalistic classroom context. In my study, the awareness of being observed by the researcher during the main tasks may have led all participants, regardless of their post-task anticipation conditions (i.e. no, individual and collaborative post-task conditions), to pay a similar level of attention to accuracy, resulting in non-significant effects of group and type of post-task anticipation.

Based on the exit questionnaire and interview responses, the explanation for the non-significant effects of group and type of post-task anticipation on revision and accuracy could be related to the opportunity for error corrections during post-tasks, participants' attitudes towards chat language, the lack of threat from the post-tasks and time constraints. First, as mentioned in the previous chapter, some participants reported that they were not too concerned about accuracy during the main tasks because they could correct their language mistakes during the individual and collaborative post-tasks. Hence, neither of these language correction post-tasks may promote accuracy and revision. Rather than language correction post-tasks, transcription post-tasks that do not involve post-task language corrections (cf. Foster & Skehan, 2013) could be better alternatives for enhancing L2 revision behaviour and accuracy during the main task. Second, the qualitative data showed that some participants viewed chat as an informal type of communication in which informal and inaccurate language was acceptable. Owing to this, even with post-task anticipation, learners may not have prioritized L2 accuracy when performing tasks via

online chat. Third, it is possible that neither of the language correction post-tasks was threatening enough to yield significant effects and, as a result, participants did not pay attention to the demands of the upcoming tasks while performing the main tasks. Indeed, stress levels during the main task were found to be comparable across groups and types of post-task anticipation, which lends support to this explanation. Importantly, participants' interview responses revealed that the threat of the collaborative language correction post-task may be diminished by participants' familiarity with their partner. In an environment where students are not familiar with their partner, they might be more careful not to make errors during the main task for fear of embarrassment of having to discuss their errors with someone they do not know well during the collaborative language correction post-task. If the collaborative post-task had been utilized in such a context, it might have resulted in a significant increase in accuracy and the number of revisions compared to the individual post-task. Finally, regarding time constraints, because the main tasks of this study were timed, participants may have been so preoccupied with finishing the main tasks in time that they did not think about the individual or collaborative post-task. Alternatively, even if participants did think about these post-tasks, they may not have had enough time to revise their text or correct their L2 errors due to the limited time allowed. Indeed, for both types of post-task anticipation, most participants perceived a moderate and strong influence of post-task awareness in encouraging them to be more accurate during the main task, and only a few thought these post-tasks had no impact on them. Thus, it is possible that although their awareness of the post-tasks motivated them to be accurate, they did not have enough time to rectify their mistakes.

7.3 Research Question 3: Relationship between L2 revision behaviours and text accuracy in chat scripts

The third research question asked about how L2 revision behaviours were related to text accuracy in chat scripts. To address this question, a series of Spearman's correlation analyses were performed to gauge the relationship between different measures of revision and total errors per 100 words in the final text. Following the main analyses, Spearman's correlation analyses were performed to investigate the relationships between participants'

proficiency scores and 1) final text accuracy and 2) the revision measures found to significantly correlate with final text accuracy.

The main analyses found no significant correlation between final text accuracy and the focus of revisions. A weak significant correlation was observed between final text accuracy and frequency of total revisions. Perhaps because most revisions were made at the word level, the frequency of this subcategory of revisions also weakly and significantly correlated with final text accuracy. These results show that decreased accuracy in the final text could be linked to higher frequency of revisions. When learners revised more, errors occurred more frequently in the final text. In addition, higher frequency of error-triggered revisions weakly and significantly correlated with decreased final text accuracy. Follow-up analyses found that frequencies of total, word-level and error-triggered revisions were not related to L2 proficiency. It is still not clear what could cause the significant relationships between these revision frequencies and final text accuracy. It may be that the significant findings can be attributed to learners' lack of familiarity with the task content. In the questionnaire and interviews, a number of participants reported that they had difficulties choosing the words and language structures to describe the old product designs they had never seen before. For example, one participant wrote in the exit questionnaire that "Both the camera and television tasks were difficult. I have never seen those products before. It took me a very long time to think of the words to describe their components." When learners are not familiar with task content, they may not be able to use language elements relevant to the task accurately, and revise more due to their doubts about their own language use.

While this study found some significant relationships between accuracy and revision, Révész, Michel, and Lee's (2017) individual computer-assisted writing study examined the data of learners with high L2 proficiency and detected no significant relationship between accuracy, as measured by errors per 100 words in the text, and the product/ process word or character ratio or frequency of revisions made to any linguistic unit. In their study, all participants were international students from a university in the UK who were already familiar with the task utilized in the study – a computer-based IELTS academic writing

test. Therefore, based on my hypothesis above, it may be that the familiarity with the task among all participants contributed to the non-significant correlations observed between accuracy and revision behaviours in their research.

Similar to the findings related to revision frequencies, this study observed a weak significant correlation between final text accuracy and the percentage of successful error revisions. Increased accuracy in the final text weakly related to higher success rates of error revisions. The result offers support for Stevenson et al.'s (2006) hypothesis, which assumes that the outcome of revisions might be important for the improvement of text quality. The results of the follow-up analyses showed significant correlations between higher proficiency and 1) greater error revision success and 2) increased final text accuracy. It is possible that proficiency explains the significant relationship between final text accuracy and success of error revisions. Learners with higher L2 proficiency are likely to correct errors more successfully and are able to produce a more accurate text.

Compared to revision frequencies and the successful outcome of error revisions, the rates of error corrections were found to be stronger predictors of final text accuracy. This study found that, of all the revision measures, the rates of grammatical and total error corrections were the strongest predictors of final text accuracy. Increased accuracy in the final text significantly and moderately correlated with higher grammatical and total error correction rates. As with the rate of successful error revisions, the results of follow-up analyses showed that error correction rates were significantly related to proficiency. Higher grammatical and total error correction rates significantly correlated with higher proficiency. Drawing on Kellogg's (1996) cognitive model of L1 writing processes, which highlights the importance of working memory in every writing process, it is possible that learners with high L2 proficiency, whose language processing is highly automatized, have more attentional resources available to monitor their language use while writing compared to those with low proficiency. Due to this and their broader linguistic knowledge, learners with higher proficiency may be able to detect more errors, leading to higher error correction rates and therefore more accurate final text.

Interestingly, while final text accuracy correlated moderately and significantly with the grammatical error correction rate, the correlation was weak and non-significant where the rate of lexical error corrections was concerned. This can be explained by the higher number of grammatical errors made during chat (M = 17.30, SD = 7.95) compared to lexical errors (M = 9.61, SD = 5.09). Because most of the errors that participants made related to grammar, it is not surprising that final text accuracy, operationalized as the frequency of total errors in the final text, correlated more with the grammatical than the lexical error correction rate.

7.4 Research Question 4: Relationship between learners' typing ability and L2 revision in text chat

The last research question asked how learners' typing ability was related to L2 revision in text chat. This question was investigated through a series of Spearman's correlation analyses that gauged the relationship between measures of revision and typing test scores. In addition, follow-up Spearman's correlation analyses were carried out to investigate the relationships between typing test scores and accuracy, speed fluency and text quantity.

The main analyses revealed that most revision measures were not significantly correlated with typing ability. Typing scores were not related to error revision success, how frequent learners revise, or the triggers and linguistic unit of revisions. Significant correlations were only found between typing ability and 1) frequency of lexical revisions and 2) error correction rates, with higher typing ability related to a higher frequency of lexical revisions and higher error correction rates. However, all of these correlations were weak.

Regarding the result concerning the outcome of error revisions, error revision success may not be related to typing ability as much as L2 proficiency. In other words, while typing ability does not predict the outcome of error revisions, learners with higher proficiency and more language knowledge may be able to correct their errors with greater success. From the follow-up analysis for RQ3, it was found that higher proficiency significantly correlated with greater error revision success. In addition, in Stevenson et al.'s (2006) study on

individual computer-assisted L2 writing, the rate of successful error corrections in L1 was higher than that in L2, lending support to this assumption.

A possible explanation for the non-significant relationships between typing ability and 1) quantity of revision and 2) frequencies of revisions made due to different types of triggers and to different linguistic units is that the nature of L2 revision may be influenced by complex interactions between a host of factors, but cannot be predicted by one specific factor – typing ability alone. Indeed, cognitive models of L1 writing processes (Chenoweth & Hayes, 2001; Hayes, 1996, 2012; Hayes & Flower, 1980) assume the influence of several factors on L1 writing and revision, e.g. task materials, composing medium, the writer's beliefs about writing and working memory capacity. Although the findings of this research showed significant relationships between high typing ability and high frequency of lexical revisions and error correction rates, the correlations were weak. Overall, the findings of this research are, to some extent, consistent with the findings of Barkaoui (2016), who revealed that L2 revision behaviours (i.e. the quantity, temporal location, action and target of revisions) of participants with low and high typing ability were not significantly different. Because Barkaoui's research was done in an individual computer-assisted writing context, it is possible that L2 revision during online chat and during individual computer-assisted writing have similar relationships with typing ability.

Although weak, the significant positive correlations observed between typing test scores and the rates of lexical, grammatical and total error corrections suggest a possible advantage to having good L2 typing skills. According to Kellogg's (1996) cognitive model of L1 writing processes, the writer's working memory resources are allocated to multiple writing processes which can operate simultaneously during writing. When the writer has poor motor skills, e.g. typing skills, the execution process will demand more cognitive resources (Kellogg, 1996), resulting in fewer cognitive resources being available for other processes, such as ideas generation, language processing and language monitoring. Based on the findings of my research, it can be hypothesized that when L2 learners have increased typing ability, they have more cognitive resources available to attend to their lexical and grammatical errors, leading to increased rates of error detection and correction.

Moreover, the benefit of having good L2 typing ability is also revealed by the results of follow-up analyses, which found that higher typing ability significantly, although weakly, correlated with increased final text accuracy and accuracy during writing, as measured by frequencies of grammatical and total errors as typed. As mentioned above, increased typing ability may free up the cognitive resources required to attend to L2 output. As a result, learners with better typing ability may be able to engage more in internal language monitoring prior to L2 inscription, resulting in lower rates of errors as typed and, in turn, more accurate final output.

Compared to the findings relating to the frequency of lexical revisions, error correction rates and accuracy, stronger relationships were observed between typing ability and 1) speed fluency and 2) text quantity in the follow-up tests. Higher typing ability significantly and moderately correlated with increased speed fluency and higher quantity of text produced and text in chat logs. Since typing scores were calculated based on learners' ability to type a given text quickly, it is not surprising that learners with higher typing scores could type more fluently and produce a greater quantity of text. The findings pertaining to text quantity are, again, in line with the finding of Barkaoui's (2016) individual computer-assisted L2 writing study, which found a significantly greater length of text produced by learners in a high typing ability group compared to that generated by a low-level group.

7.5 Summary

The results of my research make several contributions to the understanding of L2 writing processes in a computer-assisted context. First, the frequent occurrences of both typographical/ spelling and content-related revisions observed indicate that, in the context of my research, typographical/ spelling mistakes do not interfere with high-level processing during computer-assisted writing. It is possible that while writing on a computer, learners with A2 to B2 proficiency levels can attend to typographical/ spelling errors and content at the same time. Second, the predominance of revisions carried out to low-level linguistic units is in line with the findings of individual computer-assisted L2 writing research, indicating that computer-assisted writing may restrict learners' attention to lower linguistic

units, such as the letter, word or phrase being typed. Content revisions, which usually entail extensive changes, as well as grammatical and lexical revisions, were found to occur mostly to lower-level linguistic units in this study. Third, the significant relationships found between better L2 typing ability and 1) higher error correction rates and 2) increased accuracy suggest that learners can benefit from having good L2 typing skills when writing L2 text on a computer. Due to their automatized typing processes, learners with high L2 typing ability might have more resources available to engage in internal language monitoring and attend to language errors. Fourth, the non-significant correlations found between L2 typing ability and most online chat revision measures are in line with the findings of previous individual computer-assisted L2 writing research, which found no significant effect of typing ability on revision behaviours. Based on these findings and the hypothesis previously made vis-à-vis L1 writing, I speculated that revision during computer-assisted L2 writing might be influenced by complex interactions between a host of factors, not just typing ability. A summary of the significant relationships found between L2 typing ability and different measures examined can be found in Table 26.

Table 26: Significant correlations between typing ability (as measured by typing scores) and different measures investigated

Revision	**Strength of relationship**	**Direction of relationship**
Frequency of lexical revisions	Weak	+
Grammatical, lexical and total error correction rates	Weak	+
Accuracy	**Strength of relationship**	**Direction of relationship**
Frequencies of total and grammatical errors as typed	Weak	-
Frequency of total errors in the final text	Weak	-
Speed fluency	**Strength of relationship**	**Direction of relationship**
Minutes/character (including spaces) typed	Moderate	-
Mean characters between pauses	Moderate	+
Text quantity	**Strength of relationship**	**Direction of relationship**
Words typed	Moderate	+
Words in the final text	Moderate	+

Note. + = positive relationship; - = negative relationship

Regarding L2 online chat, this study found a high rate of error revision success, indicating that chat might allow sufficient time for learners with A2 to B2 proficiency levels to diagnose problems during error corrections. The high frequency of total revisions detected also suggests that learners pay close attention to their L2 in this medium. However, because there was a much higher frequency of content-related revisions than language form revisions, learners' attention may be mainly drawn to the meaning-related aspect of language during online chat. This contradicts the claims commonly made regarding the potential of this medium for promoting learners' attention to language form. Adding to previous L2 online chat research that found a higher frequency of grammatical repairs than lexical repairs, my study found evidence that this medium might be more suitable for promoting the development of grammatical structures than lexical development. There were more frequent revisions of grammar than lexis, more error-triggered grammatical revisions than error-triggered lexical revisions, and higher rates of corrections and successful corrections of grammatical errors than lexical errors.

In terms of post-task anticipation, comparing the findings of this study to previous research utilizing oral production main tasks, I hypothesized that post-task anticipation might not have an effect on L2 learners' fluency regardless of the task medium. In addition, the results of this study suggest that post-task anticipation may not influence revision or accuracy in text chat. While Foster and Skehan (2013) hypothesized the importance of the threat of post-tasks and their emphasis on language correctness, in the discussion I argue that learners' relationship can influence this threat and that time constraints of the main task also contribute to the effectiveness of post-task anticipation in promoting attention to form during the main task.

As regards the relationship between L2 revision and final text accuracy, the findings of this study support Stevenson et al.'s (2006) hypothesis about a possible link between the outcome of revisions and text quality, with a significant correlation observed between greater error revision success and increased accuracy in the final text. Of all the revision measures, the strongest predictors of final text accuracy were correction rates of errors. The

significant correlations between final text accuracy and different revision measures examined are summarized in Table 27.

Table 27: Significant correlations between final text accuracy (as measured by frequency of errors in the final text) and revision measures

Revision	Strength of relationship	Direction of relationship
Frequencies of total, word-level and error-triggered revisions	Weak	+
Error revision success	Weak	-
Grammatical and total error correction rates	Moderate	-

Note. + = positive relationship; - = negative relationship

Chapter 8: Conclusion

In the previous chapter, the findings of my research were discussed. This chapter will conclude the thesis by presenting a summary of the findings of each research question, and discussing the methodological and theoretical contributions made and the pedagogical implications of this research. At the end of the chapter, the limitations of this study are considered and suggestions for future research are made.

8.1 Summary of findings

RQ1: What are the L2 revision behaviours of Thai learners of English with A2 to B2 proficiency levels during text chat?

This study revealed a high frequency of total revisions, suggesting that learners pay close attention to their L2 output in this medium. Comparing the finding to that of previous research in individual computer-assisted L2 writing, it is still difficult to conclude whether L2 revision quantity in text chat is different from that in individual computer-assisted writing. However, comparisons with L2 speech research are possible and my findings in this regard show a much higher mean frequency of revisions compared to the mean frequency of repairs detected in L2 speech research. This difference might be due to the difference in the turn-taking mechanism across modes. While interlocutors usually take turns speaking, all chat users can compose their messages simultaneously in their own dialogue boxes. A new message received from a chat partner may trigger the revision of a message being typed. Apart from the high total revision frequency, another important finding of my study is the very high rate of success in error revision. This indicates that, despite the time pressure during chat, participants can successfully draw on their L2 knowledge to improve form-related errors in their output. The amount of time available for L2 production during chat may be sufficient for learners with A2 to B2 proficiency levels to reflect on L2 errors and access their linguistic resources when making corrections.

Overall, the findings of this study regarding frequencies and triggers of grammatical and lexical revisions, the outcome of error revisions and error correction rates reveal that text chat might be more suitable for fostering grammatical than lexical development. There was evidence suggesting that, in this medium, learners pay more attention to grammatical than lexical features, notice and correct more grammatical errors compared to lexical ones, and tend to correct grammatical errors more successfully than lexical errors. The results concerning the frequency of grammatical and lexical revisions confirm the findings of Smith's (2008, 2009) research, which found that participants made more grammatical than lexical repairs during L2 online chat. Interestingly, the findings of the current study regarding the triggers of grammatical and lexical changes and error correction rates differ from the common findings of research on L2 speaking. Studies on L2 speaking have revealed more repairs to lexical than grammatical errors (Fathman, 1980; Lennon, 1984) and a higher correction rate for lexical than grammatical errors (Kormos, 2000). One can hypothesize that learners are more likely to attend to their grammatical mistakes than lexical ones during chat, but more to errors related to lexis than grammar during speech. Alternatively, it is possible that these differences are caused by the increased time pressure during oral communication compared to that during online chat. Due to the high time pressure in the oral mode, learners may choose not to repair grammatical errors when their output is still understandable. This time pressure may also hinder lexical retrieval during speech production, leading to frequent lexical errors, which necessitate repair. Moreover, the differences across modes may be attributed to the nature of the tasks utilized. It may be that the tasks administered in previous speech research predisposed learners to attend to lexical items rather than grammatical items, while the tasks in this research had the opposite effect.

Regarding the focus of revisions, a great majority of revisions were content-related and typographical/ spelling revisions. Since these types of revisions frequently occurred at the same time, the findings do not support Whalen and Ménard's (1995) hypothesis, which assumes that L2 linguistic processing hinders processing at more global levels of discourse. In the context of this study, typographical/ spelling revisions might not substantially interfere with high-level (content) processing, and L2 learners' attention can be directed to

typographical and spelling errors and content simultaneously during chat. The frequent occurrence of typographical revisions observed in this research is in line with the findings of individual computer-assisted L2 writing studies. Based on Barkaoui's (2016) assumption, it may be that time pressure during computer-assisted writing, low typing skills and non-automatic L2 orthographic processing account for the frequent occurrence of this type of revision.

Although students may pay attention to grammatical items and revise frequently during chat, the dominance of content-related over form-related revisions found in the current research suggests that learners may pay attention primarily to the meaning-related aspect of language rather than to form. Comparing the findings of this study to those of individual computer-assisted L2 writing research, one can assume that the interactive nature and relatively rapid speed of chat promote increased attention to meaning compared to individual computer-assisted writing. In addition, the lower frequency of form-related revisions compared to content revisions is perhaps a characteristic of peer-peer online chat among Thai learners, or this could be attributed to a learners' attitude that views language with reduced accuracy as acceptable in online chat.

With regard to the nature of revisions, this study found that most changes (about 99 per cent) were made during the drafting process, rather than to the already-sent text. The data concerning the linguistic unit of revisions show that when chatting, learners' attention may only be limited to the letter, word or phrase they are typing. Most grammatical, lexical and content revisions were made to low-level linguistic units. Because these findings resemble those of individual computer-assisted L2 writing research, it is possible that computer-assisted writing promotes local revisions. Alternatively, it could be that learners revise only short stretches of text because their chat messages tend to be short, or the relatively high speed of chat interaction limits the time available for revising at the level of larger units.

