UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

An Exploratory Study of Primary School Children's Writing Processes in Digital Environments:
The Use of Models as Written Corrective Feedback

Un Estudio Exploratorio de los Procesos de Escritura de Niños de Primaria en Entornos Digitales:
el Uso de Modelos como Feedback Correctivo Escrito

D. Aitor Garcés Manzanera

2021

UNIVERSITY OF MURCIA

Aitor Garcés Manzanera

An exploratory study of primary school children's writing processes in digital environments: The use of models as written corrective feedback

Un estudio exploratorio de los procesos de escritura de niños de Primaria en entornos digitales: el uso de modelos como feedback correctivo escrito

Doctoral Thesis

Supervisors: Dr. Dª. Raquel Criado Sánchez
Dr. Dª. Rosa María Manchón Ruiz

English Department

Faculty of Arts

2021

A mis padres y hermana

To my parents and sister

ACKNOWLEDGEMENTS

The consecution of the present study has an emotional and professional meaning for me. Since my childhood, I have always been a very curious kind of person, and research has provided me with the opportunity of satisfying this curiosity intellectually. Throughout this path, which now comes to an end, I have been able to overcome any difficulty thanks to the support of a number of very special people.

First and foremost, I am immensely indebted to the continuous and valuable guidance from my mentors, Dr. Raquel Criado and Dr. Rosa Manchón, to whom I would like to express my eternal gratitude for having introduced me to the fascinating world of Applied Linguistics. I feel that without their invaluable wisdom, guidance, and support, as well as their precious and constructive suggestions throughout the process, much of this work would have been impossible to start with, and more importantly, to finish. They have well cared for all the decisions I have taken and have provided me with incalculable advice to cover these knowledge gaps which I strongly needed to finish this dissertation. Their exhaustive and priceless feedback along with their insightful suggestions have been an enormous help to me. Thank you for putting your faith in me. Any errors in this dissertation remain my own.

I would like to thank the rest of my colleagues in the research group for their constant show of support. I wish to acknowledge the constructive and stimulating comments received by other colleagues at the international conferences I had the opportunity to attend. Special thanks are due to Sophie McBride, for her help in proofreading the whole text, and her availability. Also, I would also like to give special thanks to Dr. Luke Plonsky, from Northern Arizona University, and Kristin Rock, from the University of Hawai'i, for kindly assisting me with the myriad of data analyses. Also, I owe thanks to my colleagues at the department of English Language and Literature Teaching for their support and encouragement during the process.

I wish to acknowledge that part of the research for this dissertation was done under the auspices of the nationally funded research projects 20832/PI/18 (Aprendizaje Lingüístico en L2 a través de la escritura en entornos tradicitonales y virtuales [writing and L2 learning in pen-and-paper and computer-mediated environments], and PID2019-104353GB-I00 (Procesos y productos de escritura en L2. Descripción y entrenamiento).

I would also like to wholeheartedly thank Gloria and the elementary school for giving me access to the classroom, and allowing that this PhD could be a reality. Without their accessibility, this dissertation would not have been possible.

This study has been a rewarding experience thanks to the constant and lifelong support of my parents, Jose Luis Garcés and María Dolores Manzanera, who have been encouraging me at every moment even when I was at a low ebb and was not able to summon sufficient strength to continue my path. Equally important has been the support of my sister, Mª Ángeles, who has always believed in me.

Similarly, I wish to express my heartfelt gratitude to all my friends for their lifelong support and their endless patience. They have suffered my absence on numerous occasions during this long process. Once again, thanks for your constant support and for helping me endure this stressful and seemingly never-ending process of completing a dissertation. Also, I wish to give a hug to the heavens in honor of our much-beloved Ivan Esono. We miss you a lot.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

L2 writing has been one of the major areas of interest in the recent decades (see Leki et al., 2008; Manchón & Matsuda, 2016), with much of the research focusing on the written product, the language learning affordances of written corrective feedback (WCF) (see Bitchener & Storch, 2016; Manchón, 2009a), or the renewed interest in L2 writing processes from an SLA perspective (see Barkaoui, 2019; Gánem-Gutiérrez & Gilmore, 2018). Research on WCF has been concerned with shedding light on how the provision or not of different types of feedback may affect the quality of the L2 text, as well as providing empirical evidence to the disciplinary debates on the short-term versus long-term effects of the provision and processing of WCF (see Truscott, 1999, 2007; Ferris, 2004). Regarding WCF, few studies (see Coyle & Cánovas, 2019; Luquín & García Mayo, 2021) have explored the role of models as WCF, a highly ecological technique, in young learners' revised texts, and more importantly, how the resulting noticing of linguistic features (when existing) may affect writing processes. Additionally, research in L2 writing processes has been traditionally conducted in pen-and-paper environments (e.g. Roca de Larios, Coyle & Nicolas Conesa, 2016), but novel methodological procedures using keystroke-logging tools (see Lindgren & Sullivan, 2019) have expanded ways to observe the writing process and pausological behavior in an unobtrusive way. In light of this, our study intends to add new empirical evidence to two virtually unexplored research areas when jointly considered: (1) the implementation and temporal distribution of writing processes in an underrepresented population, i.e. young learners, in both studies with writing processes, and in digital environments, and (2) the implementation of model texts as WCF in an attempt to observe how the management of writing processes vary in terms of lexical choices and the generation of ideas based on previous research (e.g. Coyle et al., 2018). We attempted to shed light on the extent to which the provision of model texts as WCF affect the management of young learners' writing processes and children's pausological behavior. To provide an answer to our objectives, we designed a study with 18 Primary school children (aged 10-11), who took part in a three-stage feedback classroom-based experimental research. These were assigned to an experimental (n= 10) and a control (n= 8) group. Using Inputlog 8.0. (Leijten & Van Waes, 2013), children wrote their initial texts on the computer in the first stage.

Then, the experimental group was provided with a model text, and the control group self-edited their texts without feedback. In the third stage, children were asked to rewrite their initial texts. In the analysis, we operationalized the writing process in terms of planning, formulation and revision, using measures of frequency, total duration, edits and words produced. To analyze the pausological behavior, we used a wide range of measures as in previous research including time on task, pausing time, pause frequency, pause duration, pause distribution, and pause location.

Our results revealed marked differences in the writing processes, namely planning and revision. In the case of the former process, young learners receiving WCF made more frequent use of the planning process and increased the time spent on it in comparison with participants in the control group (who did not have access to feedback). This has indicated that children spend more time to planning than what resaerch with adult or high school EFL learners has revealed. As for the effects on formulation, the greatest impact was observed on the frequency of formulation episodes, which was larger in the group receiving WCF, and on the number of edits and words produced during formulation. In terms of revision, the effect of the WCF was more clearly observed in the measure of frequency of revision episodes, and there was an overall increase of time spent on macroscopic revisions in the data of the experimental group. All in all, our study has revealed that children strategically decided where and when to locate each of these writing processes. Regarding pausological behavior, our results revealed that the effect of WCF was more marked in pauses at word boundaries as well as sentence boundaries. WCF seems to mediate a great part of the aforementioned behavioural patterns. These findings were discussed from the perspective of i) enhancing our knowledge about young learners' writing processes and pausological behaviour and thus to complement related previous theories, based on adult learners; ii) shedding light on the potential role of models as WCF to mediate such processes.

Keywords: *L2 writing, model texts, young EFL learners, writing processes, pausological behavior*.

CHAPTER 1. INTRODUCTION

The present dissertation focuses on digital writing as performed by young children. The ultimate aim is to contribute to previous empirical work in two related areas, namely, the study of L2 writing processes and the effects of written corrective feedback (WCF). In this sense, the research conducted intends to shed light on the characterization of the temporal dimension of writing processes of young EFL learners (a neglected population in previous research on the temporal dimension of writing processes), and on the manner in which their writing processes and pausing behavior may be mediated by the provision of written corrective feedback (henceforth, WCF) in the form of models. In what follows we synthesize the rationale behind our global research aims, outline the theoretical and empirical intended contribution of our work, and provide a general outline of the structure of the thesis together with a synthesis of the content of the different chapters in it.

1.1. Rationale

Writing in an additional language (L2) has received considerable scholarly attention in recent decades (as comprehensively reviewed in Leki, Cumming & Silva, 2008; Manchón & Matsuda, 2016; Manchón & Polio, 2021). Much of L2 writing research has dealt with the written product and the purported benefits of different forms of WCF (see Bitchener & Storch, 2016; Manchón, 2009a, 2009b). As more fully elaborated in later chapters, globally considered, this research has shed a strong light on the characteristics of the texts written by L2 users in a variety of settings, as well as the range of potential intervening variables, including variables related to the provision (or lack of) of different types of WCF.

The study of L2 writing has also been approached from the perspective of the writing process itself. In this case, the research intent has been to identify and characterize the different writing processes involved in the production of L2 texts (see reviews in Roca de Larios, Nicolás-Conesa & Coyle, 2016; Michel, Stiefenhöher, Verspoor & Manchón, 2021). Key concerns in this body of work are the study of L2 writers' strategic allocation of attention to writing processes, or the analysis of the complexity of the problem-solving nature of writing. More recently the study of writing processes has been approached from the perspective of the language learning potential of the processing dimension of writing (see Manchón, 2011, 2020a, b; Manchón & Vasylets, 2019; Manchón & Williams, 2016; Williams, 2012). This line of research represents the intersection between writing studies and second language acquisition (SLA studies). The present thesis is framed in this SLA-oriented, L2 writing research strand.

The study of L2 writing (including both the act of writing itself and the appropriation of WCF) as a potential site for language learning has generated growing interest in the recent SLA studies and the Instructed Second Language Acquisition (ISLA) scholarly literature, (as discussed in Bitchener & Storch, 2016; Manchón, 2011a; Manchón & Vasylets, 2019). Accordingly, a dual distinction in the study of writing has been established (Manchón, 2011a): The "learning-to-write" (LW) perspective focuses on the intricacies of the process of gaining knowledge about the written conventions or style and developing the skill of writing in an additional language. In contrast, a "writing-to-learn" (WL) perspective would seek to understand the way in which the act of writing itself and of appropriating feedback can contribute to language learning. Three characteristics have been proposed to explain the benefits of language learning through writing and WCF appropriation (Manchón & Vasylets, 2019; Manchón & Williams, 2016): the availability of time while writing and while processing feedback, the permanence of writing and the type of WCF, and finally, the relevance of the problem-solving nature of writing as well as the depth of processing (DoP) while processing WCF (for a theoretical and empirical state-of-the-art, see Leow, 2020; Leow & Mercer, 2015).