RQ2: What are the effects of post-task anticipation and type of post-task anticipation on revision behaviours, speed fluency and accuracy?

No significant effect of post-task anticipation or type of post-task anticipation was observed on any measure: speed fluency, revision behaviours or accuracy. As the results concerning speed fluency are consistent with the findings of previous studies conducted using oral production tasks, post-task anticipation may not be related to fluency, regardless of the task medium. The findings suggest that post-task anticipation might not significantly influence L2 accuracy and revision. Therefore, they do not offer support for Skehan's (1998) and Foster and Skehan's (2013) hypotheses that assume that L2 learners' awareness of a post-task may redirect their focus from fluency to accuracy during the main task, and encourage them to monitor their performance more closely. However, I speculated that there could be other reasons that contribute to the non-significant effects observed on accuracy and revision. These include reasons related to the experimental research setting, the time constraints of the main tasks, participants' attitudes towards chat language, and the nature and threat of language correction tasks.

RQ3: How are L2 revision behaviours related to text accuracy in chat scripts?

This study found that final text accuracy was related to revision frequencies, the rate of successful error revisions and error correction rates, the last of which being the strongest predictors of final text accuracy. The significant correlation observed between greater error revision success and increased accuracy in the final text offers support for Stevenson et al.'s (2006) hypothesis about a possible link between the outcome of revisions and text quality. Based on the results of the follow-up analyses, it was assumed that proficiency influences final text accuracy and L2 revision in terms of error revision success and error correction rates. Learners with higher L2 proficiency potentially have the ability to detect more errors, correct errors more successfully and produce more accurate text.

RQ4: How is learners' typing ability related to L2 revision in text chat?

There was only partial evidence of a relationship between typing ability and L2 revision. Most revision measures did not significantly correlate with typing ability, and all the significant relationships detected were weak. These findings are generally consistent with

the findings of Barkaoui's (2016) individual computer-assisted L2 writing research, indicating that L2 revision during online chat and individual computer-assisted writing may have similar relationships with typing ability. It was hypothesized based on the results of the follow-up analysis and the findings from individual computer-assisted L2 writing research that the success of error revisions depends more on L2 proficiency than on typing ability. Based on the cognitive models of L1 writing processes which assume the influence of several factors on writing and revision (Chenoweth & Hayes, 2001; Hayes, 1996, 2012; Hayes & Flower, 1980), the nature of L2 revision may be influenced by complex interactions between a plethora of factors and not determined only by typing ability. An important implication of the significant relationships observed between typing scores and 1) error correction rates and 2) accuracy is that L2 learners with better typing ability may have more cognitive resources available to attend to L2 output, leading to increased detection and correction of their linguistic errors and increased internal L2 monitoring. This assumption corresponds with Kellogg's (1996) hypothesis, which assumes the influence of the writer's motor skills on the cognitive load during writing processes. Because, in this study, the typing test assessed learners' ability to type a given text quickly, the significant correlations observed between high typing ability and increased 1) speed fluency and 2) text quantity were not unexpected.

8.2 Methodological contributions

As mentioned in Chapter 5, Section 5.2, this research project was motivated by several methodological shortcomings of previous studies exploring L2 revisions during text chat. There are three main methodological contributions of this research to the field. First, this project utilized a novel data collection method − keystroke logging − which to the best of my knowledge has not previously been used in this area of research. While many previous studies within this domain have relied on chat logs as the only data source, my research adopted other data collection methods which obtain information about learners' drafting process, including computer screen recording and keystroke logging. Indeed, a very high proportion of covert revisions was found in this study, justifying the investigation of L2 revision during the message drafting process. In addition to allowing researchers to explore

revisions made during drafting, the data from keystroke logs can also be used to gauge process-based accuracy and fluency, as explained in Chapter 5, Sections 5.11.2.2 and 5.11.2.3. To the best of my knowledge, previous L2 chat studies have only explored fluency and accuracy by examining chat logs and employing product-based measures.

Second, another methodological contribution of this study is that it is the first L2 online chat revision research that takes learners' self-reports into account. I utilized verbal protocols (stimulated recall interviews) to gauge learners' underlying communicative intention when revising in online chat and to ensure the accurate classification of types of error corrections and revisions. Although stimulated recall interviews were not conducted on all task performances due to time constraints, this study demonstrates the usefulness of verbal protocols as complementary tools that can be used to confirm the reliability of the coding of L2 chat revisions prior to conducting an analysis. In addition, as discussed in Chapter 7, Section 7.1, this study found that learners' verbalization can uncover the complexity of L2 chat revisions, showing the benefit of adopting verbal protocols when exploring L2 chat revisions.

Finally, this research addresses a problem observed in previous online chat studies that mainly investigated L2 corrections by counting the total number of error corrections or comparing the frequencies or proportions of different types of corrections. My research gauges error correction rates, which can yield better insights into how much attention learners pay to each type of error during chat sessions. To date, only one L2 online chat study has examined error correction rates (Yuan, 2003), and that study explored error corrections only from the chat scripts of just two participants and did not classify the types of errors into comprehensive categories. The present research addresses these methodological shortcomings by exploring lexical and grammatical error correction rates through both chat logs and the drafting-process data of 84 participants.

8.3 Theoretical contributions

The main theoretical contribution of the current study is that, in this research, I took the initiative to investigate linguistic changes made during L2 chat within the framework of L2

writing revision. As mentioned in Chapter 5, Section 5.2, I speculated that the psycholinguistic processes involved in L2 production during chat are similar to those during writing. While previous studies have only compared L2 adjustments made during online chat with those made in speech, I also compared my findings to those of individual computer-assisted L2 writing research. This study found that, in line with the findings of individual computer-assisted L2 writing studies, L2 revisions in online chat tended to occur to lower-level linguistic units and there was a high occurrence of typographical revisions. Comparing the findings of this study to those of individual computer-assisted writing research, I also hypothesized that L2 revision during online chat and during individual computer-assisted writing might have similar relationships with typing ability. However, the dominance of content-related over form-related revisions found in my research contrasts with the trend observed in individual computer-assisted L2 writing studies, possibly due to the highly interactive nature of chat, which may promote more attention to meaning.

Moreover, by uncovering how learners revise during chat, this research contributes to a better understanding of learners' attention to form in this medium. While the central claims associated with the benefit of online chat for L2 acquisition concern its ability to promote learners' attention to form, this study found evidence that learners' attention may be mainly drawn to meaning during L2 online chat. Although there was a high frequency of revisions indicating learners' close attention to their L2 output, it is likely that this attention was mostly paid to the meaning-related aspect of language rather than to form.

Another theoretical contribution of my research concerns the relationship between L2 revision and text quality. This study is, to the best of my knowledge, the first empirical study to test Stevenson et al.'s (2006) hypothesis about a possible link between the outcome of L2 revisions and text quality. Stevenson et al. (2006) speculated that the outcome of revisions (i.e. whether revisions could improve the text) is an important indicator of text quality. In this research, a significant correlation was found between greater error revision success and increased final text accuracy, supporting Stevenson et al.'s hypothesis. However, among all the revision measures, error correction rates were found to be the

strongest predictors of final text accuracy. Because previous research exploring the connection between L2 revision and accuracy is scant and did not employ these correction measures, the findings of this study contribute new knowledge regarding this relationship.

Importantly, my research makes several contributions to the understanding of L2 writing processes in the computer-assisted context, as discussed in the summary of the previous chapter. The findings of this study indicate that learners with A2 to B2 proficiency levels can attend to typographical/ spelling mistakes and content at the same time, while writing on a computer, and that computer-assisted writing might restrict learners' attention to low-level linguistic units. Moreover, my research contributes to a better understanding of the link between typing ability and revision in computer-assisted L2 writing by exploring their relationship in the context of text chat. To the best of my knowledge, the relationship between these two variables has only been investigated in one computer-assisted L2 writing study (Barkaoui, 2016), and none of the L2 online chat research has examined participants' typing ability. In general, revision behaviours during computer-assisted L2 writing are probably not influenced by typing ability alone but could be affected by complex interactions between a host of factors. While most revision measures were not significantly correlated with typing ability, the significant findings concerning error correction rates and accuracy suggest that learners with better L2 typing skills potentially have more attention available for internal language monitoring and attending to language errors.

As far as post-task anticipation is concerned, my research provides information about the effects of post-task anticipation on the main task performance, which has rarely been examined. While all previous studies in this domain investigated the effects of post-task anticipation on L2 complexity, accuracy and fluency during oral production tasks, this research extends the scope of investigation to online writing (i.e. text chat), focusing on its effects on revision, as well as on accuracy and fluency. In addition, my research employed a new type of post-task – language correction post-task – which has not been utilised in other post-task anticipation studies. Overall, the findings reveal that in the context of my study, post-task anticipation did not significantly influence L2 accuracy, fluency or revision during online chat.

8.4 Pedagogical implications

The pedagogical implications of this study concern how educators can utilize text chat tasks and post-task anticipation for L2 learning. There are three major pedagogical implications pertaining to the use of L2 text chat tasks. First, as discussed in Chapter 7, the high proportion of content revisions compared to form-related revisions found in this research shows that learners' attention may be primarily drawn to meaning while performing information-gap tasks via online chat. This suggests the importance of promoting learners' attention to form in this medium, for instance through consciousness raising. To benefit from the text-based nature of chat, learners might have to be reminded prior to tasks to pay special attention to the L2 output and linguistic errors they see on screen while chatting. Although the main focus in information-gap tasks is on communication, in order to enhance L2 development, students can be asked to pay more attention to form when these tasks are utilized via chat. Because many participants reported that they thought chat was an informal type of communication in which reduced language accuracy was acceptable, it could also be beneficial to remind learners that they should try to use L2 as accurately as possible when communicating in this medium. In addition, attention to form may be fostered by encouraging learners to use online resources, such as dictionaries, to find out how to correct their L2 errors or avoid making mistakes while chatting. When learners' attention to form is promoted, it is likely that the frequency of form-related revisions will increase. Additionally, in the previous chapter, I speculated that non-significant effects of post-task anticipation on accuracy and revision were observed due to learners' attitudes, which view chat language as informal. It is possible that, by raising learners' consciousness towards L2 accuracy in this medium, the effectiveness of post-task anticipation in promoting accuracy and revision will increase.

Second, the findings of this study indicate the importance of teaching learners how to revise during L2 online chat task performance. The findings which concern the linguistic unit of revisions suggest a possible limitation of text chat in fostering learners' attention to their overall output. In line with the findings of individual computer-assisted L2 writing research, it was found that the focus of students' attention can be restricted to the letter,

word or phrase being typed, instead of the whole clause or sentence, during chatting. Most grammatical, lexical and even content revisions were carried out to low-level linguistic units. Hence, to ensure learners pay attention to the whole text being produced and make changes to larger-level linguistic units, it may be necessary to teach learners online chat revision strategies that aim to improve the text as a whole. This proposal corresponds with Cumming and So's (1996) assumption, which hypothesizes that students' revision behaviours are influenced by instruction, and their call for explicit instruction in revision. Additionally, it is consistent with Hayes' (1996) hypothesis about L1 cognitive writing processes, which assumes the impact of task schema, a set of knowledge acquired through practice, on revision. According to this hypothesis, the processes of revision are governed by, for instance, knowledge about the actions needing to be taken when revising and the goal of revision. When learners receive training on revision strategies, they will be equipped with a new set of knowledge related to revision, which can promote more extensive changes and improved text quality.

Third, another pedagogical implication of this research concerns the selection of L2 online chat tasks. As discussed in Chapter 7, Section 7.1, the disparity between the frequency of changes detected in this study and Smith's (2008) L2 online chat study could be due to the different types of tasks utilized in the research. The nature of tasks, such as their goals and characteristics, may determine learners' L2 revision behaviours in text chat. Therefore, when the goal is to foster learners' L2 revisions and attention to their own output during chat, educators must be careful about their choice of tasks. It could be that tasks that necessitate accuracy will promote accuracy and revision. For instance, information-gap picture description tasks that require accurate and detailed descriptions may encourage increased accuracy and revision, compared to tasks that simply require students to collaboratively sequence given scenes.

Apart from the pedagogical implications pertaining to the use of L2 text chat tasks outlined thus far, this study also has pedagogical implications related to post-task anticipation. This research found that learners' anticipation of an individual and a collaborative language correction post-task did not have a significant influence on accuracy or revision during the

main tasks. As discussed in Chapter 7, language correction post-tasks might not be suitable for promoting accuracy and revision as the opportunity for L2 corrections during post-tasks could discourage learners from making L2 corrections during the main tasks, and the post-tasks may not be threatening enough to boost learners' attention to their demands. In addition, the limited time given for completing the main tasks and learners' familiarity with each other might reduce the effectiveness of post-task anticipation in promoting learners' attention to language form. In this study, learners were given limited time to perform the main tasks and were familiar with their partner. It could be that participants did not have enough time for revision, or were too preoccupied with finishing the main tasks in time to be concerned with linguistic form. Their familiarity with their partner might also reduce the threat of the collaborative language correction post-task and, hence, its effectiveness. Therefore, this indicates the importance of careful selection of post-tasks and careful consideration of the time constraints for the main task and the social context when the goal is to raise learners' attention to form through post-task anticipation. Regarding post-task selection, it is possible that language correction post-tasks in which errors are seen and discussed by more people than in the post-tasks of this study, e.g. by the whole class, are more face-threatening. Thus, the utilization of these post-tasks may potentially encourage learners to be more accurate and revise more during the main task.

8.5 Limitations and directions for future research

Although there are several contributions made and implications provided by this research, it is not without its limitations. The first limitation of the present study lies in the data collection methods. First, the number of participants could have been higher to ensure that individual differences did not obscure the findings which might have been revealed with a larger sample size. Future research could replicate my study by recruiting more participants to confirm the findings.

The second shortcoming of this research, regarding data collection, is that the revisions examined in this study were only those changes occurring during drafting and made to already-sent text. Data from screen recording, keystroke logging and chat logs did not provide information about language repairs that took place internally. Investigation of these

repairs would contribute to a more comprehensive understanding of L2 monitoring during online chat. Although stimulated recall interviews allow researchers to gain insights into internal language repairs, they were only carried out with some participants to ensure accurate classification of types of error corrections and revisions. Recall interviews were conducted to tap into what participants were thinking when they made changes to their text, they did not aim to collect information about internal L2 repairs. Future works exploring L2 revisions during online chat should address this problem by conducting stimulated recall interviews to investigate the thoughts of all participants during all task performances. To gauge their internal repairs, researchers can ask participants to verbalize what they are thinking at the time they make changes to their text and pause during chat. Stimulated recall which probes learners' thoughts during repairs and pauses has been commonly used to gauge internal language production processes in previous L2 speech and writing research (e.g. Ahmadian & Tavakoli, 2014; Ahmadian, Abdolrezapour, & Ketabi, 2012; Kormos, 2003; Révész, Kourtali, & Mazgutova, 2017).

Third, while the present study used several methods to collect the data, these did not include eye-tracking technology, which can reveal chat users' eye gaze during writing and revision processes. Psychological and L2 studies employing eye tracking have assumed a close connection between humans' eye movements and their minds. They link the direction of gaze to mental focus (Leow, Grey, Marijuan, & Moorman, 2014; Reichle, Pollatsek, & Rayner, 2006) and use eye gaze information, such as the number and duration of eye fixations, to determine participants' attention paid to particular objects. If this research had incorporated eye-tracking technology, it might have yielded additional insights into learners' attention to form during L2 revision processes. For instance, L2 error corrections, evidence of learners' attention to form during online chat, were investigated through drafting-process data and chat logs in this study. Through these data, corrections made during drafting and to already-sent text could be examined. However, these data did not reveal whether learners attended to uncorrected errors or not. Indeed, these errors may be noticed by learners but they might choose not to correct them due to, for instance, their attitudes, which view text chat as an informal mode of communication, and their personal orientation towards correcting particular types of errors over others. If information about

participants' eye gaze had been obtained, it would have been possible to discern if uncorrected errors were attended to, and the findings could have led to a clearer understanding of learners' error corrections and attention to form. Future L2 online chat revision and writing research should use eye gaze data to triangulate the data obtained from other methods. To this end, researchers could follow the guidelines proposed by O'Rourke (2008), which outline several stages of online chat in which eye-tracking technology could be used to examine the writer's gaze. This includes exploring gaze while drafting, after drafting but before sending, and after sending chat messages. Following these guidelines, researchers can discern whether learners actually attend to their mistakes or look back to previous messages to compare text during drafting, and whether they pay attention to their messages before and after sending them.

The fourth limitation of my study concerning data collection is that, in this project, data were gathered from participants who performed tasks in an experimental setting. As this might limit the ecological validity of this research, future studies should explore L2 online chat revision in a more naturalistic setting to check if the findings of my study are consistent across settings.

Apart from the data collection methods, another limitation concerns how pauses were determined. This research used mean characters between pauses, which indicate the length of translating episodes or pause bursts, as a measure of speed fluency. The pause threshold was set at 2,000 milliseconds in conformity with common practice in computer-assisted writing studies (e.g. Révész, Kourtali, & Mazgutova, 2017; Spelman Miller et al., 2008; Sullivan & Lindgren, 2002). Although this length of time has been most frequently used as a pause threshold in writing research (Van Waes & Leijten, 2015), some researchers have set a different minimum length of pause, such as 3,000 milliseconds (Abdel Latif, 2009), or 200, 500, 1,000 or 5,000 milliseconds (Van Waes & Leijten, 2015). Having a predetermined cut-off duration of pauses has also been criticized as being insensitive to variations in writers' typing speed (Wengelin, 2006). According to Wengelin (2006), a pause is "a transition time between two keystrokes that is longer than what can be expected to be necessary for the time needed to merely find the next key" (p. 126). As the normal

transition time between keystrokes probably varies depending on the typing speed of the writer, Wengelin proposes that each writer requires their own pause criterion based on their typing speed, which could be three times the mean transitioning time between two keystrokes. In future studies exploring pauses in computer-assisted L2 writing or L2 online chat, the pause threshold could be set by taking writers' variation into account.

Thus far, I have discussed the limitations of this research pertaining to data collection methods and the measure of speed fluency. Another limitation of the current study lies in the generalizability of its findings. The first research question of this study enquired into the L2 revision behaviours of Thai learners of English with upper-elementary to upper-intermediate proficiency levels during text chat. Because the study was conducted with L2 learners of specific proficiency levels, in a particular research setting and by utilizing a certain type of task, the findings may not be generalizable to L2 online chat revision behaviours in other contexts.

One direction for future studies on L2 online chat revision could be to investigate the effect of different variables on revision behaviours. As discussed in the literature review (Chapter 2), to date, many suggestions have been made regarding the possible influence of several factors on writing, including the individual, task and social environment factors. However, there has been a lack of research exploring how different variables affect L2 revision during online chat. To the best of my knowledge, only one study (Smith, 2009) has examined this by exploring the relationship between scrolling and L2 repairs during chat. Similarly, in the domain of L2 writing research, the influence of numerous factors on L2 revision behaviours has been insufficiently explored. Therefore, investigation of the effect of different variables on L2 revision would allow researchers to unveil the complexity of L2 revision and broaden our understanding of how L2 learners write and revise. In addition, the findings of research in this domain might also inform L2 educators how to draw learners' attention to their linguistic output during writing and promote revision.

It would be interesting for future L2 online chat studies to examine the influence of task characteristics, instructional methods, individual differences, the social environment and

the use of writing tools on revision. First, for task characteristics, researchers could compare the effects of different types, topics and time constraints of task. As far as task type is concerned, research could be conducted to compare how learners revise in different types of tasks, such as narrative, decision-making and argumentative tasks. In addition, future work could compare the effect of focused and unfocused tasks − those with and without specified linguistic targets − on L2 online chat revision. Probably, focused tasks that are designed to elicit specific L2 features would promote revision of those features. Regarding the topics and time constraints of tasks, investigation of how learners' familiarity with task topics and the amount of time given for task completion influence L2 online chat revision behaviours could be informative. It is possible that learners revise their output differently depending on their knowledge of the topic, and that they revise and correct their L2 errors more frequently when given more time to complete tasks.

With regard to instructional methods, future L2 online chat research can be carried out to examine how consciousness-raising activities, such as a pre-task language review or explicit instruction about particular language features, affect L2 revision during online chat tasks. Learners' consciousness raised towards specific L2 features potentially promotes L2 revision related to them. Moreover, it may be informative to explore whether task instructions that explicitly direct learners to pay attention to their L2 mistakes and output during online chat tasks encourage L2 revisions and error corrections, and whether revision training has an effect on how learners make changes to their text.

In terms of individual differences, the influence of both cognitive (e.g. working memory and attentional control) and affective (e.g. task motivation and writers' beliefs about revision) individual differences on L2 online chat revision could be investigated. In relation to this, it would be particularly interesting to explore the effect of learners' attentional control. Attentional control is related to the efficiency of an attentional system. It is determined by one's capacity to maintain one's attention (alertness), direct one's attention to the task at hand (orientation) and notice input stimulus (detection) (Simard, Bergeron, Liu, Nader, & Redmond, 2016). It has been assumed that L2 speech production is an activity requiring the control of attention as speakers have to manage their limited cognitive

resources by allocating them to multiple parallel language production processes, and that one's ability to control one's attention influences L2 speech repair behaviours (Simard et al., 2016; Simard, Fortier, & Zuniga, 2011; Zuniga & Simard, 2019). Previous research findings have revealed that increased attentional control has a significant correlation with decreased frequency of L2 speech repairs (Simard et al., 2016; Zuniga & Simard, 2019), indicating that good attentional control may facilitate pre-articulatory planning (Zuniga & Simard, 2019). It is probable that, as in speech production, L2 learners can benefit from having good attentional control when communicating via online chat. The speed of communication in online chat is relatively high, as in speech. It is possible that L2 learners who can efficiently allocate their attention to different language processes within this time constraint of online chat are able to engage in better L2 planning and, in turn, revise their text less frequently in this medium.

Regarding the social environment, future work on L2 online chat could investigate the effect of different types of pairings on L2 revision behaviours, e.g. by gauging whether learners revise differently when chatting with a native versus a non-native peer, with a peer versus a teacher, or with a peer of the same, higher or lower L2 proficiency level.

Lastly, exploration of the influence of writing tools, such as automatic language checkers and online language resources, on L2 revision could be useful. Access to online dictionaries and other online language resources may facilitate L2 error detection and successful correction of language problems. An automatic grammar checking feature potentially promotes increased noticing and correction of grammatical errors. Indeed, this suggestion corresponds to Barkaoui's (2016) call for investigation of how editing and writing tools affect L2 revision behaviours. The author assumes that spelling and grammar checkers can reduce the cognitive demands required for revising these L2 features, thereby allowing learners to devote their attention to other language aspects.

Essentially, this research was carried out to investigate L2 revision during computer-based online chat. Since mobile-assisted text chat has become common following the evolution and increased popularity of mobile technologies, future studies could examine L2 revision

during mobile-based text chat to discern whether the nature of L2 revision in these two modes is different.

In addition to the directions for future studies on L2 online chat revision outlined above, there is still room for more research on the other topics explored in this study. As part of RQ2, this study investigated the effects of post-task anticipation and found that its effect on accuracy was not consistent with that found in some previous research. Due to this inconsistency and a severe lack of post-task anticipation research in SLA, with only two studies carried out prior to this project, more studies in this area are needed to gain insights into possible benefits of post-task anticipation. First, as the findings of Skehan and Foster (1997) and Foster and Skehan (2013) suggest that post-task anticipation may have a greater influence on accuracy in decision-making tasks than narrative and information-exchange tasks, future work in this domain could explore post-task anticipation by utilizing different types of main tasks. Additionally, researchers could compare the effectiveness of post-task anticipation when different amounts of time are given to perform the main task. As discussed earlier, I hypothesized that the limited time for completing the main task potentially reduces the effectiveness of post-task anticipation in terms of promoting learners' attention to language form. When the time for the main task is limited, learners may not be much concerned with language form as they are too preoccupied with finishing the task in time. Third, future SLA research exploring post-task anticipation could examine whether the anticipation of different types of post-tasks has a different influence on main task performance. This study utilized an individual and a collaborative language correction post-task and found that they might not be suitable for promoting accuracy and revision during the main task. The collaborative post-task was performed in pairs by learners who knew each other prior to the experiment. Compared to these post-tasks, transcription post-tasks (cf. Foster & Skehan, 2013) or more face-threatening language correction tasks in which errors are seen and corrected by more people, such as the whole class, or by someone learners are not familiar with might be better alternatives for enhancing L2 revision and accuracy. Future research could utilize more than one type of post-task and compare the effectiveness of these post-tasks in raising attention to form.

Importantly, because there has been little research exploring how tasks could be implemented to promote heightened attention to form in L2 chat, more studies like this one would be beneficial. Researchers could investigate how anticipation of other post-tasks draws learners' attention to form while performing the main task via online chat, or examine other task implementation methods that have the potential to foster attention to form during L2 chat interaction.

As discussed in Chapter 5, Section 5.2, the justification for investigating the relationships between L2 revision and 1) accuracy and 2) typing ability in RQ3 and RQ4 lies in the lack of research in these areas. Prior to this study, very few studies were conducted to examine these relationships in L2 writing and none in L2 online chat. Thus, more studies that seek to examine these relationships are needed. Indeed, knowledge about how learners' revision is related to accuracy can contribute to our understanding of how L2 accuracy can be promoted. By exploring the connection between L2 typing ability and revision, researchers will enhance their comprehension of L2 writing processes.

In conclusion, despite the limitations discussed in this section, this study makes several methodological and theoretical contributions to the field of SLA and has important pedagogical implications pertaining to how educators can use text chat tasks and post-task anticipation for L2 learning. My research addresses the methodological shortcomings of previous studies that explored L2 changes during online chat by employing error correction rates and utilizing novel data collection methods (keystroke logging and verbal protocols) in this field. Thanks to keystroke logging, the investigation of accuracy and fluency through process-based measures was possible. Unlike in previous L2 online chat studies, L2 changes were investigated under the writing revision framework based on the assumption that the psycholinguistic processes involved in L2 production during online chat are similar to those involved during writing. In summary, the findings reveal that, contrary to claims regarding the benefit of text chat in promoting attention to form, learners' attention could be primarily drawn to meaning in this medium. However, when learners of A2 to B2 proficiency levels actually correct erroneous linguistic form, the outcome is likely to be successful. In line with previous research exploring L2 repairs during online chat, this study

found that learners may devote more attention to grammatical items than lexical items. In addition, error correction rates are probably the strongest predictors of final text accuracy among all revision measures. Contributing to the knowledge of L2 writing processes in the computer-assisted context, my study reveals that typographical/ spelling revisions may not substantially interfere with content processing, and that learners' attention is possibly restricted to low-level linguistic units in this context. During computer-assisted writing, an automatized L2 typing process might free up the cognitive resources required for internal L2 monitoring and attending to L2 errors. However, L2 revision may not be influenced by typing ability alone but also by a host of other factors. In terms of the effect of post-task anticipation on attention to form, this study extends the previous scope of investigation from oral production tasks to online chat tasks, and finds that anticipation might not significantly influence the main task performance in this medium. It can be hypothesized that the effectiveness of post-task anticipation is influenced by many factors, including the time constraints of the main task. Overall, the findings of this research enhance our understanding of L2 revision and attention to form during online chat, the effects of post-task anticipation, the relationship between L2 revision and text quality, the influence of L2 typing ability on writing and L2 writing processes in a computer-assisted context.

References

Abdel Latif, M. M. (2009). Toward a new process-based indicator for measuring writing fluency: Evidence from L2 writers' think-aloud protocols. *The Canadian Modern Language Review, 65*(4), 531–558.

Abdel Latif, M. M. (2013). What do we mean by writing fluency and how can it be validly measured? *Applied Linguistics, 34*(1), 99–105.

Adams, R. (2006). L2 tasks and orientation to form: A role for modality? *ITL: International Journal of Applied Linguistics, 152*, 7–34.

Adams, R., & Nik, N. (2014). Prior knowledge and second language task production in text chat. In M. González-Lloret & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 51–78). John Benjamins Publishing.

Adams, R., Nik, N., & Newton, J. (2015). Task complexity effects on the complexity and accuracy of writing via text chat. *Journal of Second Language Writing, 29*, 64–81.

Adams, R., & Ross-Feldman, L. (2008). Does writing influence learner attention to form? In D. Belcher & A. Hirvela (Eds.), *The oral-literate connection: Perspectives on L2 speaking, writing, and other media interactions* (pp. 243–266). Ann Arbor: University of Michigan Press.

Ahmadian, M. J., Abdolrezapour, P., & Ketabi, S. (2012). Task difficulty and self-repair behavior in second language oral production. *International Journal of Applied Linguistics, 22*(3), 310–330.

Ahmadian, M. J., & Tavakoli, M. (2014). Investigating what second language learners do and monitor under careful online planning conditions. *The Canadian Modern Language Review, 70*(1), 50–75.

Allan, D. (2004). *Oxford placement test 2*. Oxford: Oxford University Press.

Allford, D., & Pachler, N. (2007). *Language, autonomy and the new learning environments* (1st ed.). Frankfurt am Main: Peter Lang.

Arslanyilmaz, A. (2012). An online task-based language learning environment: Is it better for advanced- or intermediate-level second language learners? *The Turkish Online Journal of Educational Technology*, *11*(1), 20–35.

Atlas.ti 8 [Computer software]. (2017). Retrieved from: http://atlasti.com

Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, *36*(3), 199–223.

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

Baralt, M. (2010). *Task complexity, the cognition hypothesis, and interaction in CMC and FTF environments* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Baralt, M., Gurzynski-Weiss, L., & Kim, Y. (2016). Engagement with the language: How examining learners' affective and social engagement explains successful learner-generated attention to form. In M. Sato and S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 209–239). Amsterdam: John Benjamins.

Barkaoui, K. (2007). Revision in second language writing: What teachers need to know. *TESL Canada Journal*, *25*(1), 81–92.

Barkaoui, K. (2014). Examining the impact of L2 proficiency and keyboarding skills on scores on TOEFL-iBT writing tasks. *Language Testing*, *31*(2), 241–259.

Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal*, *100*(1), 320–340.

Barkaoui, K., & Knouzi, I. (2018). The effects of writing mode and computer ability on L2 test-takers' essay characteristics and scores. *Assessing Writing*, *36*, 19–31.

Bartlett, E. J. (1982). Learning to revise: Some component processes. In M. Nystrand (Ed.), *What writers know: The language, process, and structure of written discourse* (pp. 345–363). New York: Academic Press.

Beauvois, M. H. (1992). Computer-assisted classroom discussion in the foreign language classroom: Conversation in slow motion. *Foreign Language Annals*, *25*(5), 455–464.

Beauvois, M. H. (1997). Computer-mediated communication (CMC): Technology for improving speaking and writing. In M. D. Bush (Ed.), *Technology-enhanced language learning* (pp. 165–184). Lincolnwood, Illinois: National Textbook.

Benson, P. (2000). *Autonomy in language learning*. Harlow: Longman.

Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language*, *27*, 261–278.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Blake, R. (2007). New trends in using technology in the language curriculum. *Annual Review of Applied Linguistics*, *27*, 76–97.

Broekkamp, H., & van den Bergh, H. (1996). Attention strategies in revising a foreign language text. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), *Theories, models and methodology in writing research* (pp. 170–181). Amsterdam: Amsterdam University Press.

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

Chapelle, C. (1998). Analysis of interaction sequences in computer-assisted language learning. *TESOL Quarterly*, *32,* 753–757.

Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge: Cambridge University Press.

Chapelle, C. (2004). Technology and second language learning: Expanding methods and agendas. *System*, *32*, 593–601.

Chen, W. C. (2008). *Noticing in text-based computer-mediated communication: A study of a task-based telecommunication between native and nonnative English speakers* (Unpublished doctoral dissertation). Texas A&M University.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*(1), 80–98.

Chenoweth, N. A., & Hayes, J. R. (2003). Inner voice in writing. *Written Communication, 20*, 99–118.

Choi, Y. H. (2007). On-line revision behaviors in EFL writing process. *English Teaching*, *62*(4), 69–93.

Chun, D. M. (1994). Using computer networking to facilitate the acquisition of interactive competence. *System*, *22*(1), 17–31.

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

Conacher, J. E., & Kelly-Holmes, H. (2007). *New learning environments for language learning: Moving beyond the classroom?* Frankfurt am Main: Peter Lang.

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, *7*(4), 482–511.

Cumming, A., & So, S. (1996). Tutoring second language text revision: Does the approach to instruction or the language of communication make a difference? *Journal of Second Language Writing*, *5*(3), 197–226.

Darhower, M. (2002). Interactional features of synchronous computer-mediated communication in the intermediate L2 class: A sociocultural case study. *CALICO Journal*, *19*, 249–277.

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

Doughty, C., & Long, M. H. (2003). Optimal psycholinguistic environments for distance foreign language learning. *Language Learning & Technology*, *7*(3), 50–80.

Ellis, R. (2006). The methodology of task-based teaching. *Asian EFL Journal*, *8*(3), 19–45.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, *19*(3), 221–246.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, *35*(3), 407–432.

Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, *32*(4), 400–414.

Fathman, A. K. (1980). Repetition and correction as an indication of speech planning and execution processes among second language learners. In H. W. Dechert & M. Raupach (Eds.), *Towards a crosslinguistic assessment of speech production* (pp. 77–85). Frankfurt: Peter D. Lang.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage.

Fitzgerald, J. (1987). Research on revision in writing. *Review of Educational Research*, *57*, 481–506.

Flashback Express 5 [Computer software]. (2017). Retrieved from: https://www.flashbackrecorder.com/express

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, *18*(3), 299−324.

Foster, P., & Skehan, P. (2013). Anticipating a post-task activity: The effects on accuracy, complexity, and fluency of second language performance. *The Canadian Modern Language Review*, *69*(3), 249–273.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, *21*(3), 354–375.

Fredriksson, C. (2015). The influence of group formation on learner participation, language complexity, and corrective behaviour in synchronous written chat as part of academic German studies. *ReCALL*, *27*(2), 217–238.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Knowing what to write: Conceptual processes in text production* (pp. 139–160). Amsterdam: Amsterdam University Press.

Galbraith, D. (2009a). Cognitive models of writing. *German as a Foreign Language Journal*, (2–3), 7–22.

Galbraith, D. (2009b). Writing as discovery. *British Journal of Educational Psychology Monograph Series*, *2*(6), 5–26.

Galbraith, D., & Baaijen, V. M. (2018). The work of writing: Raiding the inarticulate. *Educational Psychologist*, *53*(4), 238–257.

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

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

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.

Genesee, F. (1987). *Learning through two languages*. Boston: Heinle & Heinle.

Golonka, E. M., Tare, M., & Bonilla, C. (2017). Peer interaction in text chat: Qualitative analysis of chat transcripts. *Language Learning & Technology*, *21*(2), 157–178.

Gong, W., & Ooi, V. (2008). Innovations and motivations in online chat. In S. Kelsey & K. St. Amant (Eds.), *Handbook of research on computer mediated communication* (Vol. 2, pp. 917–933).

Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in FTF and CMC modes. *Studies in Second Language Acquisition*, *36*, 1–37.

Hall, C. (1990). Managing the complexity of revising across languages. *TESOL Quarterly*, *24*(1), 43–60.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*(3), 369–388.

Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J. F., & Carey, L. (1987). Cognitive processes in revision. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics, Vol. 2: Reading, writing, and language learning* (pp. 176–240). Cambridge, UK: Cambridge University Press.

Herring, S. (1996). *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives*. Amsterdam: John Benjamins Publishing.

Herring, S. (2001). Computer-mediated discourse. In D. Schiffrin, D. Tannen, & H. Hamilton (Eds.), *The handbook of discourse analysis* (pp. 612–634). Oxford: Blackwell.

Herring, S. (2012). Grammar and electronic communication. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Hoboken: Wiley-Blackwell.

Inoue, C. (2016). A comparative study of the variables used to measure syntactic complexity and accuracy in task-based research. *The Language Learning Journal*, *44*(4), 487–505.

Iwasaki, J., & Oliver, R. (2003). Chat-line interaction and negative feedback. *Australian Review of Applied Linguistics*, *17*, 60–73.

Jepson, K. (2005). Conversations and negotiated interaction in text and voice chat rooms. *Language Learning & Technology*, *9*(3), 79–98.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kellerman, E. (1985). If at first you do succeed... In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 345–353). Rowley, MA: Newbury House.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Mahwah, NJ: Lawrence Erlbaum Associates.

Kelm, O. (1992). The use of synchronous computer networks in second language instruction: A preliminary report. *Foreign Language Annals*, *25*(5), 441–454.

Kelm, O. (1996). The application of computer networking in foreign language education: Focusing on principles of second language acquisition. In M. Warschauer (Ed.), *Telecollaboration in foreign language learning: Proceedings of the Hawaii symposium* (pp. 19–28). Honolulu: University of Hawaii Press.

Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quantity and characteristics of language production. *The Modern Language Journal*, *79*(4), 457–476.

Kern, R., & Warschauer, M. (2000). Theory and practice of network-based language teaching. In M. Warschauer & R. Kern (Eds.), *Network-based language teaching: Concepts and practice* (pp. 1–19). New York: Cambridge University Press.

Kitade, K. (2000). L2 learners' discourse and SLA theories in CMC: Collaborative interaction in internet chat. *Computer Assisted Language Learning*, *13*(2), 143–166.

Kormos, J. (1998a). A new psycholinguistic taxonomy of self-repairs in L2: A qualitative analysis with retrospection. *Even Yearbook, ELTE SEAS Working Papers in Linguistics*, *3*, 43–68.

Kormos, J. (1998b). The use of verbal reports in L2 research: Verbal reports in L2 speech production research. *TESOL Quarterly*, *32*(2), 353–358.

Kormos, J. (1999a). Monitoring and self-repair in L2. *Language Learning*, *49*(2), 303–342.

Kormos, J. (1999b). The effect of speaker variables on the self-correction behaviour of L2 learners. *System*, *27*, 207–221.

Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning*, *50*(2), 343–384.

Kormos, J. (2002). *The structure of L2 self-repairs in the speech of Hungarian learners of English*. Unpublished manuscript.

Kormos, J. (2003). Attention and monitoring in a second language: A qualitative analysis. *Fremdsprachen Lehren und Lernen*, *32*, 116–132.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*, 390–403.

Krashen, S. D. (1978). Individual variation in the use of the monitor. In W. C. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 175–183). New York: Academic Press.

Lai, C., & Li, G. (2011). Technology and task-based language teaching: A critical Review. *CALICO Journal*, *28*(2), 498–521.

Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning & Technology*, *10*(3), 102–120.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lantolf, J. P., & Thorne, S. L. (2007). Sociocultural theory and second language learning. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 201–224). Mahwah, NJ: Lawrence Erlbaum Associates.

Larsen-Freeman, D. (1991). Second language acquisition research: Staking out the territory. *TESOL Quarterly*, *25*(2), 315–350.