As noted earlier, SLA-oriented L2 writing studies have attempted to shed light on the language learning potential of writing processes (hence following Cumming's [1990] pioneering formulation. See analysis in Manchón, 2020c), and of the appropriation of WCF (as recently discussed by Coyle et al, forthcoming; Roca de Larios & Coyle, 2021). Regarding the latter, the key issues in disciplinary discussions are related to the so called "error correction debate" (Truscott, 1999, 2007; Ferris, 1999, 2004), which ultimately tries to discern the beneficial short-term and long-term effect of the provision and processing of WCF. This distinction has been more reinterpreted within what Manchón (2011b) calls "feedback for accuracy", with a focus on a more immediate effect of the WCF on the L2 learners' accuracy, and "feedback for acquisition", where the aim would be to ascertain the way in which providing learners with WCF and encouraging feedback processing, self-reflection and new output can have longer-lasting effects in the process of language learning.

Research on L2 writing processes from the dual perspective of processes while writing and while processing feedback are of paramount interest in the present disssertation. Regarding the former, the scholarly interest in the description and analysis of writing processes (i.e. planning, formulation, revision, and monitoring) and of online behaviors (especially pausing behavior and writing fluency) has generated a growing amount of research that ultimately responds to a number of theoretical and practical concerns. Research on L2 writing processes has concentrated on the learners' use of cognitive processes while producing a text on the basis of diverse models of writing (e.g. Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes & Flower, 1980; Kellogg et al., 2013; Kellogg, 1996). Substantial empirical evidence has attempted to shed light on the strategic behavior of teenager and adult L2 writers' writing processes, for instance, planning (e.g. see Johnson [2020] for a recent review on the topic), formulation (e.g. Roca de Larios et al., 2006; Zimmerman, 2000) or revision (e.g. Faigley & Witte, 1981 or Stevenson et al., 2006; Xu, 2018, for recent studies on this writing process). These studies have provided empirical support to previous predictions that writing is a recursive process, highly proficiency-dependent, cyclical, and adaptive. An important gap in this body of work is the lack of attention to younger learners. Yet, L2 writing is part of educational policies for the teaching of second languages, and hence part of the language learning experience of thousand of young learners across the world.

Methodologically, earlier studies fundamentally relied on verbal protocols, such as think-aloud and stimulated recalls, and mainly looked at pen-and-paper writing (e.g. Roca de Larios et al., 2008). More recently, digital writing has been added to research agendas and with it a whole set of novel methodological procedures have been added, including diverse forms of keystroke logging tools (see review in Lindgren & Sullivan, 2019; Van Waes et al, 2012). As noted by Van Waes et al (2012), these new technologies have "created new possibilities for writing researchers to investigate writing as it unfolds in real time" (p. 507). The use of keystroke-logging software (henceforth, KLS. Leijten & Van Waes, 2013) has become specially popular, at times strategically combined with other techniques, for instance, by triangulating the keystroke data with eye-gaze recordings or screen capture technologies (Chukharev-Hudilainen et al., 2019; Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2017. See also contributions to Lindgren & Sullivan, 2019; Manchón, 2020; Manchón & Roca, forthcoming/2021),

Regarding the processing of feedback, the bulk of extant research has been framed in both cognitive (for instance, the Noticing Hypothesis, the Output Hypothesis, Focus-on-Form or Skill Learning Theory) and sociocultural theories of L2 learning (as reviewed in Bichener & Storch, 2016; Leow, 2020; Leow & Shu, forthcoming/2021; Polio, 2012). Methodologically, this research has relied on a range of procedures intended to foster and track language learning processes by fostering deeper linguistic processing. In the case of younger learners, such methodological procedures have included alternative techniques such as prompts for guided and intentional WCF processing (e.g. Santos et al., 2010) as well as unguided noticing (e.g. García-Mayo & Labandibar, 2017).

In light of this, recent theoretical accounts of WCF processing have emerged (see Leow 2020, Leow & Suh, 2021). As mentioned in the previous paragraph, the main focus has been placed on the manner in which WCF "has been attended to, detected, or noticed" (Leow, 2020, p. 104), and on the set of potential intervening variables, such as the attentional resources used while processing feedback, the depth of processing, or the level of awareness (Leow, 2020). These are relevant endeavors given the purported connection between depth of processing and potential language learning. As noted by Manchón & Vasylets (2020), "the deeper linguistic processing associated with the meaning-making activity [...] will prompt L2 users to engage in crucial language learning processes" (p. 344). Empirical evidence has supported this view by attesting that learners engage in noticing while processing WCF (see Qi & Lapkin, 2001; Swain, 1995; Swain & Lapkin, 1995). Such processing has been found to be dependent on some mediating variables, such as the type of feedback and the characteristics of the task (see Bitchener & Storch, 2016; Storch, 2016). Under this premise, a recent scholarly interest focuses on the use of diverse types of feedback and a growing number of studies have specifically focused on the use of models as WCF (e.g. Coyle & Cánovas, 2019; García-Mayo & Labandibar, 2017). This type of feedback has been reported to foster a very specific type of cognitive processing, namely, cognitive comparison (Ellis, 2003) as L2 writers are expected to draw on the information contained in a text (see Hanaoka, 2007) and compare it with their own output, thus creating cognitive conflict (Doughty & Williams, 1998). On this basis, studies have endeavored to observe WCF processing by identifying the so-called "Language Related Episodes" (LREs) through collaborative writing, with a special focus on the features noticed in the model (for instance, García-Mayo & Labandibar, 2017; Hanaoka, 2006, 2007; Yang & Zhang, 2010) by adopting guided or unguided noticing strategies.

However, as is also the norm in many other areas of SLA research, as well as SLA-oriented L2 writing research, the available empirical evidence derived primarily from studies conducted with adult (often university) L2 users is an evident problem for the generalization of findings and the advancement of research. Following this trend, very few studies (see Cánovas et al., 2015; Coyle & Cánovas, 2019; Coyle et al., 2018; Coyle et al., 2014; Coyle & Roca de Larios, 2014; Roca de Larios et al., 2015) have investigated the role of models as WCF by young learners. More importantly, few studies on feedback appropriation (and none of the studies on the use of models, with the exception of Coyle et al. [2019]) have investigated L2 writers' noticing as reflected in the management of the writing process. Building on the purported cognitive effort of comparing the model with the learner's own output, this comparison is supposed to generate new ideas obtained from the model text provided, and activate higher-level processes (such as planning or revision) while writing (Schmidt, 2001), hence the key concern of observing the impact of the provision and appropriation of WCF on the time allocation of writing processes as well as the pausological behavior (see Lindgren & Sullivan, 2019, for a wealth of studies on writing processes).

Importantly, studies of WCF have not paid attention to the manner in which the provision of feedback may influence the temporal dimension of writing processes, i.e. whether the temporal distribution of writing processes varies or remains the same in subsequent iteration of writing of the same text. This is an important gap in our knowledge of writing processes and their temporal distribution.

In view of the above considerations, our study is an attempt to fill gaps in previous research regarding the implementation and temporal distribution of writing processes by young learners, an unexplored population in writing processes studies which, as noted earlier, , has focused predominantly on adults (for instance, Roca de Larios et al., 2008 or Révész et al., 2019). Additionaly, we intend to do so by KLS (Leijten & Van Waes, 2013) as our main data collection instrument, and by implementation model as our feedback strategy. The use of KLS responds to calls (add relevant references) for applying new methodological tools to observe children's writing processes and the impact of feedback thereon. The decision to choose models is based on the predicted language affordances of model texts for children on account of their selective attention to lexical choices and the generation of ideas (see empirical studies such as Hanaoka, 2007 or more recent ones such as Coyle et al., 2018, dealing with children's writing). This strategic behavior during feedback use may influence the way in which writing processes are allocated across the time-distributed nature of writing and, very importantly, the pausological behavior associated with it. This way, specific insights are expected to be gained into how young EFL learners produce texts in a digital environment, and how the process of writing is affected by the provision of WCF in the form of models, a feedback strategy whose potential influence on the temporal dimension of writing processes has not been studied so far.

The present study is exploratory in nature given the absence of studies dealing with the influence of WCF in the shape of models on the temporal dimension of the writing process and, more generally, the scarcity of studies on children's L2 writing processes.

1.2. General aim of the thesis

The contribution of this dissertation is envisaged as being theoretical, empirical and methodological in nature. From the theoretical and empirical standpoints, the contribution of this PhD is related to the disciplinary debates on (1) the temporal dimension of writing processes and pausological behavior in young EFL learners (as detailed in Chapter 2), (2) the nature of writing processes in a digital environment, and (3) the potential influence of WCF on L2 writing processes. The methodological contribution of this dissertation is closely related to the light that can be shed on the affordances of the use of a keystroke-logging software (Leijten & Van Waes, 2013), to capture the underlying writing processes in children’s writing. With this we will also follow recent trends to new approaches for analyzing the writing processes by triangulating the data obtained from *Inputlog*, allowing us to uncover the underlying processes. This will be made possible by taking an inferential stand on the pausological data (Barkaoui, 2019; Medimorec & Risko, 2017; Michel et al., 2019). As noted by Matshuhashi (1981), "length of pauses ... and their location in the text ... provide a temporal taxonomy or description of ... aspects of written ... production from which inferences about planning and decision can be made" (p. 114). Adding to the intended contribution of this PhD, most of the studies on writing processes with young EFL learners were traditionally centered on bilingualism (e.g. Edelsky, 1982 or Fagan & Hayden, 1988), or more recently, on writing performance based on the working memory (e.g. Michel et al., 2019). Thus, the assumed relevant contribution of this study gains ground in the exploration of how young learners manage their writing processes in a digital environment, and the impact of WCF on the said processes.

1.3. General outline

This dissertation consists of two major parts, and it is divided into six chapters. **Part one** offers the theoretical background to the empirical study reported in Part 2. **Part two** reports the empirical study, whose main elements (aims, method, results, and discussion) are reported in chapters 3 to 5. Chapter 6 summarizes the main findings and contribution of this dissertation, limitations and future avenues in research.

Chapter 2 presents presents a description of the dimension of writing, with the aim of framing the study in the dimension of "writing-to-learn language". First, we offer an overview of the cognitive theories supporting writing as a language learning site and, more specifically, on the role of WCF in its various types promoting language learning. We outline factors mediating the provision of WCF, and then introduce theoretical and empirical research supporting the use of model texts as a WCF strategy. A review of the relevant empirical to our study is presented, establishing connections between the impact of model texts during the processing and use of feedback. Our analysis of previous empirical research on models points to a predominance of studies of writing in traditional environments (i.e. pen-and-paper), and a focus on written product. We identify this aspect as a research gap that our study aimsto fill with its focus on digital writing and writing processes. Second, an extensive theoretical description of writing processes as a theoretical and empirical construct is provided, outlining the models of L1 writing and L2 writing. Given the nature of our study, we also review previous empirical research on writing processes carried out in traditional and more recentsettings. Building on the recent trend in the use of new methodological approachesin the study of writing processes, we explore, from a theoretical perspective, the intricacies of logging software for the field of writing. Finally, previous research on writing processes, and pausological behavior using keystroke-logging software is put under scrutiny. The analysis of previous empirical research in keystroke-logging studies makes it evident that children are an underrepresented population, which we identify as a gap that our study intends to fill.

Chapter 3 presents the main aims of this PhD dissertation along with the Research Questions (RQs) guiding our study. All the three RQs address three different relevant dimensions, respectively: i) the temporal dimension of writing processes from a global perspective; ii) the potential strategic modifications in the temporal dimension of the writing process after the provision of WCF; and iii) the influence of WCF in the shape of models on pausological behavior.

Chapter 4 describes in detail the methodological basis of this dissertation. First, all aspects regarding the research design are explained plus a global overview of the data collection procedure. Detailed information about relevant methodological aspects are presented: (a) the participants involved in this study (given their nature as young EFL learners); (b) the data collection instruments (specifying their role in the research design); (c) a thorough description of the data collection procedure; (d) the variables implied in the empirical study along with their operationalization and the measures; and (e) the procedure for the data analysis.

Chapter 5 presents the results, which are discussed following the structure of the RQs. Firstly, quantitative results are presented regarding the potential effect of the independent variable (the provision or not of model texts as WCF) on both the temporal dimension of writing processes (RQ1) and pausological behavior (RQ2). These results are reported with descriptive data along with the effect sizes (Hedges's g for the within-groups analyses, and Hedges's $g_{adjusted}$ for the between-groups analyses).

Chapter 6 presents a discussion of the results Firstly, we attempt to shed light on why the alteration of some writing processes (RQ1) and pausological behavior (RQ2) might have occurred as mediated by the presence or absence of WCF. For RQ1, links are established with previous research for each of the writing processes (i.e. planning, formulation and revision). For RQ2, we attempt to establish a connection between the previous research on pausological behavior and the potential impact of model texts as WCF on it. The way in which our research contributes to previous work is also discussed.

Chapter 7 closes the present dissertation by synthesizing the main contribution of the research reported outlining a number of methodological limitations encountered throughout the process, and drawing pedagogical implications on the basis of the findings presented. Finally, a number of suggestions for future research are put forward.

CHAPTER 2. LITERATURE REVIEW

2.1. Introduction

This chapter is organized into three main parts. First, we describe the main trends in the study of L2 writing and we thus identify the writing-to-learn-language (WLL) dimension. An overview of the role of WCF within the SLA-oriented strand in which this PhD is framed will be presented, together with theories on the value of L2 writing as a site for language learning. Additionally, some relevant aspects related to mediating variables (i.e. the type of task, and the type of feedback) in the process of WCF are reviewed. Research on models as a WCF strategy is also reviewed by analyzing what models conceptually represent, their language learning potential, and the specific type of cognitive processing involved in the use of model texts. Second, a review of research on L2 writing processes is provided, with a separate conceptualization of each subprocess (i.e. planning, formulation, and revision). Finally, the last section explores the methodological and research affordances of logging tools, and more specifically, Inputlog. Also, previous research on pausological behavior and keystroke-logging software is reviewed. The chapter ends with a summary of the theoretical and empirical underpinnings reviewed along with the relationship they have with the intended contribution of this PhD.

2.2. Trends in the study of L2 writing: Learning-to-write and writing-to-learn

Writing has been traditionally considered as a visual form of communication (Leow, 2020), which in the case of a second or foreign language (L2), is a multi-faceted phenomenon (Cumming, 2016), which explains the abundant scholarly theoretical and empirical attention paid to L2 writing (see Leki et al., 2008; Manchón & Matsuda, 2016 for overviews), including the links between L2 writing and SLA research (see Manchón & Williams, 2016; Manchón & Polio, In press/2021).

Recent theorizing (see Manchón, 2011a) has resulted in a dual orientation in L2 writing research, namely, research focused on learning to write, on the one hand, and writing to learn language, on the other. This two-fold orientation has been identified as "learning-to-write" (LW) and "writing-to-learn" language (WLL), and as noted by Manchón (2011a), these perspectives "traverse L2 writing scholarship and practice, although they have developed independently from each other" (p. 3). As Manchón (2020) refers to, the latter perspective owes greatly to previous pioneering research in L2 writing, especially Harklau's (2002) work Cumming's (1990) pioneering work on L2 writing as a site for language learning.

In what follows, both dimensions, i.e. LW and WLL, will be further discussed. An outline of their main postulates will be presented along with an overview of the empirical strands into which each of the dimensions has translated.

►***Learning-to-write perspective***

The first of these perspectives, *learning-to-write* (LW), is mainly concerned with the intricacies of gaining knowledge about written conventions or style, and with developing the skill of writing in an additional language. In other words, this dimension involves writing as an end in itself where the aim is to foster the writer's development of literacy skills for diverse purposes (Manchón, 2018). Hyland (2011) distinguishes three foci within research on this dimension, namely one centered on the writer, a second one focused on the product, and a third concerned with the reader. In the case of studies on the writer, the research intent has been to provide them with strategies used by good writers.

The learners' use of different cognitive-related strategies has been a key concern in L2 writing research (as reviewed in Manchón, 2018; Manchón et al., 2007). As noted by Hyland (2011), "what has evolved is a model which emphasizes a planning-writing-reviewing framework" (p. 18). This is in line with what models of L1 writing have propounded in the last three decades (e.g. Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes, 2012). The complexity, as well as the mental process involved in the writing process, make it a cognitively complex and problem-solving activity. Accordingly, from a processing perspective, writing is regarded as a "problem-solving activity rather than an act of communication" (Hyland, 2011, p. 18).

In light of this, the second perspective on writing is concerned with the written product itself, i.e. the text produced by the learner and its complexities. When the focus is placed on the text as an entity (i.e. *object*), the research intent has been mainly centered on the regularities found in texts, with studies on aspects such as stance (Biber, 2006), temporal frequency (Kennedy, 1987), and impersonality and hedging (Shaw & Liu, 1998). As noted by Hyland (2011), "this kind of research has been valuable in revealing salient features of texts and how learners gain control over them" (p. 22). Seen from this perspective, the writer uses the conventions of a text to mold their objectives socially. In Hyland's (2011) words, "treating texts as discourse means seeing writing as social action" (p. 23) and therefore acknowledging that a set of social constraints have to be borne in mind in order to achieve communicative purposes. This has given rise to a considerable body of research on the intricacies of genre approaches from both academic (Swales, 2004) and professional (Bhatia, 2008) perspectives. Based on the concept of “visible pedagogy” (Bernstein, 1990), writing instruction is expected to be more effective if a more explicit approach is taken, enabling students to consciously manipulate their linguistic affordances and decisions (e.g. Hasan, 1996). This allows writers to explore the development of a text as per the purpose, audience, and message. In this respect, "genre approaches encourage us to look for organizational patterns, reminding us that [...] we want our reader to recognize our purpose" (Hyland, 2011, p. 24). As reviewed in Hirvela et al. (2016), genre approaches have also been criticized for the dominance of ideology of a certain culture (see Benesch, 2009), although genre advocators have contended that no teaching approach is exempt from this cultural bias. They have stressed out that, despite the obvious constraining nature of genres and their pattern structure, writers may opt for the best options to ease their expression (Hyland, 2007).

Finally, the third orientation in the study of L2 writing places emphasis on the readers, and how the writer produces the text to address this audience. When one writes a text, this creates a series of expectations for the reader, who has to interpret what the writer's purpose is (Hirvela et al., 2016). This unmistakably leads the writer to produce a thorough text where the assumptions are recognized by the reader, who can comprehend them. Thus, as noted by Hirvela et al. (2016), "LW [*learning-to-write*] involves creating a text that the writer assumes the reader will recognize and expect and the process of reading ... drawing on assumptions about what the writer is trying to do" (p. 50). That is why, pedagogically, this perspective of focusing on the reader has given rise to contextualizing the writing activity by reflecting real-life uses (Hyland, 2011), and has contributed to the extended practice of exposing students to sample texts of the target genre along with peer and teacher feedback (Hirvela et al., 2016). Focusing on the reader entails that writing is globally considered as a social activity, and how it is used by social groups (Hyland, 2011). This view accommodates that writing is not merely a way of representing the socially-constructed world, but also a way of constituting it, which goes in line with the view that LW encourages the discovery of texts as socially interwoven with the purposes of the society and discourse community into which it is integrated.