Lee, H., Joo, K., Moon, J., & Hong, Y. (2007). EFL learner performance variation as the effect of interlocutor type. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use, learning and teaching* (pp. 127–138). Brussels: KVAB.

Lee, J. (2000). *Tasks and communicating in language classrooms*. Boston: McGraw-Hill.

Lee, L. (2001). Online interaction: Negotiation of meaning and strategies used among learners of Spanish. *ReCALL*, *13*(2), 232–244.

Lee, L. (2002). Synchronous online exchanges: A study of modification devices on non-native discourse. *System, 30*, 275–288.

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, *30*(3), 358–392.

Leijten, M., Van Waes, L., Galbraith, D., & Torrance, M. (2011, February). *The influence of working memory on error correction strategies during sentence production*. Paper presented at the 4th International Conference on Writing Research: Writing Research Across Borders II, Washington, DC.

Lennon, P. (1984). Retelling a story in English as a second language. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 50–68). Tübingen, Germany: Günter Narr.

Leow, R. P., Grey, S., Marijuan, S., & Moorman, C. (2014). Concurrent data elicitation procedures, processes, and the early stages of L2 learning: A critical overview. *Second Language Research*, *30*(2), 111 –127.

Li, J. (2006). The mediation of technology in ESL writing and its implications for writing assessment. *Assessing Writing*, *11*, 5–21.

Li, J., & Cumming, A. (2001). Word processing and second language writing: A longitudinal case study. *International Journal of English Studies*, *1*(2), 127–152.

Liao, P. L., & Fu, K. (2014). Effects of task repetition on L2 oral (in written form) production in computer-mediated communication. *International Journal of Humanities and Arts Computing*, *8*, 221–236.

Lindgren, E., & Sullivan, K. P. H. (2006). Analysing online revision. In K. P. H. Sullivan & E. Lindgren (Eds.), *Studies in writing, Vol. 18, Computer keystroke logging and writing: Methods and applications* (pp. 157–188). Oxford: Elsevier.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, *4*, 126–141.

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 77–99). Clevedon: Multilingual Matters.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: John Benjamins.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). New York: Academic Press.

Long, M. H. (1998). Focus on form in task-based language teaching. *University of Hawaii Working Papers in ESL*, *16*(49–61).

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Hoboken, NJ: John Wiley & Sons.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, *26*(1), 27–56.

Long, M. H., & Norris, J. (2000). Task-based language teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597–603). London: Routledge.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–167). Clevedon: Multilingual Matters.

MacArthur, C. A., Graham, S., & Fitzgerald, J. (2008). *Handbook of writing research*. New York: The Guilford Press.

Manchón, R. M. (2011). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins.

Manchón, R. M., Roca de Larios, J., & Murphy, L. (2009). The temporal dimension and problem-solving nature of foreign language composing processes: Implications for theory. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 102–129). Clevedon: Multilingual Matters.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, *20*, 83−108.

Motteram, G., & Thomas, M. (2010). Afterword: Future directions for technology-mediated tasks. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 218–237). London/New York: Continuum.

Murray, D. E. (2000). Protean communication: The language of computer-mediated communication. *TESOL Quarterly*, *34*(3), 397–421.

New, E. (1999). Computer-aided writing in French as a foreign language: A qualitative and quantitative look at the process of revision. *The Modern Language Journal*, *83*(1), 80–97.

Nik, N. (2010). *Examining the language learning potential of a task-based approach to synchronous computer-mediated communication* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand.

Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness, 18,* 384–402.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.

O'Connor, N. (1988). Repairs as indicative of interlanguage variation and change. In T. J. Walsh (Ed.), *Georgetown University Round Table in Languages and Linguistics 1988: Synchronic and diachronic approaches to linguistic variation and change* (pp. 251–259). Washington, DC: Georgetown University Press.

O'Rourke, B. (2008). The other C in CMC: What alternative data sources can tell us about text-based synchronous computer mediated communication and language learning. *Computer Assisted Language Learning*, *21*(3), 227–251.

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: Second language research in honor of Susan M. Gass* (pp. 226–253). New York: Routledge.

Ortega, L., & González-Lloret, M. (2015). Staking out the territory of technology-mediated TBLT. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 59–86). John Benjamins Publishing.

Palviainen, Å., Kalaja, P., & Mäntylä, K. (2012). Development of L2 writing: Fluency and proficiency. *AFinLA-E: Soveltavan Kielitieteen Tutkimuksia, 4,* 47–59.

Payne, J. S., & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal*, *20*(1), 7–32.

Piolat, A. (1997). Writers' assessment and evaluation of their texts. In C. Clapham & D. Corson (Eds.), *Encyclopaedia of language and education* (Vol. 7, pp. 189–198). Dordrecht, Netherlands: Kluwer Academic.

Plonsky, L., & Oswald, F. L. (2014). How big is "Big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*(4), 878–912.

Porte, G. (1996). When writing fails: How academic context and past learning experiences shape revision. *System*, *24*(1), 107–116.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, *9*(5), 1–6.

Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, *37*(3), 439–467.

Razagifard, P. (2013). The impact of text-based CMC on improving L2 oral fluency. *Journal of Computer Assisted Learning*, *29*, 270–279.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2006). E–Z reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7*, 4–22.

Révész, A., Kourtali, N.-E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208–241.

Révész, A., Michel, M., & Lee, M. (2017). *Investigating IELTS academic writing task 2: Relationships between cognitive writing processes, text quality, and working*

memory. Retrieved from: https://www.ielts.org/-/media/research-reports/ielts_online_rr_2017-3.ashx

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*(2), 283–331.

Robinson, P. (2001a). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287–318). Cambridge: Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27–57.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 631–678). Malden, MA: Blackwell.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, *43*(1), 1–32.

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Putz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind* (pp. 243–269). Philadelphia, PA: John Benjamins.

Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance*. Philadelphia/Amsterdam: John Benjamins.

Robinson, P. (2015). The Cognition Hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and*

directions in the development of TBLT: A decade of plenaries from the international conference (pp. 87–121). Amsterdam: John Benjamins.

Roca de Larios, J., Manchón, R., Murphy, L., & Marín, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing, 17*, 30–47.

Rogers, T. T., & McClelland, J. L. (2014). Parallel distributed processing at 25: Further explorations in the microstructure of cognition. *Cognitive Science, 38*(6), 1024–1077.

Rouhshad, A., Wigglesworth, G., & Storch, N. (2016). The nature of negotiations in face-to-face versus computer-mediated communication in pair interactions. *Language Teaching Research, 20*(4), 514–534.

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Parallel distributed processing models of schemata and sequential thought processes. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition, Vol. 2*. Cambridge, MA: Bradford Books.

Salaberry, M. R. (2000). L2 morphosyntactic development in text-based computer-mediated communication. *Computer Assisted Language Learning, 13*(1), 5–27.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. London: Palgrave Macmillan.

Satar, H., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal, 92*(4), 595–613.

Sauro, S. (2009). Computer-mediated corrective feedback and the development of L2 grammar. *Language Learning & Technology, 13*(1), 96–120.

Sauro, S., & Smith, B. (2010). Investigating L2 performance in text chat. *Applied Linguistics*, *31*(4), 554–577.

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics: Reading, writing, and language learning* (Vol. 2, pp. 142–175). Cambridge: Cambridge University Press.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*(2), 129–158.

Schmidt, R. (1994a). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, *11*, 11–26.

Schmidt, R. (1994b). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–209). London: Academic Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

Seliger, H. (1980). Utterance planning and correction behavior: Its function in the grammar construction process for second language learners. In H. W. Dechert & M. Raupach (Eds.), *Towards a crosslinguistic assessment of speech production* (pp. 87–99). Frankfurt: Peter D. Lang.

Shekary, M., & Tahririan, M. H. (2006). Negotiation of meaning and noticing in text-based online chat. *The Modern Language Journal*, *90*(4), 557–573.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, *27*(4), 657–677.

Simard, D., Bergeron, A., Liu, Y.-G., Nader, M., & Redmond, L. (2016). Production d'autoreformulations autoamorcées en langue seconde: Rôle de l'attention et de la mémoire phonologique. *The Canadian Modern Language Review, 72*(2), 183–210.

Simard, D., Fortier, V., & Zuniga, M. (2011). Attention et production d'autoreformulations autoamorcées en français langue seconde, quelle relation? *Journal of French Language Studies, 21*(3), 417–436.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*, 38–62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2003a). Focus on form, tasks, and technology. *Computer Assisted Language Learning, 16*(5), 391–411.

Skehan, P. (2003b). Task-based instruction. *Language Teaching, 36*(1), 1–14.

Skehan, P. (2007). Task research and language teaching: Reciprocal relationships. In S. Fotos & H. Nassaji (Eds.), *Form-focused instruction and teacher education: Studies in Honor of Rod Ellis* (pp. 55–69). Oxford: Oxford University Press.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532.

Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–260). Amsterdam: John Benjamins.

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics, 36*, 34–49.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, *1*(3), 185–211.

Smith, B. (2003). Computer-mediated negotiated interaction: An expanded model. *The Modern Language Journal*, *87*(1), 38–57.

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly*, *39*(1), 33–58.

Smith, B. (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning & Technology*, *12*(1), 85–103.

Smith, B. (2009). The relationship between scrolling, negotiation, and self-initiated self-repair in an SCMC environment. *CALICO Journal*, *26*(2), 231–245.

Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In F. M. Hult (Ed.), *Directions and prospects for educational linguistics* (pp. 79–97). New York: Springer.

Smith, B., & Renaud, C. (2013). Using eye-tracking as a measure of foreign language learners' noticing of recasts during computer-mediated writing conferences. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 147–165). Amsterdam: John Benjamins.

Sotillo, S. (2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology*, *4*(1), 82–119.

Spelman Miller, K. (2000). Academic writers on-line: Investigating pausing in the production of text. *Language Teaching Research*, *4*(2), 123–148.

Spelman Miller, K., Lindgren, E., & Sullivan, K. P. H. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42*(3), 433–454.

Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, *15*, 201–233.

Sullivan, K., & Lindgren, E. (2002). Self-assessment in autonomous computer-aided second language writing. *ELT Journal*, *56*(3), 258–266.

Sullivan, N., & Pratt, E. (1996). A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom. *System*, *29*(4), 491–501.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*(3), 320–337.

Thomas, M., & Peterson, M. (2014). Editorial for the special issue: Web 2.0 and language learning: Rhetoric and reality. *CALICO Journal*, *31*, i–iii.

Thomas, M., & Reinders, H. (2010). Deconstructing tasks and technology. In M. Thomas & H. Reinders (Eds.), *Task-based language learning and teaching with technology* (pp. 1–13). London/New York: Continuum.

Thorne, S. L. (2013). Digital literacies. In M. Hawkins (Ed.), *Framing languages and literacies: Socially situated views and perspectives* (pp. 192–218). New York: Routledge.

Thorson, H. (2000). Using the computer to compare foreign and native language writing processes: A statistical and case study approach. *The Modern Language Journal*, *84*(2), 155–170.

Tillema, M. (2012). *Writing in first and second language: Empirical studies on text quality and writing processes* (Unpublished doctoral dissertation). Utrecht Institute of Linguistics, Netherlands.

Tudini, V. (2003). Using native speakers in chat. *Language Learning & Technology*, *7*(3), 141–159.

van Bruggen, J. A. (1946). Factors affecting regularity of the flow of words during written composition. *Journal of Experimental Education*, *15*(2), 133–155.

Van den Branden, K., Bygate, M., & Norris, J. M. (2009). *Task-based language teaching: A reader*. Amsterdam: John Benjamins.

van Hest, E. (1996). *Self-repair in L1 and L2 production*. Tilburg: Tilburg University Press.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, *12,* 287–301.

VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features and research methodology. *AILA Review, 11,* 27–36.

VanPatten, B. (1996). *Input processing and grammar instruction*. New York: Ablex.

VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Lawrence Erlbaum Associates.

Van Waes, L., & Leijten, M. (2015). Fluency in writing: A multidimensional perspective on writing fluency applied to L1 and L2. *Computers and Composition*, *38*, 79–95.

Van Waes, L., Leijten, M., & Van Weijen, D. (2009). Keystroke logging in writing research: Observing writing processes with Inputlog. *German as a Foreign Language*, *2*(3), 41–64.

Van Waes, L., Leijten, M., Wengelin, A., & Lindgren, E. (2012). Logging tools to study digital writing processes. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 507–533). New York/Sussex: Taylor & Francis.

Van Waes, L., & Schellens, P. J. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics*, *35*, 829–853.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*, *13*, 7–26.

Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The Modern Language Journal*, *81*(4), 470–481.

Warschauer, M. (1999). *Electronic literacies: Language, culture, and power in online education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wengelin, A. (2006). Examining pauses in writing: Theory, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Studies in writing, Vol. 18, Computer*

keystroke logging and writing: Methods and applications (pp. 107–130). Oxford: Elsevier.

Whalen, K., & Ménard, N. (1995). L1 and L2 writers' strategic and linguistic knowledge: A model of multiple-level discourse processing. *Language Learning*, *45*(3), 381–418.

Wickens, C. (1984). Processing resources in attention. In R. Parasuraman & D. Davies (Eds.), *Varieties of attention* (pp. 63–101). New York: Academic Press.

Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 671–691). Mahwah, NJ: Lawrence Erlbaum Associates.

Willis, J. (1996). *A framework for task-based learning*. Essex: Longman.

Yates, S. J. (1996). Oral and written linguistic aspects of computer conferencing: A corpus-based study. In S. Herring (Ed.), *Computer-mediated communication: Linguistic, social and cross-cultural perspectives* (pp. 29–46). Amsterdam: John Benjamins.

Yilmaz, Y., & Yuksel, D. (2011). Effects of communication mode and salience on recasts: A first exposure study. *Language Teaching Research*, *15*, 457–477.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*(1), 1–27.

Yuan, K.-H., & Maxwell, S. (2005). On the post hoc power in testing mean differences. *Journal of Educational and Behavioral Statistics*, *30*(2), 141–167.

Yuan, Y. (2003). The use of chat rooms in an ESL setting. *Computers and Composition*, *20*, 194–206.

Yuksel, D., & Inan, B. (2014). The effects of communication mode on negotiation of meaning and its noticing. *ReCALL*, *26*(3), 333–354.

Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly*, *17*(2), 165–187.

Zeng, G. (2017). Collaborative dialogue in synchronous computer-mediated communication and face-to-face communication. *ReCALL*, *29*(3), 257–275.

Ziegler, N. (2016). Taking technology to task: Technology-mediated TBLT, performance, and production. *Annual Review of Applied Linguistics*, *36*, 136–163.

Ziegler, N., & Mackey, A. (2014, September). *Pre-task planning, performance, and individual differences in L2 text chat*. Paper presented at EUROSLA conference, York, UK. Abstract retrieved from: https://www.york.ac.uk/education/research/crellu/eurosla24

Zimmerman, R. (2000). L2 writing: Sub-processes, a model of formulating and empirical findings. *Learning and Instruction*, *10*, 73–99.

Zuniga, M., & Simard, D. (2019). Factors influencing L2 self-repair behavior: The role of L2 proficiency, attentional control and L1 self-repair behavior. *Journal of Psycholinguistic Research*, *48*(1), 43–59.

Appendices

Appendix A: Participant information sheet

Participant information sheet (Translated from Thai)

Title: Language learning through online chat tasks

Researcher: Vararin Charoenchaikorn, PhD student
Department of Linguistics and English Language, Lancaster University

You are invited to take part in this research study. Please take time to read the following information carefully before you decide whether or not you wish to take part.

What is the study about?
I am carrying out this study as part of my Doctoral studies in the Department of Linguistics and English Language, Lancaster University. The aim of the study is to investigate second language learning through text chat. My study will involve a series of English pedagogic tasks performed in online chat.

Why have I been invited?
I have approached you because I am interested in understanding how students who learn English can benefit from performing tasks via text chat. I would be very grateful if you would agree to take part in my study.

What will I be asked to do if I take part?
If you decided to take part, you will be asked to complete 1) English language tasks, 2) a brief typing test, 3) a background and an exit questionnaire, and you will also be interviewed by me.

The experiment will take approximately 2 hours and a half. During the experiment, your computer screen and keystrokes (keys pressed on your keyboard) will be recorded by using a specialized software. I will use an audio recorder, and in some cases a video camera, to make recordings during the interview. In the video recordings, your face will not be shown. During some English language tasks, your voice may also be recorded.

Who will conduct the study?
The experiment will be conducted by me.

What are the possible benefits from taking part?
Taking part in the experiment will allow you to experience an online platform for language learning and reflect on your own experiences of learning English. Also, to compensate for your time, you will be given 300 baht. You will not lose this payment even if you decide to withdraw from the experiment.

Do I have to take part?
No. It's completely up to you to decide whether or not you take part.

What if I change my mind?
You are free to withdraw from the study at any time during the experiment and up to 2 weeks after taking part in this study, and you do not have to give a reason. If you withdraw within this period, I will not use any of the information that you provided for my project and your data will be destroyed. If you withdraw after 2 weeks, your data will be used for my project.

How will my data be stored?
All the information obtained during the course of the research will be kept securely. Any paper-based data will be kept in a locked cupboard which I am the only person with the key to. Electronic data will be stored on a password protected computer and an encrypted external hard drive. Files containing personal data will be encrypted. Data that can identify you such as age will be kept strictly confidential and separately from non-personal information. In accordance with University guidelines, I will keep the data securely for a minimum of 10 years. The data may be seen by my supervisor, Professor Judit Kormos, who will give me guidance regarding my research.

What will happen to the results of the research study?
The results of the study will be used for academic purposes only. This will include my PhD thesis and other publications, for example journal articles. I am also planning to present the results of my study in academic conferences and classes I teach.

Will my data be identifiable?
Any identifying information will be anonymised when writing up and presenting the findings from this study. I will only use anonymised quotes from the interview with you, so you cannot be identified in publications. Your video and voice recordings will be shown/ played during presentations and teaching only if you give consent to it.

What if I have a question or concern?
If you have any queries or if you are unhappy with anything that happens concerning your participation in the study, please contact myself or my supervisor, Professor Judit Kormos, Department of Linguistics and English Language, Lancaster University, LA1 4YL, UK, j.kormos@lancaster.ac.uk.

If you have any concerns or complaints that you wish to discuss with a person who is not directly involved in the research, you can also contact Professor Elena Semino, Department of Linguistics and English Language, Lancaster University, LA1 4YL, UK, +44(0)1524 594176, e.semino@lancaster.ac.uk.

Vararin Charoenchaikorn, PhD Student
Department of Linguistics and English Language,
Lancaster University,
LA1 4YL, UK
v.charoenchaikorn1@lancaster.ac.uk, cvararin@hotmail.com

This study has been reviewed and approved by the Faculty of Arts and Social Sciences and Lancaster Management School's Research Ethics Committee.

Thank you for considering your participation in this project.