In short, addressing writing from the perspective of LW means that the act of writing is in itself both an internal cognitive process, and a social practice that, requires the interaction between the writer and the audience. Thus, research has attempted to analyze not only the act of writing by the writer (including the study of the processing dimension of writing) but also the linguistic characteristics of texts and the way of expressing these ideas.

► ***Writing-to-learn perspective***

The second approach to L2 writing corresponds to that of “writing-to-learn” language (WLL), which, as noted above, builds on the pioneering work of Cumming (1990) and ought to be linked to tenets of the Output Hypothesis (e.g. Swain, 1985). Cumming (1990) suggested that the act of writing in an L2 could contribute to the development of L2 linguistic knowledge. He envisioned writing as an activity where writers were prompted to monitor their production, an activity that could help them "analyze and consolidate second language knowledge" (p. 483). Cumming's (1990) work attempted to theorize why attention to form-meaning connections foster the refinement of the L2, thus contributing to language learning

The pioneering study by Cumming (1990) was replicated partially by Swain & Lapkin (1995). Their research focused on the process of noticing during writing and the potential influence that variables such as proficiency and task complexity might have in the process. Their study yielded compelling results for the role of noticing in helping language learners modify their written output, and helped to shed light on the problem-solving nature of writing. This study constituted the baseline for subsequent research which attempted to verify previous predictions on the relationship between writing and L2 development (see Manchón & Vasylets, 2019 for a review).

As reviewed by Leow (2020), the WLL perspective gathered momentum with influential publications (see Manchón & Roca de Larios, 2007, for a position paper, or Manchón, 2009, 2011b) which have endeavored to discuss, theoretically and empirically, why and how writing can be a site for language learning. That is why the exploration of how L2 writers make use of writing for their language learning objectives is deemed relevant in SLA studies, as noted by Manchón (2011b): "the WLL dimension deserves a more prominent place in the L2 research agenda" (p. 5). The discussion of the association of L2 writing to L2 learning has subsequently included the analysis of the role that WCF might have with regards to the purported language learning potential of L2 writing in ISLA contexts (e.g. Bitchener, 2019; Bitchener & Storch, 2016; Leow, 2020), as more fully discussed in a later section.

The affordances of writing for language learning are explained through three main characteristics (as reviewed in Manchón & Vasylets, 2019; Manchón & Williams, 2016; Williams, 2012): the availability of time while writing and during the processing of feedback, the permanent nature of the written text and the WCF provided on texts and, finally, the problem-solving nature of writing and WCF processing.

Two of the unique characteristics of writing, pace (there is more time for writing than for speaking) and permanence (texts are visible and so is feedback) might contribute to better management of writers' attentional resources, facilitating the monitoring of these resources and favoring the focus on linguistic concerns (Manchón & Williams, 2016). The permanence of writing is, furthermore, reported to benefit L2 learners by pushing "*motivated* writers to expect and demand more of themselves" (Ortega, 2012, p. 406). Besides, this permanence is thought to give L2 writers an impetus to prompt greater metalinguistic reflection, enabling them to further notice the gaps and holes in the L2 during the writing process. Similarly, when provided with WCF, L2 writers are reported to engage in noticing processes and analyze the linguistic features in their texts. This noticing effect combined with the permanence of written texts leads to a purported benefit for the L2 writer as they can engage in comparing their output with the input provided. As noted by Manchón & Vasylets (2019), "written texts are ... permanent, and hence potentially useful for testing one's knowledge of the L2 given that L2 users can engage in the process of cognitive comparison" (p. 344). These testing possibilities allow learners to prioritize the incoming input in the form of WCF, and to tap into a series of processes facilitating language learning. This visibility of both their written output and WCF might certainly facilitate the development of higher and more complex goals in L2 writing (Manchón & Vasylets, 2019) as it opens a window of possibility for testing their hypotheses about the L2, as stated in the Output Hypothesis (Swain, 1995, 2005).

An additional characteristic of writing thought to contribute to its language learning potential is the problem-solving nature of both composing and WCF processing. The act of writing as a complex and meaning-making activity induces writers to engage in deeper linguistic processing (see Byrnes & Manchón, 2014; Cumming, 1990), and when this occurs, writers are pushed to examine their explicit knowledge. Theoretical works on writing processes have demonstrated that the linguistic processing that has been observed to take place during writing manifests itself in the actions such as going backward to analyze what one has already written, reflecting upon, reviewing and editing the writer's output (see Manchón, 2011a; Roca de Larios et al., 2002). In a similar fashion, WCF processing potentially allows for cognitive comparison, prompting learners to notice the forms in the feedback (i.e. the input received) in contrast to one's interlanguage (Manchón & Vasylets, 2019).

Generating text in an L2, as anticipated previously, involves engagement in decision-making and problem-solving. The writing process, as a complex activity, ultimately possesses language learning potential because of the heightened attention devoted to language forms, as attested in previous empirical research (e.g. Granfeldt, 2008; Kormos, 2014; Kuiken & Vedder, 2011; Schoonen et al., 2009; Tavakoli, 2014). This attention to form while writing has given rise to the assumption that writing possesses a language learning advantage over speaking (Ortega, 2012). Despite the vast amount of research in the strand of WLL supporting these tenets, there are still central concerns on "whether or not this potential greater linguistic processing may result in better output" (Manchón & Vasylets, 2019, p. 345). In fact, the question of where L2 writing and the processing of WCF produce changes in the writers' knowledge repertoire has caused much debate in recent scholarly discussions, especially concerning the interface issue between explicit and implicit knowledge. Manchón and Williams (2016) commented on this central issue and suggested that writing leads to changes in linguistic knowledge. The permanent nature of writing may increase L2 writers' control of their attentional resources, prioritizing language-related concerns and placing more focus on incoming input in the form of WCF (Manchón & Vasylets, 2019). From the perspective of the problem-solving nature of writing, another possibility points to the benefits that writing might have on the retrieval of explicit L2 knowledge, contributing to either an enhancement of accuracy or the automatization of such explicit knowledge, but without obvious changes to the L2 system (Manchón & Williams, 2016). This resonates well with Bitchener's (2016) claim that the processing of WCF allows for more accurate L2 practice, which is thought to gradually convert declarative knowledge into procedural knowledge.

As anticipated in the previous paragraphs, theoretical and empirical underpinnings (see Polio, 2012 or Bitchener, 2012, for thorough reviews) have pointed to the role of WCF in developing L2 linguistic knowledge. This research is reviewed below, starting with the role of WCF in L2 writing research and then moving on to the discussion of the role that WCF possesses to further language development.

►The role of WCF

WCF is understood as a "written response to a linguistic error that has been made in the writing of a text by an L2 learner" (Bitchener & Storch, 2016, p. 1). Other definitions in the scholarly literature point to WCF as an essential and non-negotiable instructional strategy, which "help[s] L2 learners improve their writing effectiveness" (Kang & Han, 2015, p. 1). WCF has been a long-established feedback form in L2 classrooms (see Van Beuningen et al., 2012). As anticipated in the previous section, one notable reason has been its permanence and visibility, as the written output may be compared with the feedback provided, which may take different forms. Such a characteristic allows for greater availability of time, and hence for more opportunities for reflection. In contrast, oral corrective feedback does not allow for permanency unless it is recorded and subsequently corrected, which unmistakably uses up too much time (Bitchener, 2012). Additionally, the explicit or implicit presence of corrections in WCF, one of its key characteristics, increases the likelihood of noticing (Ashwell, 2000). Among the teacher's intent on facilitating language learning, WCF holds a pivotal role since providing corrective support to learners is thought to consolidate their L2 knowledge and enhance further L2 development. Therefore, feedback provision, processing, and use are key elements in the WLL dimension of L2 writing (Manchón, 2011a).

However, the debate on the learning potential of WCF has faced strident criticism from scholars, especially after Truscott's famous position paper (1996), where he argued that no compelling research evidence led to actual benefits of WCF for L2 knowledge. The origins of this objection date back to the 1970s, when North American L1 composition teachers took a process-oriented approach to writing in their composition lessons, based on Flower & Hayes's (1981) cognitive theory of writing (Hyland et al., 2016). The idea of providing feedback was further reinforced after French immersion programs taking place in Canada (see Swain, 1985) failed to ensure grammatical gains on the part of students, one of the factors being the lack of provision of WCF.

Back to the objections to the language learning affordances of WCF, Truscott (1996) argued that error correction gave the impression of putative learning instead of driving learners towards real language learning. Truscott's (1996, 2007) questioning of the purported benefits of both the provision and processing of WCF were based upon three principal arguments: (i) as learners only uptake certain information from the input for a specific moment, he argued that it would not be available for future use, (ii) learners would not be able to acquire or incorporate a certain form if their interlanguage was not at its precise developmental point, especially if no connection could be made to previously acquired knowledge, thus assuming Piennemann's (1989) readiness concept, and (iii) teachers were not sufficiently able to provide learners with accurate CF in a systematic manner. These claims did not go unnoticed in the scholarly community, and counter-arguments started to appear. Ferris (1999, 2003, 2004) contended against and dismounted Truscott's arguments by relying on previous empirical evidence suggesting that teacher correction may be fruitful and that the provision of CF clearly and selectively may help some writers. The growing body of research that emerged out of the interest of empirically testing the role of WCF on L2 enhancement (see Bitchener, 2008; Bruton, 2009, 2010; Ellis et al., 2008; Kan & Han, 2015) shed light upon the efficacy of WCF and has since then attempted to demonstrate opposing evidence to Truscott's views. The effects of WCF were tested not only on the same text or task but also on new texts (see Sheen, 2007).

Manchón (2011a) reinterpreted the debate by establishing a dichotomy in the understanding of feedback: first, *feedback for accuracy* involves "the provision of WCF to encourage accuracy in learners' use of the L2 shortly after feedback processing" (Nicolás-Conesa et al., 2019, p. n.d.). Second, *feedback for acquisition* refers to fostering longer-term language learning through the learners' involvement in feedback processing, detection and self-reflection on errors, and the new output production.

In what follows, we present the cognitive and sociocultural theories supporting the language learning potential of WCF.