Appendix B: Consent form

Lancaster University
Linguistics and English Language

Consent Form (Translated from Thai)

Project title: Language learning through online chat tasks

Researcher: Vararin Charoenchaikorn, PhD Student

Email: v.charoenchaikorn1@lancaster.ac.uk, cvararin@hotmail.com

Please tick each box:

1.) I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2.) I agree to my actions on computer being recorded i.e. computer screen and keys pressed on keyboard being recorded. ☐

3.) I agree to the recording of videos during interview. ☐

4.) I agree to the audio recording during tasks and interview. ☐

5.) I understand that the data of my actions on computer will be anonymised and may be used by the researcher in reports, publications, and presentations. ☐

6.) I understand that anonymised quotes from the interview may be used by the researcher in reports, publications, and presentations. ☐

7.) I understand that the video recordings may be used by the researcher in presentations and teaching only if I give consent to it. ☐

8.) I understand that the audio recordings may be used by the researcher in presentations and teaching only if I give consent to it. ☐

9.) I give my consent to use my video recordings during presentations and/or teaching. ☐

10.) I give my consent to use my voice recordings during presentations and/or teaching. ☐

11.) I understand that my name will not appear on the questionnaires and that any information provided by me in the questionnaires may be used by the researcher in reports, publications and presentations. ☐

12.) I understand that my participation is entirely voluntary and that I have the right to withdraw from the project at any time during the experiment and up to 2 weeks after taking part in this study without having to give any reason. If I withdraw within this period, the information I have provided will not be used for the project. ☐

13.) I understand that all the information obtained during the course of the research will be kept securely according to University guidelines for a minimum of 10 years after the end of the study. ☐

14.) I have received a copy of this consent form and of the accompanying information sheet. ☐

15.) I agree to take part in the above study. ☐

_____	_____	_____
Name of Participant	Date	Signature

I confirm that the participant was given an opportunity to ask questions about the study, and all the questions asked by the participant have been answered correctly and to the best of my ability. I confirm that the individual has not been coerced into giving consent, and the consent has been given freely and voluntarily.

Signature of Researcher_____ **Date** _____

One copy of this form will be given to the participant and the original kept in the files of the researcher at Lancaster University

Appendix C: Main tasks

Camera task instructions

กิจกรรม: พัฒนาการของกล้องถ่ายภาพ

สถานการณ์

คุณและคู่สนทนาของคุณเป็นเพื่อนร่วมชั้นในวิชา 'การออกแบบผลิตภัณฑ์' ในคาบเรียนครั้งที่แล้ว อาจารย์ของพวกคุณได้พูดถึงพัฒนาการการออกแบบกล้องถ่ายภาพในยุคเริ่มแรก โดยยกตัวอย่างรูปแบบของกล้องถ่ายภาพในยุคต้นจำนวนทั้งสิ้น 4 รูปแบบ แต่ว่าอาจารย์พูดเร็วมากจนคุณและคู่สนทนาของคุณตามไม่ทัน เนื่องจากวิชานี้จะมีการสอบในอาทิตย์หน้า คุณและเพื่อนจึงต้องสนทนาเพื่อแลกเปลี่ยนข้อมูลที่แต่ละคนมี เพื่อที่คุณทั้งสองจะได้มีข้อมูลที่ครบถ้วนเพื่อนำไปใช้สอบ

คำสั่ง

ตอนนี้คุณทั้งคู่ต่างก็มีข้อมูลของกล้องถ่ายภาพเพียงคนละ 2 รูปแบบ (จากจำนวนทั้งสิ้น 4 รูปแบบที่อาจารย์พูดถึง) แต่รูปแบบของกล้องถ่ายภาพที่คุณและเพื่อนของคุณมีนั้นบังเอิญไม่ซ้ำกัน ดังนั้น หากคุณแลกเปลี่ยนข้อมูลกัน คุณแต่ละคนก็จะมีข้อมูลของกล้องครบทั้ง 4 รูปแบบ

ในกิจกรรมนี้ โปรดคุยกับเพื่อนของคุณผ่านทางแชทเพื่อรับข้อมูลของกล้องถ่ายภาพ 2 รูปแบบที่คุณไม่มี และอธิบายรูปแบบกล้องถ่ายภาพที่คุณมีทั้ง 2 รูปแบบให้เพื่อนของคุณทราบ**อย่างละเอียดที่สุด** โดยคำนึงว่าข้อมูลนี้จะต้องนำไปใช้ในการสอบ คุณต้องอธิบายลักษณะรูปร่างกล้องถ่ายภาพโดยละเอียด รายละเอียดขององค์ประกอบภายนอกต่างๆ สี และ วัสดุที่ใช้ทำผลิตภัณฑ์

นอกจากนี้ จงปรึกษาหารือกันด้วยว่า ในจำนวนรูปแบบกล้องถ่ายภาพทั้ง 4 นี้ รูปแบบใดได้รับการออกแบบขึ้นเป็นลำดับแรกสุด และรูปแบบใดถูกพัฒนาขึ้นเป็นลำดับท้ายสุด โดยคำตอบต้องมาจากความเห็นพ้องต้องกันของคุณทั้งคู่ และมาจากข้อสันนิษฐานที่สมเหตุสมผล

คุณมีเวลาทั้งสิ้น 25 นาทีเพื่อทำกิจกรรมนี้ ผู้วิจัยจะมาเตือนคุณเมื่อเวลาผ่านไปครึ่งหนึ่ง เมื่อคุณมีเวลาเหลือ 5 นาที และเมื่อหมดเวลา ถ้าหากคุณทำกิจกรรมเสร็จก่อนหมดเวลา กรุณาแจ้งผู้วิจัยโดยทันที คุณสามารถใช้ปากกาที่ผู้วิจัยได้เตรียมไว้ให้เพื่อเขียน/วาดลงบนใบกิจกรรมได้

คำศัพท์

ทองเหลือง = Brass

ด้ามจับ = Grip

ช่องมองภาพขณะถ่าย = Viewfinder

Camera task instructions translated from Thai

Task: The Evolution of Cameras

Situation

You and your partner are classmates in the 'Product Design' course. In the last class, your teacher lectured about the evolution of camera design by talking about 4 early designs of cameras. However, the teacher spoke too fast and neither of you could follow his lecture very well. As there will be a test next week, you and your partner have to exchange the information each of you have in order to get the complete information to prepare for the test.

Instructions

Each of you currently has the information of only 2 designs (from the total of 4). Fortunately, none of the designs both of you have is the same. Therefore, once you have finished exchanging information, you will have the information of all 4 designs.

In this task, please chat with your partner to obtain information of the 2 designs you missed. In return, please describe all of the 2 designs you have to your partner **as detailed as possible** by keeping in mind that the information will be used for the test. You must describe details regarding the cameras' shape, external features and colour, and the materials they are made from.

In addition, together with your partner, discuss which camera design that both of you think is the first and the last invented among the 4 designs. The answers should be agreed upon by both of you and should be justified.

You have 25 minutes to complete the task. The researcher will notify you when you are halfway through the task, when you have 5 minutes left and when the time is up. However, if you finish early, please notify the researcher immediately after completion. During the task, you may use the pen provided to write/ draw on the task sheet.

Vocabulary

ทองเหลือง = Brass

ด้ามจับ = Grip

ช่องมองภาพขณะถ่าย = Viewfinder

Camera task: Materials for student A

รูป A

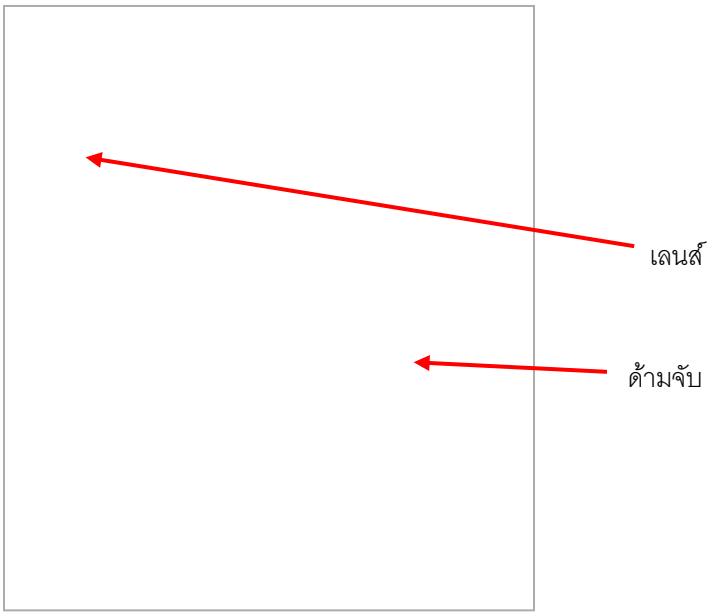


Image has been removed due to Copyright restrictions (Image from: http://www.cameramuseum.ch/data/dataimages/upload/zoom_ESCOPETTE-1889.jpg)

รูป C

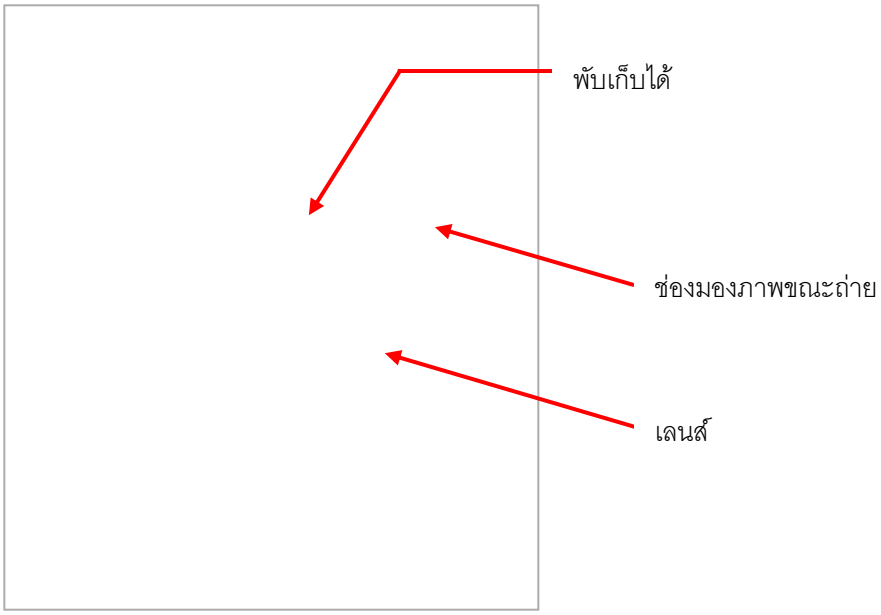


Image has been removed due to Copyright restrictions (Image from: http://www.retiredcameras.com/images/eastman/3afoldingpocketb4/3afpb4front.jpg)

รูป B (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)

รูป D (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)

จงเติมคำตอบลงในช่องว่าง โดยเติม A B C หรือ D เพียงคำตอบเดียวในแต่ละข้อ:

รูป _____ เป็นกล้องถ่ายภาพที่ได้รับการออกแบบขึ้นเป็นลำดับ**แรก**สุด

รูป _____ เป็นกล้องถ่ายภาพที่ได้รับการออกแบบขึ้นเป็นลำดับ**ท้าย**สุด

Camera task: Materials for student A translated from Thai

Picture A

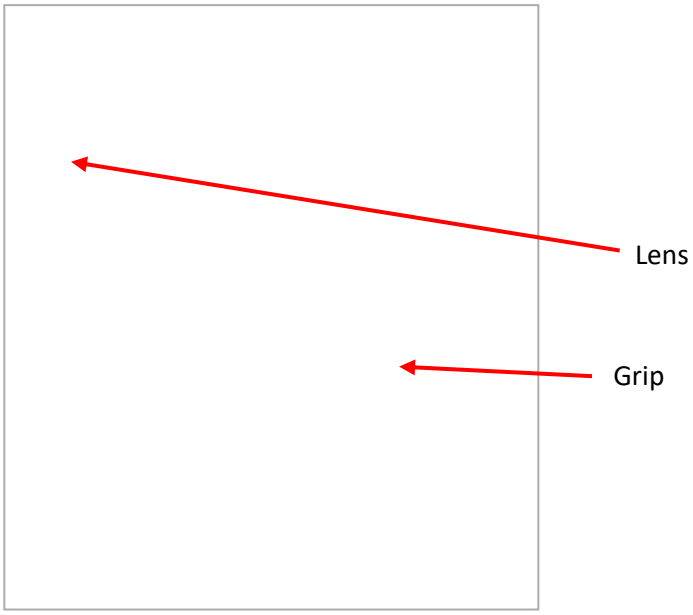


Image has been removed due to Copyright restrictions (Image from: http://www.cameramuseum.ch/data/dataimages/upload/zoom_ESCOPETTE-1889.jpg)

Picture C

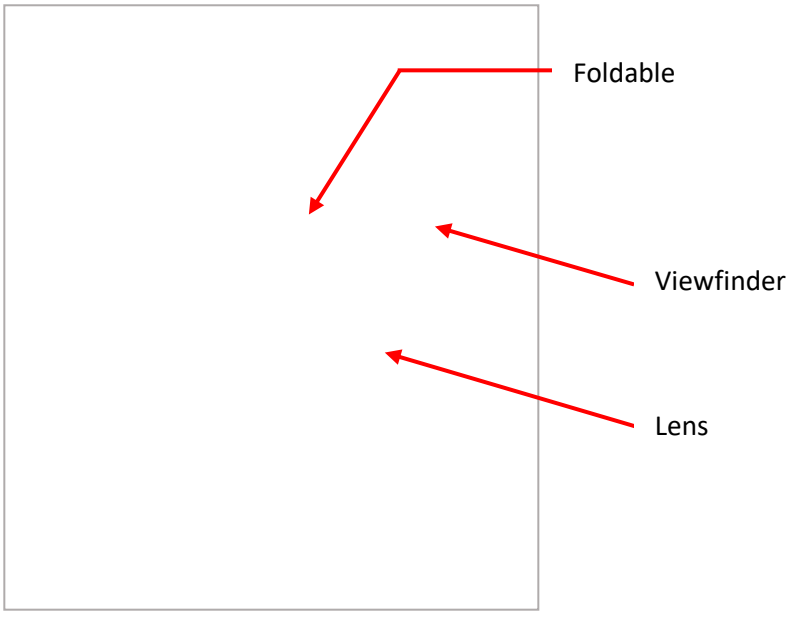
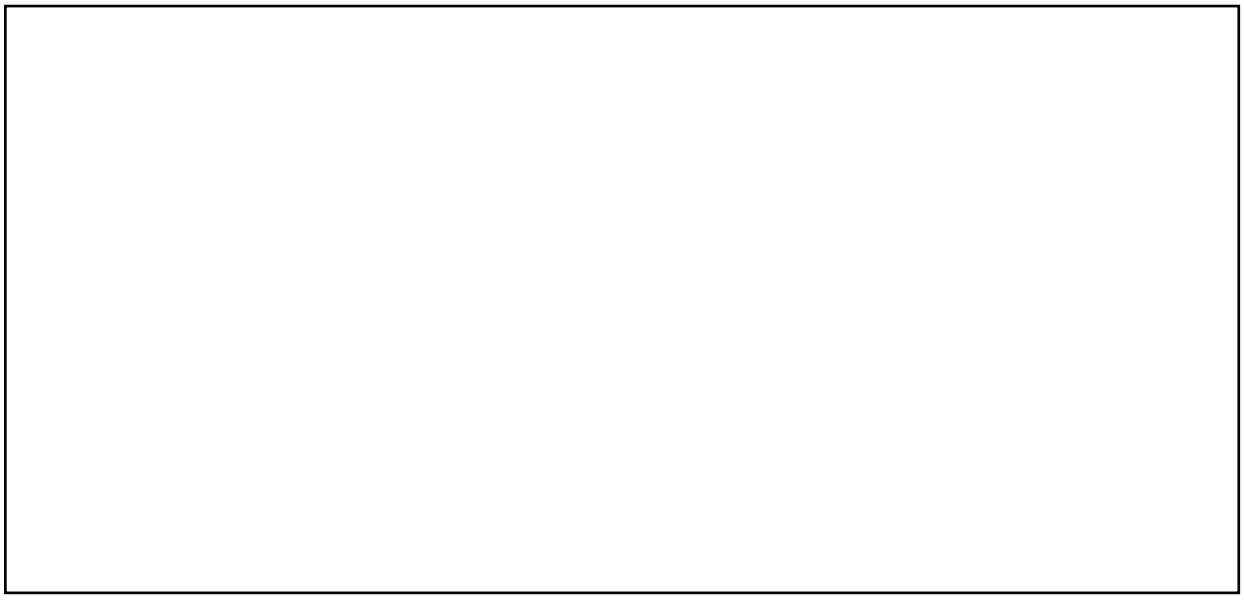
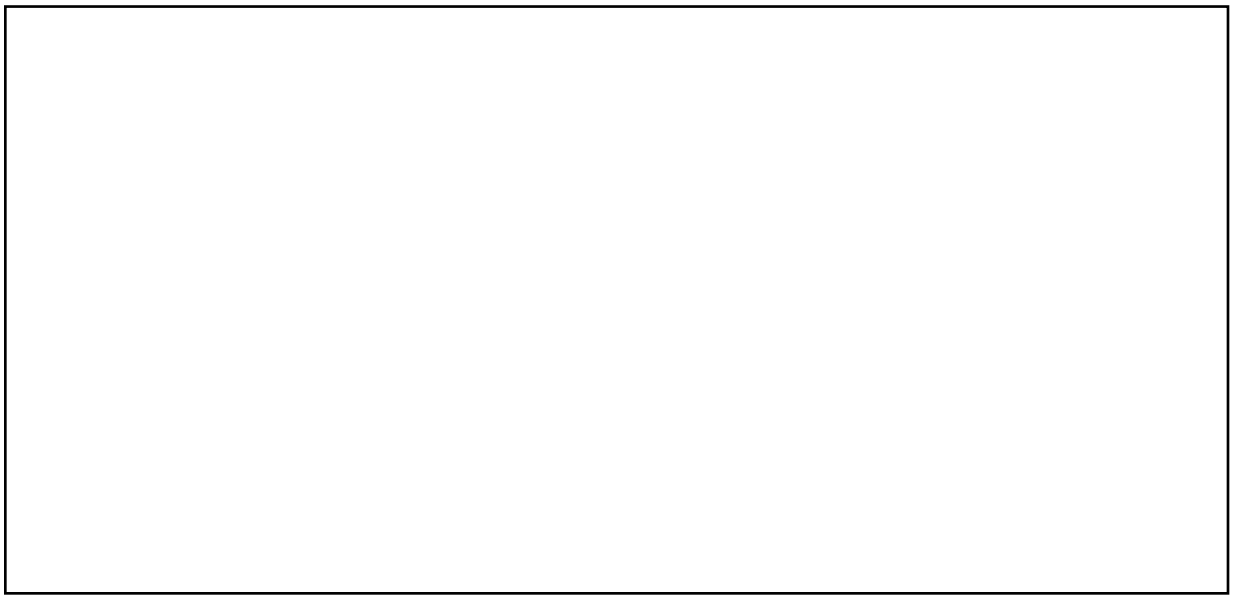


Image has been removed due to Copyright restrictions (Image from: http://www.retiredcameras.com/images/eastman/3afoldingpocketb4/3afpb4front.jpg)

Picture B (You may use the information you get from your partner to draw/ write in this box.)



Picture D (You may use the information you get from your partner to draw/ write in this box.)



Please fill ONE answer in each blank by choosing from A, B, C, or D:

Picture _____ shows the design invented **first**.

Picture _____ shows the design invented **last**.