►***The language learning potential (LLP) of WCF***

As discussed in the relevant literature (e.g. Bitchener, 2019; Leow, 2020; Polio, 2012), the purported language learning potential of WCF can be linked to a series of cognitive and sociocultural theories on learning. These cognitive theoretical postulates include relevant SLA conceptualizations such as Skill Acquisition Theory (DeKeyser, 2001, 2007, 2015), the Noticing Hypothesis (Schmidt, 1990, 1994, 2001), the Output Hypothesis (Swain, 1995, 1998, 2000, 2005). Ultimately, we include the Interaction Hypothesis (Long, 1981, 1983; Mackey & Polio, 2009) given its relevance. Krashen's theory, on the other hand, disregard the role of WCF as being beneficial for acquired knowledge, just as the Processability Theory (Pienneman, 2015). These two latter theories will be further discussed below, mainly focusing on the previous postulates, i.e. on cognitive theories, to do justice to their importance for the present dissertation. Within the sociocultural perspectives, the Socio-cultural theory of mind will be briefly explained.

2.3. Relevant theories on the language learning potential of WCF: cognitive and sociocultural perspectives

Cognitive-based theories of second language acquisition (SLA) focus upon the process that students go through when learning an L2. As mentioned previously, numerous theories can be alluded to support the pivotal role held by writing (and the subsequent feedback provided) in ensuring that L2 learning does occur. Along this line, cognitive and sociocultural perspectives provide considerable evidence-based support for the consideration of WCF as a useful instrument for language development. Given the relevance of cognitive theories for our PhD, I will focus on sociocultural perspectives and SCT very briefly, whereas the main part of this section will be devoted to the those cognitive theories that offer the theoretical rationale for the language learning potential of WCF.

SCT is the central theory in the sociocultural tradition. Vygotsky posited that cognitive development and linguistic knowledge can be constructed through social interactions, and hence in a collaborative way. SCT can be used to support the importance of WCF as a socially situated activity. As comprehensively reviewed by Bitchener (2012), SCT supports the assistance that occurs when L2 speakers concur in the same interaction activity, especially if there is a more knowledgeable speaker. This theory was developed on the basis of central tenets of psychology and education (e.g. Engeström, 2001; Rogoff, 1990; Wertsch, 1991), and SLA research (e.g. Swain, 2000; Swain et al., 2011). Regarded as a psychological theory rather than an L2 learning theory (Bitchener & Storch, 2016), the main tenet behind the SCT lies in it triggering the development of higher-order cognitive capacities (such as the ability to plan, intentional memory, among others) through the interactions between novices and expert peers within a society or community (Vygotsky, 1981, as cited in Bitchener & Storch, 2016). Based on the SCT, the individual would develop his cognitive functions as a result of social interactions, which would consequently become internalized in a transformative way (Lantolf & Thorne, 2006, as cited in Bitchener & Storch, 2016). There are three main constructs in SCT (as reviewed in Bitchener & Storch, 2016 and Storch, 2018): (1) The Zone of Proximal Development (ZPD); (2) mediation and tools; and (3) the concept of activity. The first construct refers to the capability of reaching a potential development through problem-solving while being guided by an expert peer or in collaboration with more knowledgeable partners. Related to WCF, this reflects very clearly the relevance of collaborative writing (Bitchener & Storch, 2016). Similarly, a specific view of a process approach to writing (Lee, 2014) represents the tenets of SCT and ZPD, in which the learner is provided with multiple corrections across a continuum of drafts, taking into consideration the evolving needs and abilities. Regarding the construct of mediation and tools, the use of computers (physical tools) and symbolic ones (such as language), have a certain impact on the provision of WCF (Storch, 2018). In terms of activities, Swain et al. (2011) state that humans' actions are socially mediated. In fact, such an aspect is seen in the engagement in activities promoting learning (Lantolf, 2005). In the WCF context, teachers might push learners towards a focus on WCF or the targeted structure, but the individual is held responsible for the importance given to the processing of WCF.

Certain scholars (Lantolf & Thorne, 2007 or Swain et al., 2002) have posited that higher linguistic knowledge might be reached through *scaffolding*, which includes CF. As noted by Bitchener (2012), "this 'other regulation' can eventually enable learners to be 'self-regulated', which leads to the role of collaborative writing activities (Storch & Wigglesworth, 2010; Wigglesworth & Storch, 2012). Thus, the role of proficiency in the SCT becomes essential for the effectiveness of feedback, which requires the appropriate adaptation so that L2 learners achieve their zone of proximal development (ZPD), as "each learner has [a ZPD] in which learning can take place" (Polio, 2012, p. 382). Among the empirical research on the role of SCT in WCF studies (see Polio, 2012, for a review), Aljaafreh & Lantolf's (1994) study indicated that learners automatized the structure after gradual help of the tutor when needed, thus self-regulating it. Storch and Wigglesworth (2010) also detected that, when learners were provided with two different kinds of feedback (reformulation and error codes) to work collaboratively, error coding feedback was more linguistically relevant as learners produced more language-related episodes.

As discussed before, the main connection between the potential language learning affordances of WCF and sociocultural theories falls into *four* main aspects. Firstly, the construct of scaffolding relates to WCF as a form of student-tailored assistance (Storch, 2018), which in itself may promote L2 development (Polio, 2012). Second, if the learner needs less assistance, it means that learning has occurred (see Aljaafreh & Lantolf, 1994; Lantolf & Thorne, 2007), taking it as evidence of L2 development in the field of dynamic assessment (e.g. Poehner & Lantolf, 2013; Lantolf & Poehner, 2014). Third, the diversity of forms the assistance via WCF might take are likely to have an impact on how WCF is provided and processed (Storch, 2018). In other words, the type of WCF and the mediating tool (e.g. a computer) might condition the learners' engagement (Ellis, 2020) with the feedback (see Guardado & Shi, 2007). As mentioned previously, I will focus on cognitive theories supporting WCF in the ensuing sections given their relevance for the present empirical study.

As reviewed by Bitchener and Storch (2016), L2 linguistic competence may be *acquired* or *learned*. Several cognitive theories support the view that this acquired competence (mainly implicit and unconscious) may be reached through the learned competence (namely explicit and conscious). As will be further detailed below, Krashen (1981) claims the idea that acquired competence cannot be furthered by any type of formal instruction (included feedback), but he saw a role for learning to expand the acquired competence (Bitchener, 2012). Conversely, Skill Acquisition Theory (SAT) (Anderson, 1993; McLaughlin, 1990) along with DeKeyser's (2007) studies posit that declarative knowledge may become procedural knowledge through extensive systematic practice. In other words, controlled practice (including corrective feedback) might be facilitative of the conversion of declarative into automatized procedural knowledge (Bitchener & Ferris, 2012). The concept of McLaughlin's (1987) controlled practice relates to WCF in that the comparison of this feedback with the writer's output may contribute to the restructuring of the learners' interlanguage. Finally, other cognitive and interactionist theories, namely the Noticing Hypothesis, the Output Hypothesis, and the Interactional Hypothesis, draw on the ideas that while receiving WCF, numerous cognitive operations are carried out: (i) the learner receives feedback which is input, and engage in noticing the features highlighted in the WCF, relating to the degree of attention (Schmidt, 1994), and (ii) the learner subsequently produce a rewritten version of the text by *applying* the feedback. Drawing on these concepts of the learner's processing capacity and the degree of attention to the highlighted linguistic forms in the WCF are decisive for the amount of L2 input processing that becomes intake (Bitchener & Ferris, 2012). Finally, interactionist theorists consider individual cognitive factors (e.g. the working memory or aptitude) to have an influence over learning, and subsequently, over the processing of feedback.

2.3.1. Krashen's Theory

Krashen's (1985) Monitor Model became relevant as one of the nativist theories of language acquisition, whose influence was and has been perceived until the present time in L2 teaching and acquisition research. As noted by Bitchener (2012), "each of the five hypotheses of Krashen's Monitor Model have something to say, either directly or indirectly, about the language learning potential of written CF" (p. 349). Krashen's (1982) five main hypotheses provide the basis for the Natural Approach. The first of them is the so-called Acquisition-Learning Hypothesis, whose purpose is elucidating the distinction between two ways for the development of L2 competence: *acquisition* and *learning*. Krashen (1982) understands acquisition as an implicitly-led process, in which learners interact in natural communication than in instruction-based settings. For Krashen (1982), learning is a process in which learners engage in a continuum of comprehending the L2 rules (grammar) as a result of instruction and activities with a focus on language forms, undoubtedly related to WCF (Bitchener & Ferris, 2012). Krashen (1982) regarded implicit learning as the natural way of acquiring a language as the subconscious system is involved in any utterances in the L2. His perspective on explicit learning was rooted in the idea that it entailed conscious processes.

The second hypothesis proposed by Krashen is the Monitor Hypothesis, whose purpose is to explain the role of learning in acquisition. Krashen (1982) regarded acquisition as the responsible process in charge of giving rise to utterances or fluency whilst learning aims to monitor this *acquired* process. Along this line, the monitor itself is seen as a device that alters the output produced – included within this acquired system – either *a priori* or *posteriori* (McLaughlin, 1987). Nevertheless, this monitoring process does affect as long as L2 learners are provided with sufficient time to concentrate on accuracy (Bitchener & Ferris, 2012), adding to this the necessity that the specific grammar aspect has been dealt with previously and its relevance is considerable to the learner. Thus, this points to a certain role of WCF when learners are sufficiently knowledgeable about the linguistic structure (Bitchener, 2012). In relation to this hypothesis, the issue of the interface contributes to the debate of the interaction between learning and acquisition (and thus implicit and explicit learning). There are three positions on the issue: the non-interface, the weak interface, and the strong interface. Krashen (1982, 1985) has strongly advocated the first one, as he believes there is no possible interaction between learning and acquisition since they follow two distinct paths. In this sense, Krashen disregarded the role of formal instruction in terms of acquisition, and thus rejected the fact that explicit knowledge could lead to implicit knowledge.