Camera task: Materials for student B

รูป B

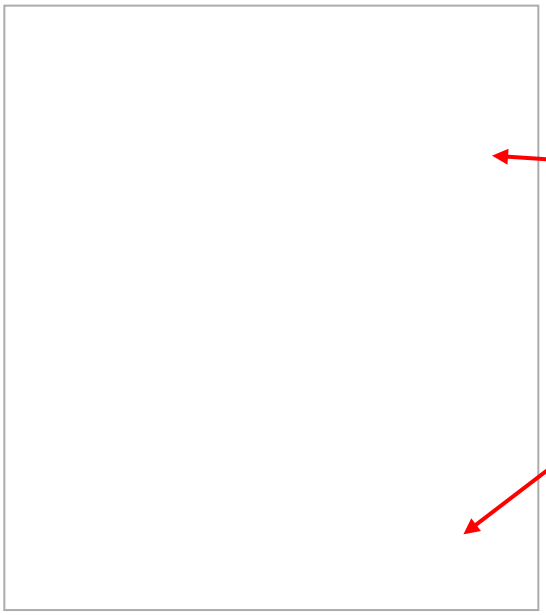


Image has been removed due to Copyright restrictions (Image from: https://collectiblend.com/Cameras/images/Voigtlander-Daguerreotype-(Metallcamera,replica).jpg)

รูป D

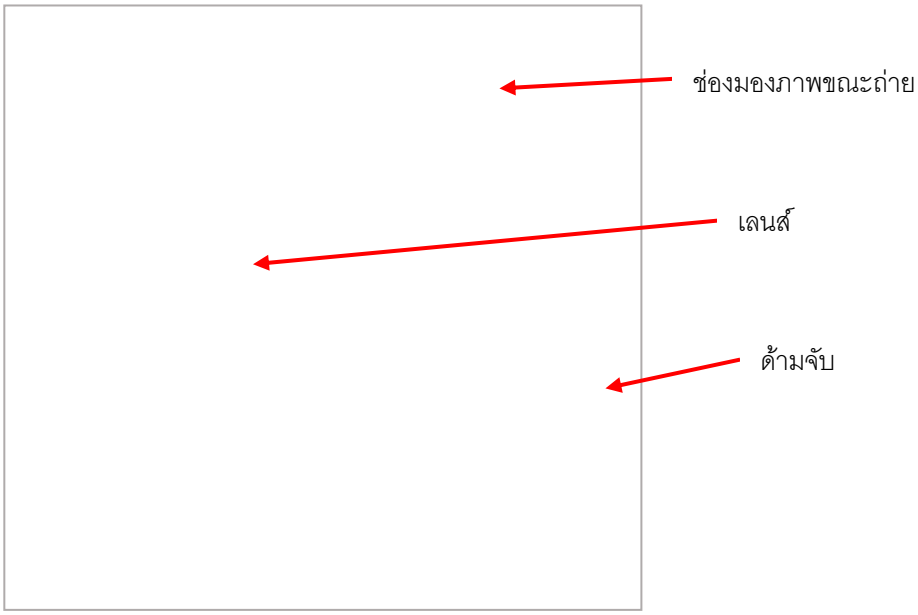


Image has been removed due to Copyright restrictions (Image from: https://blog.scienceandmediamuseum.org.uk/wp-content/uploads/2017/03/thompsons-revolver-camera.jpg)

รูป A (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)

รูป C (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)

จงเติมคำตอบลงในช่องว่าง โดยเติม A B C หรือ D เพียงคำตอบเดียวในแต่ละข้อ:

รูป _____ เป็นกล้องถ่ายภาพที่ได้รับการออกแบบขึ้นเป็นลำดับ**แรก**สุด

รูป _____ เป็นกล้องถ่ายภาพที่ได้รับการออกแบบขึ้นเป็นลำดับ**ท้าย**สุด

Camera task: Materials for student B translated from Thai

Picture B

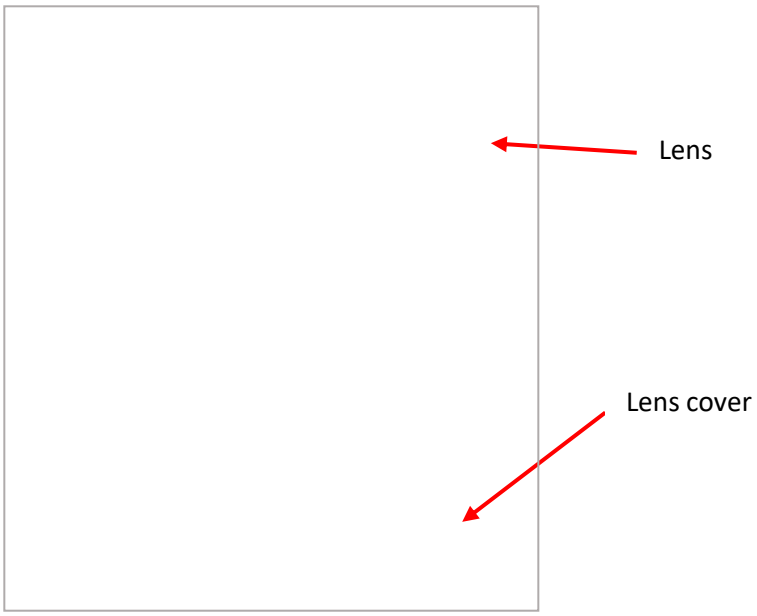


Image has been removed due to Copyright restrictions (Image from: https://collectiblend.com/Cameras/images/Voigtlander-Daguerreotype-(Metallcamera,replica).jpg)

Picture D

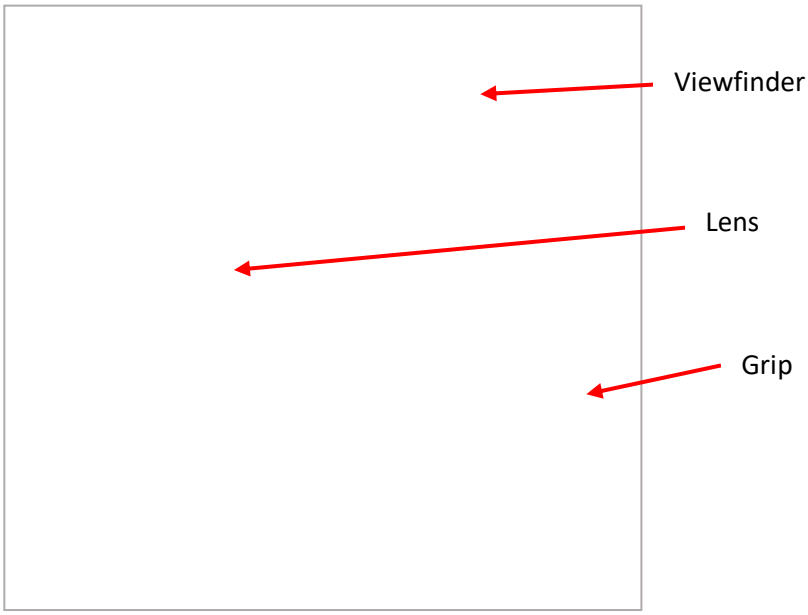


Image has been removed due to Copyright restrictions (Image from: https://blog.scienceandmediamuseum.org.uk/wp-content/uploads/2017/03/thompsons-revolver-camera.jpg)

Picture A (You may use the information you get from your partner to draw/ write in this box.)

Picture C (You may use the information you get from your partner to draw/ write in this box.)

Please fill ONE answer in each blank by choosing from A, B, C, or D:

Picture _____ shows the design invented **first**.

Picture _____ shows the design invented **last**.

Television task instructions

กิจกรรม: พัฒนาการของโทรทัศน์

สถานการณ์

คุณและคู่สนทนาของคุณเป็นเพื่อนร่วมชั้นในวิชา 'การออกแบบผลิตภัณฑ์' ในคาบเรียนครั้งที่แล้ว อาจารย์ของพวกคุณได้พูดถึงพัฒนาการการออกแบบโทรทัศน์ในยุคเริ่มแรก โดยยกตัวอย่างรูปแบบของโทรทัศน์ในยุคต้นจำนวนทั้งสิ้น 4 รูปแบบ แต่ว่าอาจารย์พูดเร็วมากจนคุณและคู่สนทนาของคุณตามไม่ทัน เนื่องจากวิชานี้จะมีการสอบในอาทิตย์หน้า คุณและเพื่อนจึงต้องสนทนาเพื่อแลกเปลี่ยนข้อมูลที่แต่ละคนมี เพื่อที่คุณทั้งสองจะได้มีข้อมูลที่ครบถ้วนเพื่อนำไปใช้สอบ

คำสั่ง

ตอนนี้คุณทั้งคู่ต่างก็มีข้อมูลของโทรทัศน์เพียงคนละ 2 รูปแบบ (จากจำนวนทั้งสิ้น 4 รูปแบบที่อาจารย์พูดถึง) แต่รูปแบบของโทรทัศน์ที่คุณและเพื่อนของคุณมีนั้นบังเอิญไม่ซ้ำกัน ดังนั้น หากคุณแลกเปลี่ยนข้อมูลกัน คุณแต่ละคนก็จะมีข้อมูลของโทรทัศน์ครบทั้ง 4 รูปแบบ

ในกิจกรรมนี้ โปรดคุยกับเพื่อนของคุณผ่านทางแชทเพื่อรับข้อมูลของโทรทัศน์ 2 รูปแบบที่คุณไม่มี และอธิบายรูปแบบโทรทัศน์ที่คุณมีทั้ง 2 รูปแบบให้เพื่อนของคุณทราบ**อย่างละเอียดที่สุด** โดยคำนึงว่าข้อมูลนี้จะต้องนำไปใช้ในการสอบ คุณต้องอธิบายลักษณะรูปร่างโทรทัศน์โดยละเอียด รายละเอียดขององค์ประกอบภายนอกต่างๆ สี และ วัสดุที่ใช้ทำผลิตภัณฑ์

นอกจากนี้ จงปรึกษาหารือกันด้วยว่า ในจำนวนรูปแบบโทรทัศน์ทั้ง 4 นี้ รูปแบบใดได้รับการออกแบบขึ้นเป็นลำดับแรกสุด และรูปแบบใดถูกพัฒนาขึ้นเป็นลำดับท้ายสุด โดยคำตอบต้องมาจากความเห็นพ้องต้องกันของคุณทั้งคู่ และมาจากข้อสันนิษฐานที่สมเหตุสมผล

คุณมีเวลาทั้งสิ้น 25 นาทีเพื่อทำกิจกรรมนี้ ผู้วิจัยจะมาเตือนคุณเมื่อเวลาผ่านไปครึ่งหนึ่ง เมื่อคุณมีเวลาเหลือ 5 นาที และเมื่อหมดเวลา ถ้าหากคุณทำกิจกรรมเสร็จก่อนหมดเวลา กรุณาแจ้งผู้วิจัยโดยทันที คุณสามารถใช้ปากกาที่ผู้วิจัยได้เตรียมไว้ให้เพื่อเขียน/วาดลงบนใบกิจกรรมได้

คำศัพท์

หน้าปัดแสดงคลื่นความถี่ = Frequency receiver dial

ปุ่มปรับ = Knob

ลำโพง = Loudspeaker

Television task instructions translated from Thai

Task: The Evolution of Televisions

Situation

You and your partner are classmates in the 'Product Design' course. In the last class, your teacher lectured about the evolution of television design by talking about 4 early designs of televisions. However, the teacher spoke too fast and neither of you could follow his lecture very well. As there will be a test next week, you and your partner have to exchange the information each of you have in order to get the complete information to prepare for the test.

Instructions

Each of you currently has the information of only 2 designs (from the total of 4). Fortunately, none of the designs both of you have is the same. Therefore, once you have finished exchanging information, you will have the information of all 4 designs.

In this task, please chat with your partner to obtain information of the 2 designs you missed. In return, please describe all of the 2 designs you have to your partner **as detailed as possible** by keeping in mind that the information will be used for the test. You must describe details regarding the televisions' shape, external features and colour, and the materials they are made from.

In addition, together with your partner, discuss which television design that both of you think is the first and the last invented among the 4 designs. The answers should be agreed upon by both of you and should be justified.

You have 25 minutes to complete the task. The researcher will notify you when you are halfway through the task, when you have 5 minutes left and when the time is up. However, if you finish early, please notify the researcher immediately after completion. During the task, you may use the pen provided to write/ draw on the task sheet.

Vocabulary

หน้าปัดแสดงคลื่นความถี่ = Frequency receiver dial

ปุ่มปรับ = Knob

ลำโพง = Loudspeaker

Television task: Materials for student A

รูป A

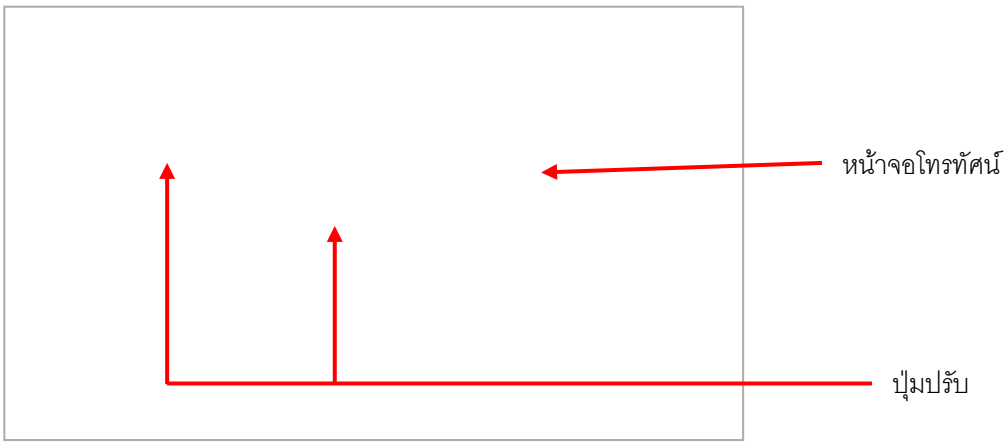


Image has been removed due to Copyright restrictions (Image from: https://s3-eu-west-1.amazonaws.com/lowres-picturecabinet.com/43/main/15/93759.jpg)

รูป C

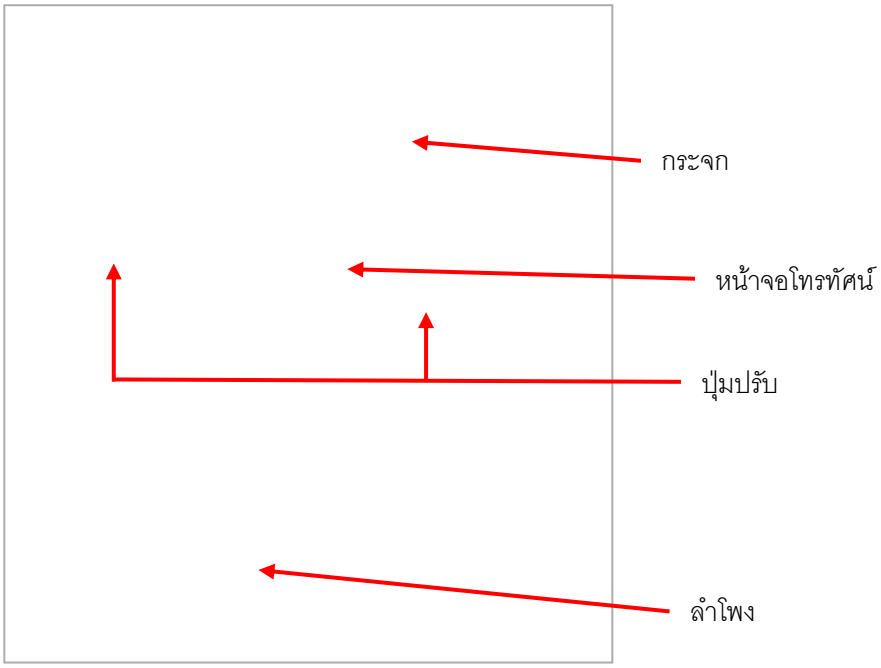


Image has been removed due to Copyright restrictions (Image from: http://www.earlytelevision.org/images/marconi_702-hd.jpg)

รูป B (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)

รูป D (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)

จงเติมคำตอบลงในช่องว่าง โดยเติม A B C หรือ D เพียงคำตอบเดียวในแต่ละข้อ:

รูป _____ เป็นโทรทัศน์ที่ได้รับการออกแบบขึ้นเป็นลำดับ**แรก**สุด

รูป _____ เป็นโทรทัศน์ที่ได้รับการออกแบบขึ้นเป็นลำดับ**ท้าย**สุด

Television task: Materials for student A translated from Thai

Picture A

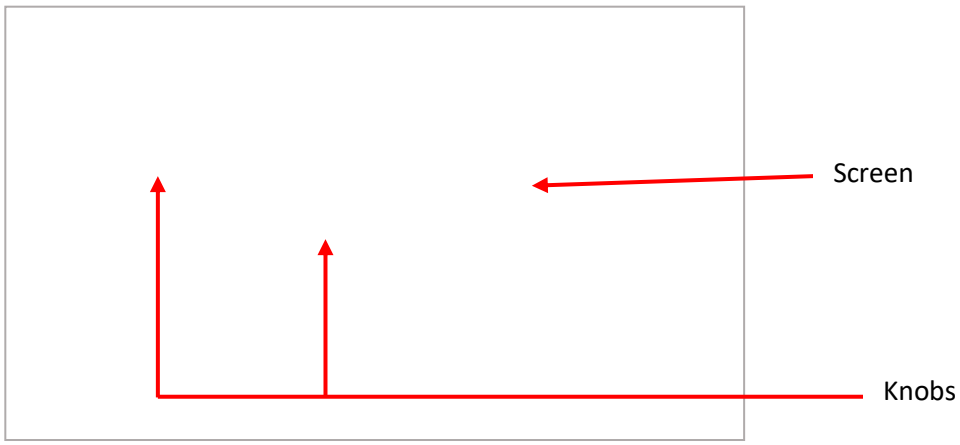


Image has been removed due to Copyright restrictions (Image from: https://s3-eu-west-1.amazonaws.com/lowres-picturecabinet.com/43/main/15/93759.jpg)

Picture C

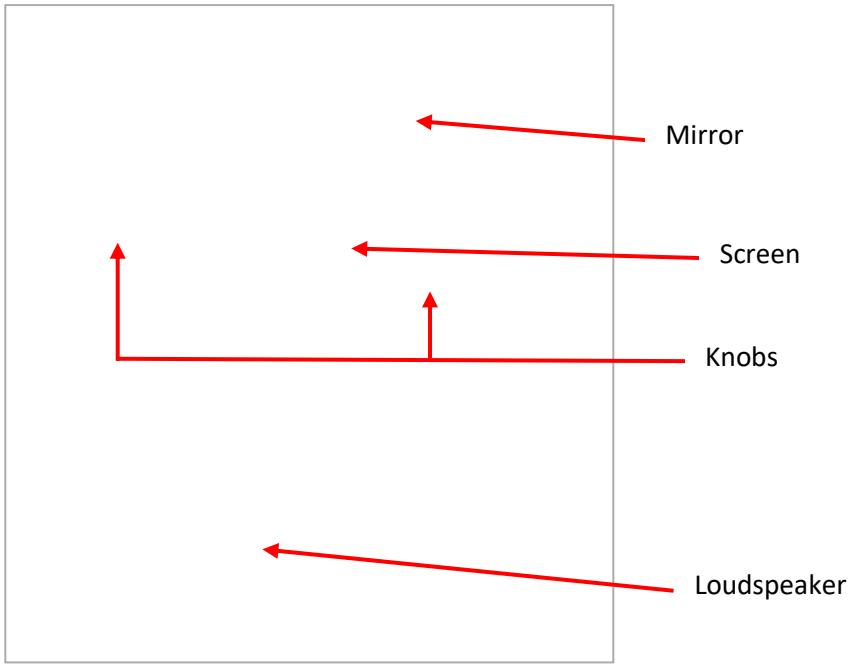
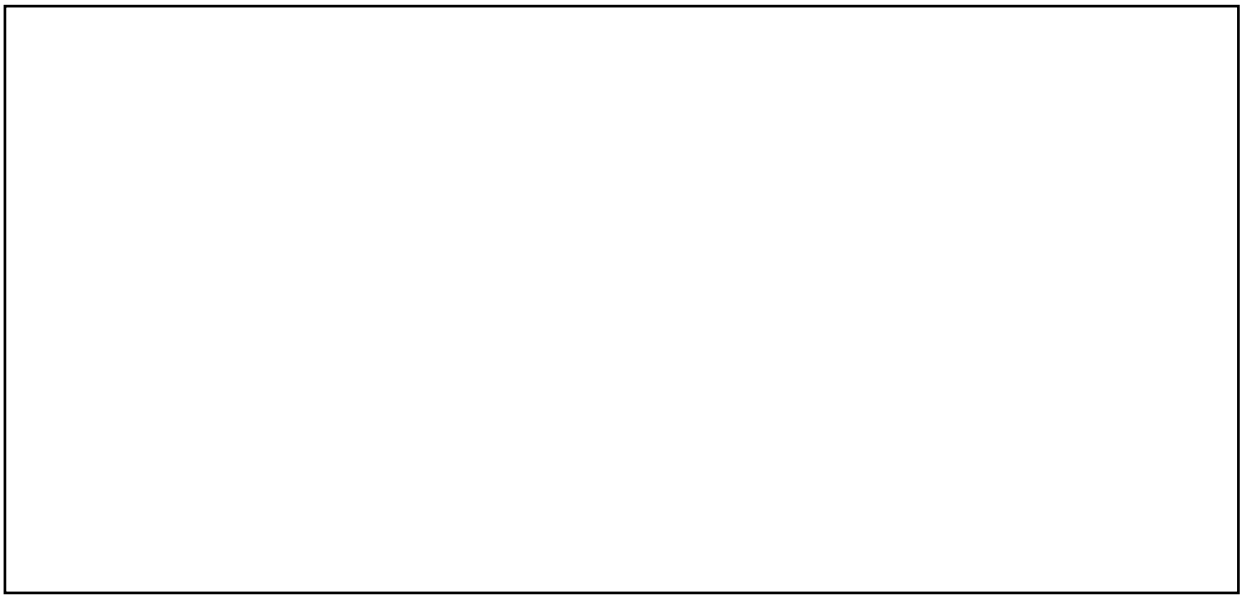
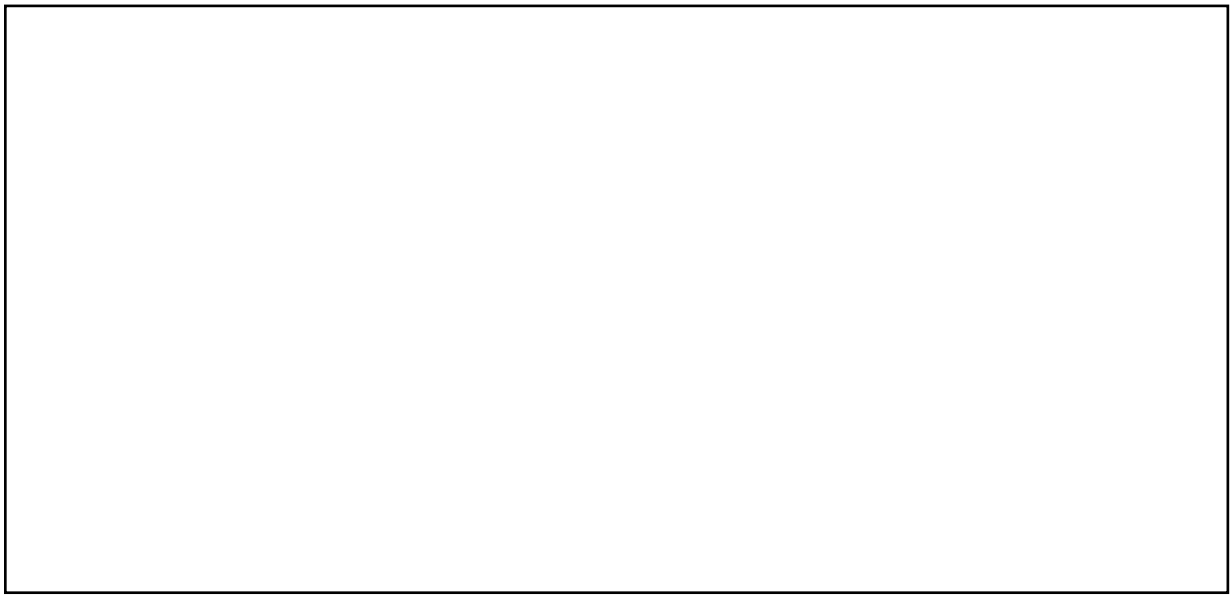


Image has been removed due to Copyright restrictions (Image from: http://www.earlytelevision.org/images/marconi_702-hd.jpg)

Picture B (You may use the information you get from your partner to draw/ write in this box.)



Picture D (You may use the information you get from your partner to draw/ write in this box.)



Please fill ONE answer in each blank by choosing from A, B, C, or D:

Picture _____ shows the design invented **first**.

Picture _____ shows the design invented **last**.

Television task: Materials for student B

รูป B

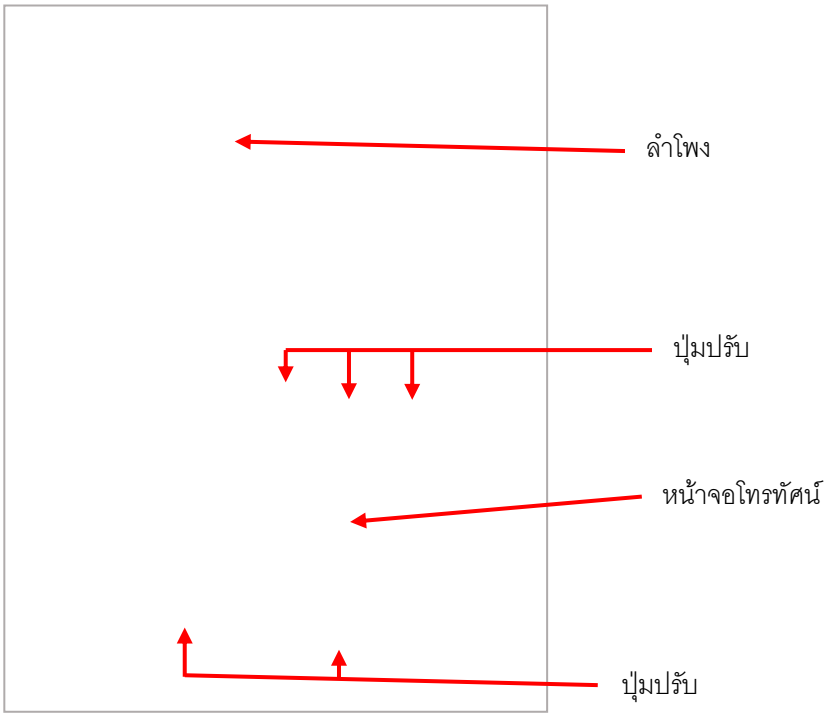


Image has been removed due to Copyright restrictions (Image from: http://image1.thematicnews.com/uploads/images/00/00/41/2015/01/31/fa9452b687.jpg)

รูป D

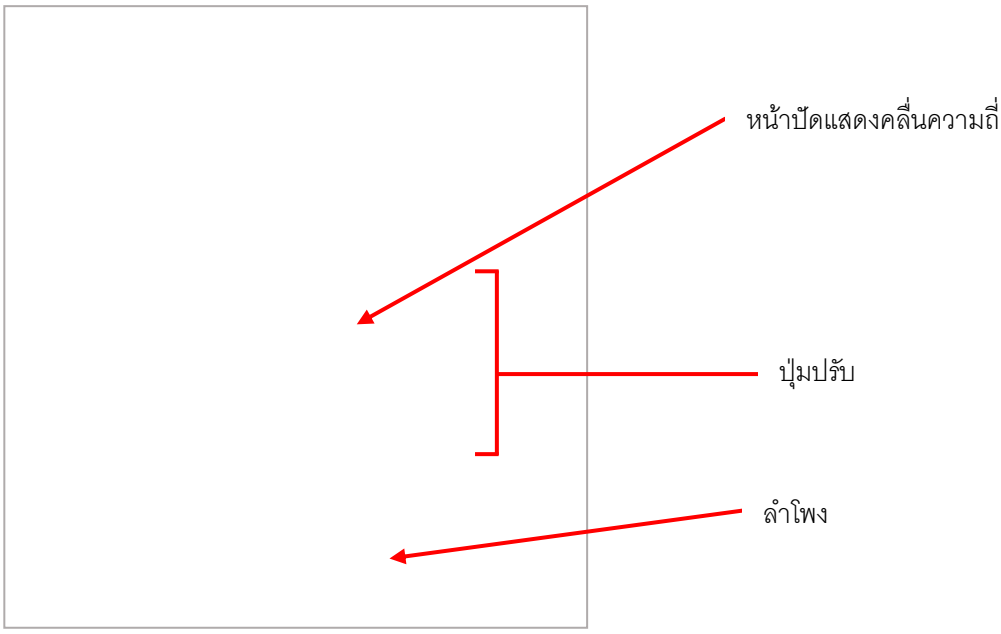
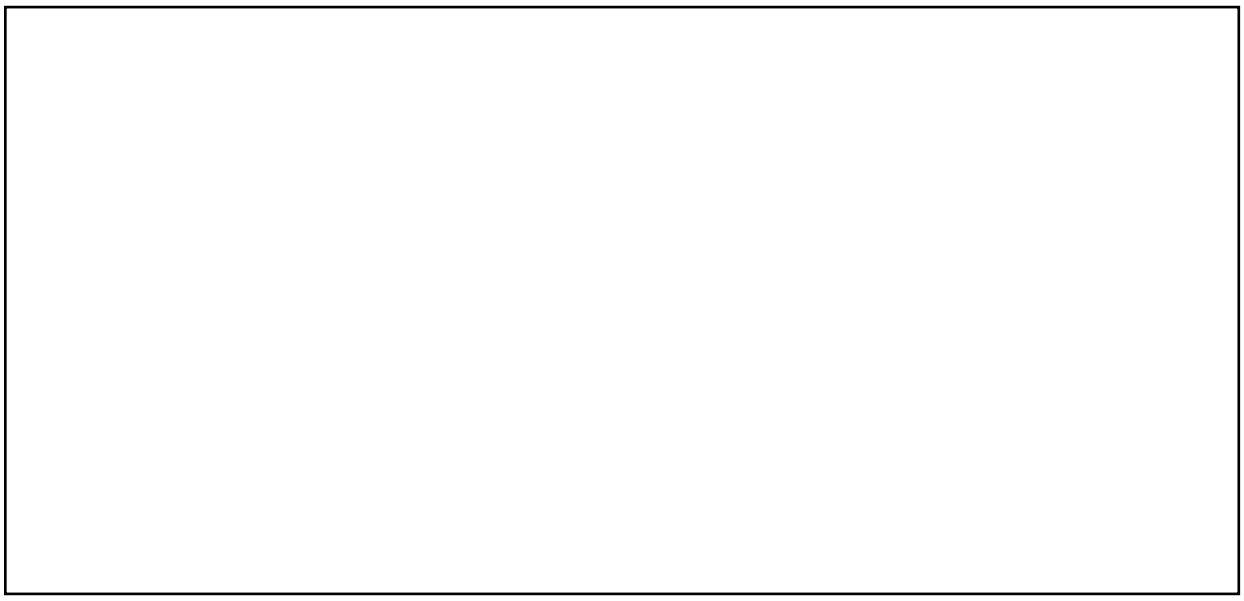
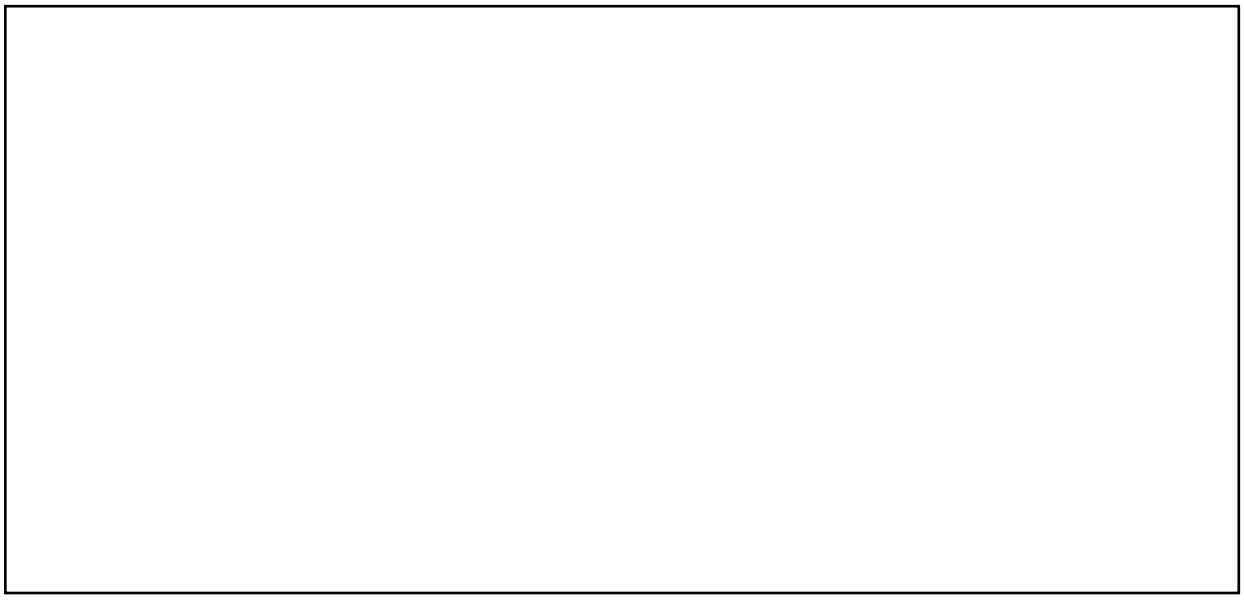


Image has been removed due to Copyright restrictions (Image from: https://www.earlytelevision.org/images/cossor_137t-hd.jpg)

รูป A (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)



รูป C (คุณสามารถนำข้อมูลที่ได้จากเพื่อน มาวาด/เขียนลงในช่องด้านล่างได้)



จงเติมคำตอบลงในช่องว่าง โดยเติม A B C หรือ D เพียงคำตอบเดียวในแต่ละข้อ:

รูป _____ เป็นโทรทัศน์ที่ได้รับการออกแบบขึ้นเป็นลำดับ**แรก**สุด

รูป _____ เป็นโทรทัศน์ที่ได้รับการออกแบบขึ้นเป็นลำดับ**ท้าย**สุด

Television task: Materials for student B translated from Thai

Picture B

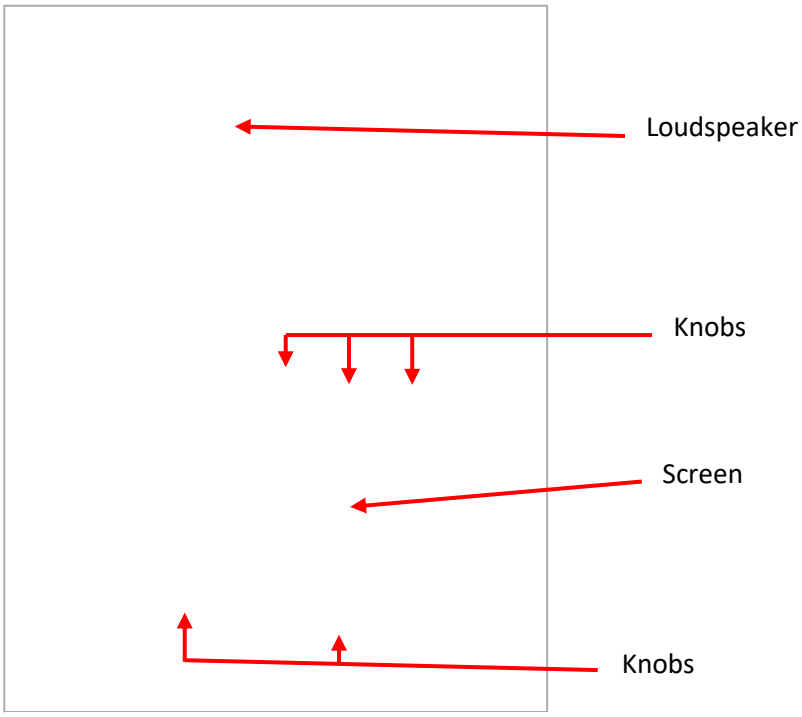


Image has been removed due to Copyright restrictions (Image from: http://image1.thematicnews.com/uploads/images/00/00/41/2015/01/31/fa9452b687.jpg)

Picture D

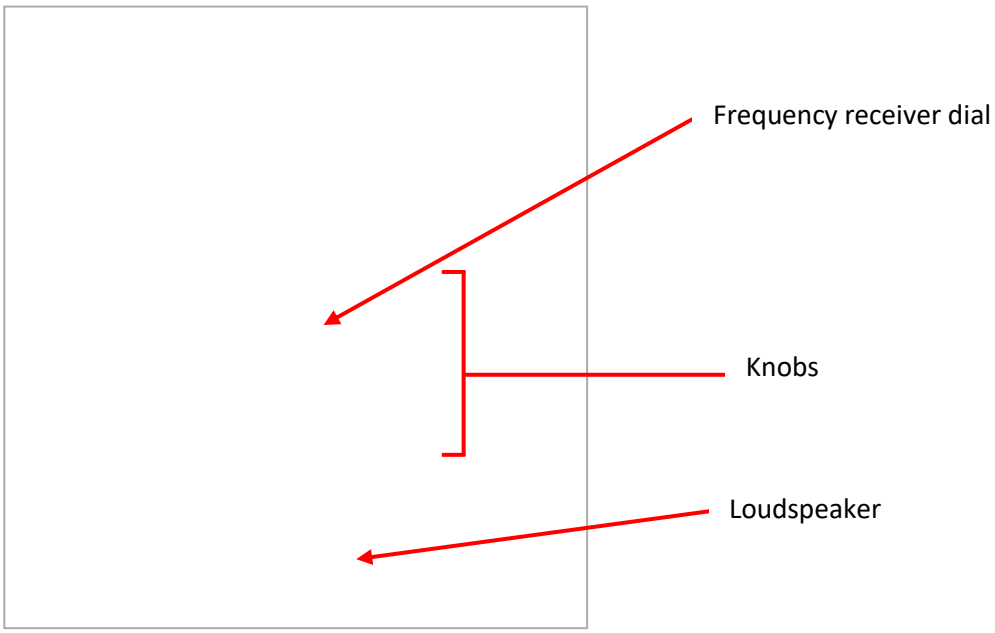


Image has been removed due to Copyright restrictions (Image from: https://www.earlytelevision.org/images/cossor_137t-hd.jpg)

Picture A (You may use the information you get from your partner to draw/ write in this box.)

Picture C (You may use the information you get from your partner to draw/ write in this box.)

Please fill ONE answer in each blank by choosing from A, B, C, or D:

Picture _____ shows the design invented **first**.

Picture _____ shows the design invented **last**.

Appendix D: Background questionnaire

Background Questionnaire (Translated from Thai)

1. Gender: ❑ Female ❑ Male

2. Age: _____ years

3. Native language(s): _____

4. How many years have you been learning English? _____ years

5. Did you, in addition to official school language instruction, learn English in a different way? If so, please tick the relevant boxes. (You may choose more than one option.)

❑ Study abroad ❑ Private tutoring / Cram school

❑ Independent study ❑ Other: _____

6. Have you spent a longer period (three months or more) in any English-speaking country? If so, please indicate the name of the country, the year(s) of visit and the length of visit for each stay.

Country	Year(s) of visit	Length of visit

7. How important are language accuracy, complexity and fluency to you, when 1) communicating in English and 2) communicating via online chat. Below are descriptions of the terms *accuracy*, *complexity* and *fluency*.