The Natural Order Hypothesis is the third proposal by Krashen (1982), which he based upon research studies carried out by scholars such as Brown (1973), Dulay and Burt (1974), Fathman (1975) or Makino (1980), where they found that acquisition may have an orderly component which made some structures be acquired earlier than others, irrespective of any other student or context-related factor (e.g. age, L1 background, L2 exposure or previous experience, among others). In the same vein, Krashen (1985) states that the acquisition order does not attend to reasons of formality or to the way in which rules are arranged to be taught in the L2 lessons. Thus, Krashen (1982) regards the order of acquisition as proceeding "in a predictable way" (p. 12)– especially grammatical sequencing – provided that the main aim of the L2 learning program is to boost acquisition and not learning. In terms of WCF, this third hypothesis would not acknowledge a role for corrective feedback as the provision of WCF introduces a change in the purported predictable order of acquisition, which the theory does not acknowledge.

Fourth, the Input Hypothesis is considered as constituting the main core of his model (Krashen, 1985). It derives directly from the previous hypothesis, and the main argument for it is that learners are prone to develop their L2 competence as they receive comprehensible input. To phrase it in another way, Krashen (1985) regards comprehensible input (known as *i+1*) as “input that contains structures at our next stage – structures that are a bit beyond our current level of competence” (p. 2). In Krashen’s (1985) words, “input is the essential environment ingredient” (p. 2), and thus a necessary condition to progress up the acquisition ladder by including understandable structures that push L2 learners’ level a little further. In this regard, the focus of this hypothesis was primarily oriented to the comprehension of the message with the help of context. Hence, the acquisition of new grammar structures would be ensured thanks to input that is both comprehensible and abundant.

Fifth and final, the Affective Filter Hypothesis is the last of Krashen’s model hypotheses. Stemming from an original proposal by Dulay and Burt (1977), Krashen (1982) brings up the idea that affective factors may exert an influence over L2 acquisition through the recognition of several affective factors: anxiety, stress, self-confidence, and motivation. Along these lines, input will be filtered according to the learners' perspectives, and consequently, learners having a strong or high filter would be more predisposed to seize incoming input, such as WCF. Learners with a higher filter are more likely to encounter problems internalizing their input in any form. Hence, the affective filter must be low for better processing of WCF.

Krashen's Hypotheses are far from being interpreted as adopting roles to corrective feedback. The first of the hypotheses does pose a problem in the concepts of 'acquisition' and 'learning' as understood by Krashen (1982). In this regard, 'acquisition' is concerned with the use of implicit knowledge in the L2 with the main aim of attaining communication within an L2 context. Conversely, 'learning' is linked to formal L2 instruction and as a consequence, to explicit knowledge. Concerning feedback, the Acquisition-Learning Hypothesis partially neglects the role of error correction as having little or no effect on the acquisition, but Krashen (1982, as cited in Bitchener, 2012) admits that it may be of some use for conscious learning, and thus explicit knowledge, which he equates with 'learning'. Along the same line, the Monitor Hypothesis does not completely rule out the role of WCF providing that the linguistic error under scrutiny is part of the explicit knowledge. Studies carried out in an attempt to verify empirically such aspects have not been successful in clarifying whether a rule is applied consciously or subconsciously. The Natural Order Hypothesis, when addressed within the framework of WCF, does not diverge from the previously explained Krashen's hypotheses insofar as it casts doubt upon the usefulness of both classroom instruction and error correction. Equally, the Input Hypothesis would also disregard the effectiveness of WCF and its potential value as comprehensible input is viewed as enough for L2 acquisition (Bitchener, 2012). Last, the Affective Filter Hypothesis is probably one of Krashen's hypotheses which considers the effect of WCF as negative on the learners' part in that they will not only show signs of avoidance at the time of making mistakes but "difficult constructions, focus less on meaning and more on form" (Krashen, 1982, p. 75). This will undoubtedly produce an overuse of the Monitor on the grounds that L2 learners are paying special attention to their output whilst resorting to their conscious declarative knowledge but at the expense of losing fluency in the L2.

2.3.2. Skill Acquisition Theory

Skill Acquisition Theory (henceforth, SAT) provides a solid basis for the development of skills from an initial stage to mastery (DeKeyser, 2007, 2014). It is a general theory from cognitive psychology applicable to all complex skills, further applied to the SLA context (see Polio, 2012) In this vein, SAT allows for the understanding of learning any skill from a sequential perspective, that is, the gradual acquisition of behavioral skills that are thought to be ultimately automatized. In light of this, SAT brings to light "two interrelated representational systems comprising declarative knowledge and procedural knowledge" (Lyster & Sato, 2013, p. 71). Thus, advocators of SAT do understand L2 learning as a transition to a more automatized use of the language itself together with the corresponding amount of meaningful practice and feedback in all forms. In what follows, we provide an account of the two models of SAT, and what role their main tenets have regarding the LLP of WCF.

►Anderson's Model

One of the main tenets of SAT has been the dichotomous but inclusive view of declarative and procedural knowledge. Such a conceptualization of the phases in learning a language was depicted and developed by Anderson (1983, 1985, 1993) in his Adaptive Control of Thought (ACT) which in further revisions (Anderson, 1993) added an *R* for Rational. Anderson (1993) proposes two types of knowledge which are declarative knowledge (as referred to as the knowledge of *that*) and procedural knowledge (as referred to knowledge of *how*), into which the former may be transformed. As mentioned previously, Anderson’s model “is a general model of skill acquisition” (Bitchener & Ferris, 2012, p. 13), which implies whatever activity which may be learned and further automatized (e.g. when one learns how to ride a bike, the first steps are always very explicit and rule-based, but once one gets accustomed to riding it after considerable practice, the process becomes automatized). This explains why Anderson's model, while not strictly related to the field of SLA, may apply to several aspects concerning L2 learning just as several researchers point out (see Johnson, 1996; O'Malley & Chamot, 1990; Schmidt, 1992; Towell & Hawkins, 1994; as cited in Bitchener & Ferris, 2012: 13). Getting back to ACT, this model sets the stage for a progression line leading declarative knowledge to procedural knowledge through a three-stage path: declarative, procedural, and automatic. the progression between one stage and the other would only happen through practice (DeKeyser, 1997). First, the declarative stage underscores the theoretical and explicit description of the procedure to be automatized. Second, the procedural stage involves turning declarative knowledge into procedural knowledge. This proceduralisation stage occurs at a relatively fast pace. Finally, the autonomous stage takes the learner a step forward by granting more rapid and automatic access to the procedure while reducing the time to access it insofar as practice is intensified. Nonetheless, as precisely indicated by DeKeyser (2015), empirical research has shown that the automatization of the procedural stage results in a much slower process. Such a stage implies "ranging from a mere speed-up of the same basic mechanisms to a speed-up of a broader task through a qualitative change in its components" (DeKeyser, 2015, p. 96).

While some scholars have contended against declarative becoming procedural knowledge, and hence have questioned the role of instruction and corrective feedback (see Bitchener & Ferris, 2012) in relation to SAT, there have been other voices (e.g. DeKeyser, 2001, 2007; Hulstijn, 1995; Schmidt, 1995; Swain, 1985; Swain & Lapkin, 1995) who have adhered to Anderson's model, and the purported contribution of controlled activities in the context of SLA. Thus, controlled activities include the use of corrective feedback as a way of proceduralizing declarative knowledge. The potential development of declarative knowledge and potentia explicit knowledge through some forms of WCF may favor focusing on a specific linguistic problem. As a result, learners may not "proceduralize inaccurate language" (Polio, 2012, p. 381), and hence they may modify their output. The facilitative mechanism for the declarative-to-procedural knowledge path has been *practice*, and hence traditionally associated with the idea of task or exercise repetition on a focused specific aspect of the L2 (Bitchener & Storch, 2016), echoing DeKeyser's (2003) claim that the gap between declarative and procedural knowledge may be bridged with practice. In fact, DeKeyser (2007) saw a role for WCF in helping learners not to produce inaccurate L2. Building on the assumption that practice should be systematic, which in Leow's (2020) view is not part of the writing curriculum, and that WCF is fundamentally explicit, the provision of corrective feedback might help consolidate and proceduralize L2 explicit knowledge, declarative in nature, furthering L2 development (Ellis, 2010; Manchón, 2010). The systematicity of practice requires the accumulation of explicit knowledge "at its various smaller stages and the speed with which it can be applied" (Bitchener, 2012, p. 350), echoing other voices (Polio, 2012) that large amounts of practice are necessary in order to proceduralize this explicit knowledge. The various stages of Anderson's model may account for this aspect since feedback may contribute to expanding declarative knowledge in the initial stage (see Evans et al., 2011 or Harthshorn et al., 2010). Since the procedural and automatic stages may promote the learners' attention to their own declarative knowledge, the presence of feedback may be instrumental in reassessing the potential scope of a targeted linguistic rule that may appear in the feedback provided (e.g. Leeman, 2007). Nonetheless, the role of WCF in Anderson's model is still inconclusive, and as Leow (2020) noted, "[it] is not well explicated nor is it well explained how L2 writers process WCF [in this theory]" (p. 103). Other scholars (e.g. Doughty & Williams, 1998), as reviewed by Van Beuningen (2010), accommodate the view that as implicit and explicit knowledge are separated (e.g. Krashen, 1985), then explicit knowledge may only help in advancing the learner's interlanguage through noticing the gap (as will be further detailed).

►McLaughlin's Model

McLaughlin (1987, 1990) developed his Information Processing Model on the basis that complex behavior - understood as automatized knowledge if considering Anderson's model - is enlarged by simple processes, echoing his claims that there was an interface position between declarative and procedural knowledge (McLaughlin, 1980, 1987). Following this idea, L2 learning may be looked through this lens since *learning* a language is certainly an activity that requires complex thought processes involving high-demanding cognitive skills. McLaughlin's model stems from tenets within cognitive psychology, which emerged from several studies (see McLaughlin, 1978, 1980, 1987, 1990; McLauglin & Heredia, 1996, as cited in Bitchener & Ferris, 2012). This cognitive perspective, as comprehensibly reviewed by Pienneman (2008), resonates with McLaughlin's main views: (i) learners are seen as limited-capacity processors as regards controlled processes, (ii) L2 processing skills are enhanced and more efficient through automatization, allowing for automatic processing without the constraints of controlled-processing.