Accuracy: When you are concerned about accuracy, you try to make as few language errors as possible.

Complexity: When you are concerned about complexity, you try to show your capability to use advanced language and your wide language knowledge by, for instance, using uncommon words and complex sentence structures.

Fluency: When you are concerned about fluency, you try to produce language without too many pauses or hesitations.

	Not important at all →						Extremely important
7.1 Communicating in English (please choose only one box in each row.)							
Accuracy	❑	❑	❑	❑	❑	❑	❑
Complexity	❑	❑	❑	❑	❑	❑	❑
Fluency	❑	❑	❑	❑	❑	❑	❑
7.2 Communicating via online chat (please choose only one box in each row.)							
Accuracy	❑	❑	❑	❑	❑	❑	❑
Complexity	❑	❑	❑	❑	❑	❑	❑
Fluency	❑	❑	❑	❑	❑	❑	❑

8. How much time **per week**, on average, do you spend using computers? Please choose one of the following:

❑ Less than half an hour ❑ Half an hour – 1 hour ❑ 1–2 hours
❑ 3–4 hours ❑ 5–6 hours ❑ More than 6 hours

9. How much time **per week**, on average, do you spend using a computer or smartphone for the following online activities? Please choose only one answer for each activity.

	Less than half an hour	Half an hour – 1 hour	1–2h	3–4h	5–6h	7–8h	More than 8 h
Instant messaging	❑	❑	❑	❑	❑	❑	❑
Video chat	❑	❑	❑	❑	❑	❑	❑
Writing/reading e-mails	❑	❑	❑	❑	❑	❑	❑
Participating (writing/reading) in online discussion boards	❑	❑	❑	❑	❑	❑	❑
Writing/reading online blogs	❑	❑	❑	❑	❑	❑	❑
Other: _____	❑	❑	❑	❑	❑	❑	❑

10. How often do you use **<u>English</u>** in the following online activities? Please choose only one answer for each activity.

	Always	Often	Sometimes	Rarely	Never
Instant messaging	❑	❑	❑	❑	❑
Video chat	❑	❑	❑	❑	❑
Using social media (e.g. Facebook)	❑	❑	❑	❑	❑
Surfing the internet	❑	❑	❑	❑	❑
Writing/reading e-mails	❑	❑	❑	❑	❑

Participating (writing/reading) in online discussion boards	❑	❑	❑	❑	❑
Watching videos or movies online	❑	❑	❑	❑	❑
Other: _____	❑	❑	❑	❑	❑

11. Please rate your English proficiency in speaking and writing based on the information below. Please choose only one box for each skill.

Speaking (Please choose only one box.)

❑ *Basic user (level 1):* I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.

❑ *Basic user (level 2):* I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.

❑ *Independent user (level 1):* I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).

❑ *Independent user (level 2):* I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.

❑ *Proficient user (level 1):* I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.

❑ *Proficient user (level 2):* I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.

Writing (Please choose only one box.)

❑ *Basic user (level 1):* I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.

❑ *Basic user (level 2):* I can write short, simple notes and messages relating to matters in areas of immediate need. I can write a very simple personal letter, for example thanking someone for something.

❑ *Independent user (level 1):* I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.

❑ *Independent user (level 2):* I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.

❑ *Proficient user (level 1):* I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select a style appropriate to the reader in mind.

❑ *Proficient user (level 2):* I can write clear, smoothly flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.

12. Please rate your competence in <u>communicating in English via online chat</u>.

Poor ⟶ Excellent

❑ ❑ ❑ ❑ ❑ ❑ ❑

13. Did you and your partner know each other before the experiment? ❑ Yes ❑ No

If yes, please indicate your relationship with him/her.

❑ Acquaintances ❑ Friends ❑ Good friends

14. Can you touch type (i.e. type without having to look at the keys on keyboard)?

❑ Yes, I can ❑ No, I cannot

Appendix E: Exit questionnaire

Exit Questionnaire (Translated from Thai)

	Very strong influence	Strong influence	Moderate influence	Slight influence	No influence
1. Before you began the (camera/ television) task, you were told that afterwards you would have to perform an individual language correction task in which you would have to re-read the chat conversation to correct your own language. Please rate the extent to which the knowledge of this task influenced you to try to use English more accurately while performing the (camera/ television) task.	❑	❑	❑	❑	❑
2. Before you began the (camera/ television) task, you were told that afterwards you would have to perform a collaborative language correction task in which you would have to re-read the chat conversation and discuss with your partner in order to make language corrections together. Please rate the extent to which the knowledge of this task influenced you to try to use English more accurately while performing the (camera/ television) task.	❑	❑	❑	❑	❑

For your answer in (1), please indicate why:

_____

_____

For your answer in (2), please indicate why:

_____

_____

3. Are there any aspects of language (e.g. vocabulary and grammar) which you tend to revise or edit when you are chatting online in English? If yes, what are they?

_____
_____

4. 4.1 Did you notice your own English language errors or instances where you needed to improve your language when chatting online during this experiment?

❑ Yes ❑ No ❑ Not sure

4.2 Do you usually notice your own English language errors or instances where you need to improve your language more during online chatting or speaking?

❑ Online chat ❑ Speaking ❑ Not sure

For your answer in (4.2), why?

_____
_____

5. Do you usually correct your English language errors more during online chatting or speaking?

❑ Online chat ❑ Speaking ❑ Not sure

Why?

_____
_____

6. While chatting online in English, which of the following influence your language revision, i.e. making changes to your language by, for example, deleting, adding, replacing or merging some parts of the text? Please also explain how each of the options you choose affects your revision. (You may choose more than one option. If you think none of these influence your revision, you may choose none of them.)

❑ Relationship between you and your chat partner
How?_____

❑ Your typing ability
How?_____

❑ Time constraints during tasks
How?_____

❑ Task type (e.g. description, argumentation and discussion)
How?_____

❑ Other(s):_____
How?_____

7. During the online chat tasks performed in this experiment, were there any instances where you realized that your English language was not accurate but you still sent messages to your partner without making corrections? If yes, why did not you make corrections?

_____

_____

8. Do you agree that online chatting can facilitate learning English?

- ❑ Strongly agree
- ❑ Agree
- ❑ Neither agree nor disagree
- ❑ Disagree
- ❑ Strongly disagree

9. Please rate each task by indicating whether you agree with each statement in the table below. (Choose only one answer for each statement.)

	Strongly agree	**Agree**	**Neither agree nor disagree**	**Disagree**	**Strongly disagree**
Camera Task					
9.1 The task is interesting.	❑	❑	❑	❑	❑
9.2 The task is useful for learning English.	❑	❑	❑	❑	❑
9.3 The task is difficult.	❑	❑	❑	❑	❑
9.4 The task is stressful.	❑	❑	❑	❑	❑
Television Task					
9.5 The task is interesting.	❑	❑	❑	❑	❑
9.6 The task is useful for learning English.	❑	❑	❑	❑	❑
9.7 The task is difficult.	❑	❑	❑	❑	❑
9.8 The task is stressful.	❑	❑	❑	❑	❑
Collaborative language correction task					
9.9 The task is interesting.	❑	❑	❑	❑	❑
9.10 The task is useful for learning English.	❑	❑	❑	❑	❑
9.11 The task is difficult.	❑	❑	❑	❑	❑
9.12 The task is stressful.	❑	❑	❑	❑	❑
Individual language correction task					
9.13 The task is interesting.	❑	❑	❑	❑	❑
9.14 The task is useful for learning English.	❑	❑	❑	❑	❑
9.15 The task is difficult.	❑	❑	❑	❑	❑
9.16 The task is stressful.	❑	❑	❑	❑	❑

For your answer in (9.3), please indicate why

_____

For your answer in (9.7), please indicate why

_____

10. Did you have enough time to complete the camera task?

❑ Yes ❑ No, because_____

Did you have enough time to complete the television task?

❑ Yes ❑ No, because_____

11. Did you experience any problems during the experiment, e.g. technical problems?

_____

_____

12. Is there anything else you would like to say about this project?

_____

_____

Appendix F: Van Waes and Schellens' (2003) taxonomy of revision behaviour and Stevenson, Schoonen, and de Glopper's (2006) multi-dimensional revision taxonomy

Van Waes and Schellens' taxonomy of revision behaviour (2003, p. 837)

1. Number of revisions
2. Type of revision: Addition, deletion, substitution, reordering
3. Level of revision: Letter, word, phrase, sentence, paragraph; layout, punctuation
4. Purpose of revision: Correction of typing errors, revision of form, revision of content
5. Location of revision: Title, first paragraph, first sentence of paragraph, elsewhere
6. Remoteness of revision: *Measured in terms of the number of lines above or below the point of inscription*
7. Temporal location of revision: Stage, segment, unit

Stevenson, Schoonen, and de Glopper's multi-dimensional revision taxonomy (2006, p. 211)

1. Orientation: Content, typing, language
   Language revisions were classified into subcategories according to the target of revision (i.e. grammar, vocabulary, punctuation, phrasing and spelling) and whether they were error- or non-error-triggered. Error revisions were further coded for outcomes: successful or not successful.
2. Domain: Clause and above, below-clause, below-word
3. Location: Pre-text, point of inscription, previous text
4. Action: Addition, deletion, substitution, other

Appendix G: Revisions classified based on surface structures vs stimulated recall comments

The table below shows how the revisions initially classified into each category were coded the second time based on stimulated recall comments. The numbers of matched classifications are marked in bold. Other numbers indicate the numbers of unmatched classifications. For instance, the first column shows that 330 revisions initially coded as content revisions were coded as content revisions again in the second coding. Three and five of the revisions initially classified as content revisions were coded the second time as lexical and rephrasing revisions, respectively. The row 'Total' shows the total number of revisions in each category which were coded again based on stimulated recall data, and the last row shows the proportion of these revisions which were classified into the same category again in the second coding.

Initially coded based on surface structures / Coded based on stimulated recall	Content	Grammar	Lexis	Reorganiza-tion	Rephrasing	Unclear
Content	**330**	3	3	0	0	15
Grammar	0	**65**	0	0	0	8
Lexis	3	2	**57**	0	0	4
Reorganization	0	0	0	**4**	0	0
Rephrasing	5	0	0	0	**8**	0
Typo-Spelling	0	3	3	0	0	3
Total:	338	73	63	4	8	30
% of correct classification:	97.63	89.04	90.48	100	100	

Appendix H: Correlations between different measures of quantity of revision, error correction rate, accuracy and speed fluency

Quantity of revision (N = 168)

			Covert revisions (per 100 words in the final text)	All revisions (per 100 words in the final text)	Characters (including spaces) in the final text/ characters (including spaces) typed
Spearman's rho	Covert revisions (per 100 words in the final text)	Correlation coefficient		**.997****	-.769**
		p (2-tailed)		.000	.000
	All revisions (per 100 words in the final text)	Correlation coefficient			-.761**
		p (2-tailed)			.000

Note. ** Correlation is significant at the 0.01 level (2-tailed).
Correlation coefficient values lower than -.9 and higher than .9 are marked in bold.

Error correction rate (N = 168)

			Grammatical error correction rate	Lexical error correction rate	Total (grammatical + lexical) error correction rate
Spearman's rho	Grammatical error correction rate	Correlation coefficient		-.056	.786**
		p (2-tailed)		.468	.000
	Lexical error correction rate	Correlation coefficient			.469**
		p (2-tailed)			.000

Note. ** Correlation is significant at the 0.01 level (2-tailed).

Accuracy (N = 168)

			Grammatical errors made (per 100 words typed)	Lexical errors made (per 100 words typed)	Total (grammatical + lexical) errors made (per 100 words typed)	Final text accuracy: Grammatical and lexical errors (per 100 words in the final text)
Spearman's rho	Grammatical errors made (per 100 words typed)	Correlation coefficient		.316**	.877**	.856**
		p (2-tailed)		.000	.000	.000
	Lexical errors made (per 100 words typed)	Correlation coefficient			.674**	.656**
		p (2-tailed)			.000	.000
	Total (grammatical + lexical) errors made (per 100 words typed)	Correlation coefficient				**.957****
		p (2-tailed)				.000

Note. ** Correlation is significant at the 0.01 level (2-tailed).
Correlation coefficient values lower than -.9 and higher than .9 are marked in bold.

Speed fluency (N = 168)

			Mean characters between pauses
Spearman's rho	Mins per character including spaces typed	Correlation coefficient	-.776**
		p (2-tailed)	.000

Note. ** Correlation is significant at the 0.01 level (2-tailed).

Appendix I: Wilcoxon signed-rank tests performed using the pilot data to compare the two main tasks (camera and television tasks)

	p	*z*	*r*
Quantity of revision			
Covert revisions (per 100 words in the final text)	.959	-0.05	-.02
All revisions (per 100 words in the final text)	.959	-0.05	-.02
Characters including spaces in the final text/typed	.114	-1.58	-.50
Linguistic unit (revisions per 100 words in the final text)			
Clause and above	.917	-0.11	-.03
Phrase	.508	-0.66	-.21
Word	.878	-0.15	-.05
Letter	.169	-1.38	-.44
Punctuation	.345	-0.94	-.30
Focus (revisions per 100 words in the final text)			
Content	.575	-0.56	-.18
Grammar	.799	-0.26	-.08
Lexis	.508	-0.66	-.21
Reorganization	.080	-1.75	-.55
Trigger (revisions per 100 words in the final text)			
Error-triggered	.333	-0.97	-.31
Non-error-triggered	.386	-0.87	-.27
Outcome of error revision (%)			
Successful error revision	.230	-1.20	-.40
Error correction rate (%)			
Grammatical	.386	-0.87	-.27
Lexical	.575	-0.56	-.18
Accuracy			
Lexical errors per 100 words typed	.386	-0.87	-.27
Grammatical errors per 100 words typed	.333	-0.97	-.31
Speed fluency			
Minutes/character (including spaces) typed	.285	-1.07	-.34
Mean characters between pauses	.230	-1.20	-.40
Perceived task difficulty			
Perceived task difficulty	.705	-0.38	-.12

Note. The p-values were computed based on the asymptotic significance, 2-tailed, test.

Appendix J: Descriptive statistics arranged by group (control and experimental) and type of post-task anticipation (individual and collaborative)

Revision behaviours

			Mean	SD
Quantity of revision				
All revisions (per 100 words in the final text)	Control		13.61	6.72
	Experimental	Individual	14.03	5.15
		Collaborative	14.14	6.36
		Total	14.08	5.76
	Total		13.93	6.08
Characters including spaces in the final text/typed	Control		0.76	0.07
	Experimental	Individual	0.77	0.08
		Collaborative	0.77	0.08
		Total	0.77	0.08
	Total		0.77	0.08
Linguistic unit (revisions per 100 words in the final text)				
Clause and above	Control		1.05	1.05
	Experimental	Individual	1.17	0.95
		Collaborative	1.31	1.14
		Total	1.24	1.05
	Total		1.18	1.05
Phrase	Control		3.43	1.97
	Experimental	Individual	3.49	1.92
		Collaborative	3.56	2.34
		Total	3.53	2.13
	Total		3.50	2.07
Word	Control		7.14	3.89
	Experimental	Individual	7.65	3.41
		Collaborative	7.47	3.91
		Total	7.56	3.65
	Total		7.42	3.73
Letter	Control		1.84	1.34
	Experimental	Individual	1.68	1.01
		Collaborative	1.77	1.25
		Total	1.73	1.13
	Total		1.76	1.20

			Mean	SD
Focus (revisions per 100 words in the final text)				
Content	Control		9.42	4.91
	Experimental	Individual	8.70	3.55
		Collaborative	9.19	4.94
		Total	8.94	4.29
	Total		9.10	4.50
Grammar	Control		1.67	1.25
	Experimental	Individual	2.30	1.43
		Collaborative	2.16	1.29
		Total	2.23	1.36
	Total		2.04	1.35
Lexis	Control		1.23	1.11
	Experimental	Individual	1.76	1.30
		Collaborative	1.59	1.21
		Total	1.68	1.25
	Total		1.53	1.22
Trigger (revisions per 100 words in the final text)				
Error-triggered	Control		1.75	1.51
	Experimental	Individual	2.18	1.31
		Collaborative	2.16	1.13
		Total	2.17	1.22
	Total		2.03	1.33
Non-error-triggered	Control		1.20	0.96
	Experimental	Individual	1.82	1.39
		Collaborative	1.54	1.19
		Total	1.68	1.29
	Total		1.52	1.21
Outcome of error revision (%)				
Successful error revision	Control		85.76	20.85
	Experimental	Individual	81.87	21.39
		Collaborative	76.22	26.62
		Total	79.07	24.18
	Total		81.34	23.26
Error correction rate (%)				
Grammatical	Control		16.53	13.90
	Experimental	Individual	16.54	10.97
		Collaborative	15.34	11.82
		Total	15.94	11.37
	Total		16.14	12.23

			Mean	SD
Lexical	Control		13.63	18.19
	Experimental	Individual	14.13	14.97
		Collaborative	15.57	13.69
		Total	14.85	14.30
	Total		14.44	15.66
Total (grammatical + lexical)	Control		14.60	10.62
	Experimental	Individual	15.60	8.43
		Collaborative	14.95	7.81
		Total	15.28	8.10
	Total		15.05	8.99

Speed fluency

			Mean	SD
Minutes/character (including spaces) typed	Control		0.01	0.00
	Experimental	Individual	0.01	0.00
		Collaborative	0.01	0.00
		Total	0.01	0.00
	Total		0.01	0.00
Mean characters between pauses	Control		13.70	5.83
	Experimental	Individual	12.00	4.70
		Collaborative	12.01	4.38
		Total	12.00	4.53
	Total		12.57	5.04

Accuracy

			Mean	SD
Accuracy during writing (errors per 100 words typed)				
Lexical errors	Control		3.77	2.09
	Experimental	Individual	4.31	1.75
		Collaborative	4.45	2.41
		Total	4.38	2.10
	Total		4.17	2.11
Grammatical errors	Control		6.93	4.33
	Experimental	Individual	7.85	3.39
		Collaborative	8.40	3.32
		Total	8.13	3.36
	Total		7.73	3.74

			Mean	SD
Total (grammatical + lexical) errors	Control		10.70	5.62
	Experimental	Individual	12.16	4.10
		Collaborative	12.85	5.01
		Total	12.50	4.57
	Total		11.90	5.00
Final text accuracy				
Total (grammatical + lexical) errors (per 100 words in the final text)	Control		10.59	6.03
	Experimental	Individual	11.70	4.48
		Collaborative	12.71	5.58
		Total	12.21	5.06
	Total		11.67	5.44

Text quantity

			Mean	SD
Total words typed	Control		253.80	82.20
	Experimental	Individual	225.41	75.17
		Collaborative	230.84	66.00
		Total	228.13	70.47
	Total		236.68	75.32
Total words in the final text	Control		212.55	70.51
	Experimental	Individual	187.95	63.54
		Collaborative	192.48	58.67
		Total	190.21	60.92
	Total		197.66	64.93

Stress during the main tasks

Code	Frequency	Percent
Control group		
1	2	3.57
2	17	30.36
3	30	53.57
4	7	12.50
5	0	0.00
Experimental group		
1	4	3.57
2	26	23.21
3	53	47.32
4	24	21.43
5	5	4.46

Code	Frequency	Percent
Individual post-task anticipation		
1	3	5.36
2	13	23.21
3	25	44.64
4	13	23.21
5	2	3.57
Collaborative post-task anticipation		
1	1	1.79
2	13	23.21
3	28	50.00
4	11	19.64
5	3	5.36
Total		
1	6	3.57
2	43	25.60
3	83	49.40
4	31	18.45
5	5	2.98

Note. Lower values indicate higher stress levels.