As anticipated, the Information Processing Model takes the view that information processing entails two central manners of information processing: controlled or automatic. Running in parallel with Anderson's claims, McLaughlin's model (1987, p. 133) does understand "[...] two notions - automatization and restructuring- [...] central to cognitive theory". In defining terms, automatization implies the reduction of some cognitive components, hence impacting the reaction time, and a reduced variability (for a comprehensive review, see DeKeyser & Criado, 2013). As regards restructuring, it implies the reorganization of declarative knowledge so that proceduralization is facilitated (McLaughlin, 1990). Nevertheless, far from being one exclusive to the other, both are interrelated as controlled processing is thought to shift towards automatic processing - in line with Anderson's (1993) views that practice may lead to proceduralized automatic knowledge. This implies that, through instruction and corrective feedback, explicit knowledge may become implicit knowledge. In fact, WCF may have a role in the controlled processing component since, as reviewed by Bitchener & Ferris (2012), it may become automatic processing through the reorganization of declarative knowledge. Such restructuring may be achieved through controlled activities, as posited by other scholars (e.g. DeKeyser, 1997, 2007; Schmidt, 1990, 1995), since this knowledge may lead to automatized knowledge as a result of this controlled phase.

McLaughlin's model relies heavily upon the limitations imposed by the different types of memories. In this regard, short-term memory (henceforth, STM) and working memory (WM), as

Baddeley (1986) initially conceived them, allow for the storing of the most immediate information upon which our attentional focus is placed. This type of information processing on the part of the learner requires a great deal of attention and thus limits the extent to store it (Bitchener & Ferris, 2012). Bearing in mind the effort L2 learners have to expend to preserve such information, it seems plausible that after sufficient repetition this immediate information may become automatic. Thus, such automatized information would be stored in long-term memory (LTM) and therefore easily accessible with minimal attentional control by the learner.

In relation to the above, the role of both the STM and LTM is conceivable during the processing of WCF, since the LTM may also store declarative knowledge. Thus, when learners compare old with new structures, they access the declarative knowledge stored in the LTM, once the linguistic gap is noticed (Schmidt, 1990, 1995, 2001). Under this premise, the role of WCF might favor L2 acquisition given the assumption that declarative knowledge may become automatized through the mentioned comparison. Provided that LTM works in parallel with STM, the restructuring of the learner's linguistic system as a result of controlled processing might boost L2 development (Bitchener & Storch, 2016).

2.3.3. The Noticing Hypothesis

Noticing is a cognitive operation that has proven very influential for different theories in the field of SLA as well as to WCF, and it is thought to be a very relevant psychological process (Izumi, 2014).

The Noticing Hypothesis (Schmidt, 1990, 1994, 1995, 2001) points to the early stages of the L2 learning process, which Leow (2018) refers to as "the input-to-intake stage" (p. 1). Globally considered, this hypothesis sets the stage for the role of input and the minimal attention that learners have to pay to linguistic features, which according to Schmidt (2001), is always conscious, and subsequently, notice these features. Noticing is the first step towards acquisition but does not guarantee it. Schmidt (1990, 1994, 1995) underpins its usefulness by presenting two different forms of this hypothesis: a strong and a weak form. In the former, learning cannot occur if noticing is not activated. The weak form, however, points out the usefulness of noticing but does not take it as inherently essential to the learning process. The weak version entails that L2 learners do not have to "understand" input at a particular level but globally - that is to say, metalinguistic awareness of linguistic forms is not requisite. In later works (Schmidt, 2001), he distanced himself from the strong interface claims and admitted the possibility that learning may occur without conscious perception.

From a global perspective, the Noticing Hypothesis underscores the fact that L2 learners do have a limited capacity to process information. As a consequence, they are not able to pay full attention to every detail of the input to which they are exposed, a context in which attention holds a pivotal role in giving access to awareness and further noticing (Leow, 2018). Thus, Schmidt (1990, 1994, 1995, 2001) does not conceive L2 learning without attention and awareness of the linguistic features, hence highlighting the role of explicit learning in SLA. He distinguished a three-level system of awareness (Schmidt, 1990). The first level is *awareness at the level of perception*, of which the learner may be conscious or not. Although some unconscious or conscious perception may appear at this stage in certain salient forms in the input, the learner does not go beyond apperceiving. The second level entails *awareness at the level of noticing*, which entails consciously noticing “the surface structures of utterances in the input” (Schmidt, 2001, p. 5), i.e. focal awareness. As noted by Leow (2020), noticing “results in intake and in item learning” (p. 100). The third level involves *awareness at the level of understanding*, which is conscious and includes comprehending linguistic rules. This level is thought to lead to a restructuring of these rules by comparing them with the input received. Concerning this, Schmidt (1990) conceives understanding as requiring a much deeper insight into the comparison between the L2, and also analyzing, comparing, and testing hypotheses as regards the linguistic input. Criticism has been levelled at this conception of understanding since the recognition of a specific aspect of the L2 with no previous knowledge of the form makes little sense as a certain understanding of the form is required (Truscott & Sharwood Smith, 2011). Also, Tomlin & Villa (1994) contended against the role of consciousness and understanding as necessary for learning to occur. In this sense, they claimed that certain information may be stored in our memory without conscious awareness, hence not positing any role for awareness in the initial stages (i.e. intake) (Leow, 2015). Notwithstanding the difficulty in the conceptualization of both understanding and awareness, Schmidt (2001, 2012, as cited in Izumi, 2014) stresses out that noticing is certainly facilitative for L2 learning.

Noticing has been tackled from very different angles, which has motivated the scientific discussion about the issue of considering several types of noticing (Izumi, 2013). First, *noticing a form(-meaning-function) relationship* constitutes the basic and most original postulation of noticing proposed by Schmidt & Frota (1986). Such a conceptualization of noticing underscores that learners' linguistic competence is fostered after noticing a particular L2 structure (i.e. form) used in the input. Not limited to a concentration on form, such noticing also entails the meaning conveyed and the context (Schmidt, 2001). Remarkably, this type of noticing does entail a noticing of the form, the meaning, and function itself, from which learners establish relationships (Izumi, 2014). This threefold association seems appropriate providing that L2 learners focus their attention upon each of these aspects. If tackled independently, learning may be partially achieved, giving rise to the wrong use of the L2 for communicative purposes. Eventually, WCF could hypothetically provide learners with the necessary tools to consolidate their understanding of this association by filling in the absence of one of the components of the three-sided relationship, or by prompting new hypotheses regarding this form-meaning-function mapping.

Second, *noticing the gap between the interlanguage and the L2* holds a very relevant role in L2 learning as learners play a leading part in detecting (and noticing) the *gap* existing between their real output, the rules as well as the pragmatic use of the L2. Within this context, Schmidt & Frota (1986) put forward the so-called *noticing the gap*, which takes into consideration the idea that there is a gap to which attention is required in such a way that input turns into intake (Schmidt, 1990). A similar concept is conceived by Ellis (1995) with his notion of cognitive comparison, which underscores the idea that learners do also have to identify similarities between their output and the input received. Along the same line, L2 learners profitably compare language forms produced with the output of a more proficient learner or even a native speaker to enhance their linguistic competence. According to him, such understanding acts as a facilitator for the intake to occur. Notwithstanding the alluded differences between noticing and understanding, Schmidt (1993) makes it clear that both processes store linguistic material in WM and LTM. In this vein, the provision of different types of WCF such as models (Hanaoka, 2007) might foster the creation of ideas, but also the improvement of linguistic forms. How this input is processed may result in the learner being able to correct one linguistic feature on its whole or merely reducing this noticing to a partial feature (Hanaoka, 2007). All in all and according to the above, the provision of WCF would be effective since learners notice the gaps in their output by comparing it with the corrective feedback provided, thus leading them to a subsequent restructuring of their developing grammar. At this point, Leow (2020) argues that, if the L2 learner notices a mismatch between his output and the linguistic information in the WCF, there should be more cognitive engagement beyond the mere noticing of this information without any further elaboration or processing. In fact, noticing has failed to "result in subsequent performance" (Leow, 2020, p. 101) if the input is not transformed into intake and further processed (Leow, 2015, as cited in Leow, 2020).

Third, *noticing holes in the interlanguage* implies a completely different perspective as to the previous type of noticing. In this regard, learners can distinguish a missing linguistic structure (either referring to vocabulary or grammar) in their interlanguage (Swain, 1998). By way of comparison, whilst noticing the gap entails that the learner identifies the presence of a form in the input, noticing holes prompts the learner to screen for the form which is absent in their interlanguage (Izumi, 2014). As seen, this type of noticing focuses much more on the relevance of output as L2 learners have to *notice* a hole to produce a form based upon this hole, thus triggering noticing in the input to look for this form. On this basis, Swain (1998) also puts forward that a previous step to noticing a form may be noticing a hole, a claim which has been tested empirically by a range of studies (see Hanaoka, 2007, 2012; Izumi, 2002; Izumi, Bigelow, Fujiwara & Fearnow, 1999). Minimally noticing feedback is necessary before restructuring the interlanguage, and as reviewed by Leow (2020), if an L2 writer notices a mismatch between his output and a different language content in the WCF, stored in the STM, thus this restructuring is more likely to happen. Nonetheless, noticing with a low level of awareness does not guarantee that deeper processing occurs (Leow, 2020, p. 101).

Finally, *noticing the gap in our own L2 ability* is the last of the proposed types of noticing, which introduces a new conception of noticing in close connection with *noticing the gap* and *noticing the hole*. The gap here is different (Izumi, 2014) as the focus, in this case, is placed on the own learners' awareness of not being able to express themselves by using the exact terms as a result of their incomplete knowledge of the word. In other words, interest is not merely directed towards the form itself, but rather towards using the appropriate expression in the L2 in a correct or precise way in accordance with the context. Thus, learners center their concerns on the intentionality of the message, hence the comparison established between the intended message and the output produced to detect the gap bridging both messages. Izumi (2013) also clarifies that this type of noticing "occurs learner-internally" (p. 28) and despite the indistinctive use of *gap* and *hole* by certain researchers (see Doughty & Williams, 1998), there is a subtle difference between one and the other. Whereas *hole* pertains to the absence of a form in the learner's interlanguage, *gap* refers to the inability to use the L2 to express ideas in a precise manner, where noticing is internal and gives rise to subsequent issues at receptive and productive levels.

In summary, one necessary condition for noticing, and hence WCF, is the prerequisite of attention for learning to be effective (Schmidt, 2001), a claim echoed by other voices (e.g. Ellis, 1995; Swain, 1985; 1995). Noticing processes undoubtedly point to mechanisms in an attempt to bridge the gap existing in the learners' interlanguage, where the LLP of WCF is likely to be pivotal. Thus, when learners are provided with WCF, and some attention is paid to the linguistic forms or the ideas present in the feedback, they might notice the mismatch between the input in the form of WCF and the written product, hence related to the previously mentioned process of *noticing-the-gap*. Schmidt (1995) conceived that conscious attention was necessary, and hence, learners did not only have to merely pay attention to errors (noticing-with-awareness) (Schmidt, 2001). Learners should engage in noticing-with-understanding in order to process input in the form of WCF deeper. This is possible given the off-line nature of writing, which, as announced before, facilitates that learners access their working memory and LTM, for instance. their declarative knowledge (DeKeyser, 2013), and hence the availability of time for the processing of input stimulate learners' engagement in cognitive comparison (Manchón, 2013), that is, identifying new knowledge in the input. This new knowledge, i.e. input in the form of WCF allows for a restructuring of the learners' interlanguage.

2.3.4. Output Hypothesis

WCF as a tool to ease the acquisition of the L2 is also supported by the commonly known Output Hypothesis (Swain, 1985, 1995, 1998, 2005). A large number of researchers have stressed the importance of output as a valuable tool and a push-forward instrument to boost the acquisition of the L2 (see Adams, 2003; Ellis, 2005; Manchón, 2011; Skehan, 1998; Swain & Lapkin, 1982; Williams, 2012). Swain (1985, 1995) argued that output is positively valuable in the sense that it pushes learners to a deeper language processing which requires more mental effort for receptive skills (input). When learners engage in producing new forms in their output, they require conscious attention which is gradually achieved (Schmidt, 1992). Thus, Swain (1985) refers to this "pushed output" as requiring not only semantic processing but also a syntactically challenging form that goes beyond mere paraphrasing.

Though comprehensible input is also invaluable to L2 acquisition, Swain (1985) clarifies that it does not provide sufficient developmental readiness to increase L2 proficiency, and therefore L2 learners need to be pushed to deliver sufficient amounts of *comprehensible output*. Bearing this in mind, it is posited that comprehensible output "may be the trigger that forces the learner to pay attention to the means of expression needed" (Swain, 1985, p. 249) so that when difficulties arise, they are pushed into perfecting their output in such a way to make it more precise and coherent. Throughout time and subsequent revisions, Swain's hypothesis (1985, 1993, 1995) has experienced extensions, and its scope has been enlarged, leading to three specific functions of output positively benefitting SLA (Izumi et al., 1999), all of which will be next looked into greater detail.

1) T*he hypothesis-testing function of output* allows learners to move from a semantically-based analysis to a syntactic one (Swain, 1985, p. 252). Through this, L2 learners can test how comprehensible and accurate their interlanguage is by comparing it with the feedback provided in the L2. As well, this way of regarding output is in close relation to comprehensible input since learners are somewhat forced to get involved in the process of negotiation of meaning, which gives rise to more accurate output (Izumi et al., 1999).
2) *The metalinguistic reflective process of output*, allowing learners to "control and internalize linguistic knowledge" (Swain, 1997, p. 119). By producing output, L2 learners engage in a more complex syntactic processing if compared with comprehension of input. Objectively, such a focus upon syntax is thought to lead learners to an ensuing modification or reprocessing of their output, possibly driving the learning process towards acquisition. Of the utmost importance is Swain's (1995) indication that this metalinguistic reflection does not necessarily entail using linguistic terminology. In this vein, the main objective lies in furthering learners' awareness and the understanding of the "forms, rules, and form-function relationships" (Izumi et al. 1999, p. 423) as long as communication is ensured by the context in which learners are making use of the L2.
3) *Output as a facilitator to noticing holes or gaps in L2 knowledge*. In this sense, Swain (1995) claims that output might potentiate a noticing or consciousness-raising function in that "the activity of producing the target language may prompt L2 learners to consciously recognize some of their linguistic problems" (p. 125) which relates output to the main tenets of the Noticing Hypothesis (Schmidt & Frota, 1986). All in all, output plays a pivotal role here in that it is a facilitator in the process of noticing. Nonetheless, it seems convenient to clarify that Swain does not put output in a privileged position as the only source of L2 acquisition (Izumi et al., 1999).

In relation to the benefits of output towards WCF, Swain (1993) stated very clearly that the absence of feedback may alter the acquisition process and the value of output may be called into question. After such a claim, Swain is pointing to the importance of WCF in this case. Other voices (e.g. Williams, 2012) have equally indicated that output is regarded as having a greater weight within the written modality at the expense of the oral modality. This is a consequence of the spontaneity of oral output, which acts as a constraint, while the written setting is characterized by its permanence.

Accordingly, learners engaging in writing have the possibility of paying close attention to those aspects of the L2 which are particularly complex, hence suggesting that learners compare their at-the-moment output with their previous knowledge. Such a comparison requires that L2 learners resort to the retrieval of declarative knowledge, which points out to the relevance of the tenets discussed and supported by the Noticing Hypothesis (Schmidtt, 1990, 1994, 2001), and as explained previously, the Output Hypothesis (Swain, 1985, 1990, 1998, 2005). In the case of the Output Hypothesis, it is suggested that L2 learners may reorganize their L2 knowledge at the time of writing by testing their hypotheses on linguistic items and ultimately resorting to metalinguistic reflection, hence related to the aforementioned hypothesis-testing function. Thus, when learners are provided with WCF, they question themselves whether their written output is adequate to what is the norm in the L2. As a consequence of this, the L2 structures provided in the WCF, and possibly stored in the working memory, might challenge the learner's interlanguage while producing newly written output, attempting to incorporate the WCF (Polio, 2012). Other voices, such as Leow's (2020), have claimed that considering that feedback is processed consciously, this hypothesis-testing function might facilitate the modification or reprocessing of learners' output. Izumi (2013) claims that, by engaging in writing, learners produce output in whose process they will certainly attend to unacquired forms (noticing), thus identifying holes within their own interlanguage while at the same time detecting those knowledge gaps to be completed.

In contrast to the hypothesis-testing, output as a facilitator in noticing the hole (Swain, 1998) might indicate that, after being provided with WCF, L2 learners may encounter other linguistic problems during the rewriting of their texts. In relation to this, Swain and Lapkin (1995) understand that noticing processes do contribute to producing modified output, suggesting further evidence of integration of new L2 knowledge or even consolidation of previous knowledge. This is congruent with the view that, when learners pay attention to the incoming input in WCF (Schmidt, 1990; Ellis, 1993), it might become intake, and subsequently internalized in the learner's interlanguage. Thus, the modified output might incorporate the linguistic structures identified in the noticing-the-gap process after the appropriation of WCF.

All these aspects considered, the role of WCF is potentially beneficial for L2 learners, and more specifically for children – who are the targeted population in this PhD – since producing written output and being provided with CF is undoubtedly facilitative of L2 development. The provision of WCF in its many forms (e.g. models) might spark learners' interest in producing new forms and structures, encouraging writers to test their hypotheses. Nevertheless, Leow (2020) claims that for the Output Hypothesis, and more specifically, the hypothesis-testing function to take place, L2 learners' motivation to engage in deep feedback processing should be high. More research is needed on how feedback is processed as well as on the impact of L2 learning on the written product (Leow, 2020), given that the most common timing of WCF, asynchronous, may impair the learners' ability to test their hypothesis while rewriting their texts.

2.3.5. The Interactional Hypothesis

First proposed by Long (1981), the Interactional Hypothesis initially posited that engaging in conversation with native speakers may lead to L2 acquisition as a consequence of this interaction. The Interactional Hypothesis has been revisited by several scholars throughout the last decades, and has received several names: Block (2003) refers to it as *input, interaction and output model*, Carroll (1999) called it the *interaction theory*, Ellis (1991) framed it within the oral dimension by naming it *the oral interaction hypothesis*, and eventually, Gass & Mackey's (2007) most recent renaming has been *the interaction approach*. The main theoretical underpinnings behind this theory are rooted in a particular emphasis on the roles of input, output, and feedback, which "occur during interaction in the L2" (Polio, 2012, p. 383). These concepts, which are drawn from other L2 learning and acquisition theories, point to the access to comprehensible input as a necessary aspect, but in greater amounts, which may lead to acquisition (Long, 1983). Long embraced the tenets of the comprehensible input theory (Krashen, 1982), but his conceptualization entailed a different perspective in which interaction held a more relevant role than mere exposure to input.

On theoretical grounds, the major component in the Interactional Hypothesis is *attention*. When learners engage in writing, they have to pay attention to certain forms to generate their output, which is what the Output Hypothesis (Swain, 1985, 1995, 2005) puts forward, as stated in the previous section. In the event that learners are provided with input, as with the provision of WCF, they are expected to attend to form. When this interaction between input and attention occurs, the act of providing feedback, either implicit or explicit, implies a form of negotiation and recasts (Polio, 2012), especially during oral communication, encouraging the learner to pay attention to form. This is understood as "a form of positive evidence [...] and negative evidence" (Bitchener, 2012, p. 351).

The Interactional Hypothesis has been framed within oral language research, but recent empirical endeavors, as reviewed by Polio (2012), have applied several aspects of this theory to the field of writing (see Swain & Lapkin, 1995; Swain, 1998; Qi & Lapkin, 2001; Sachs & Polio, 2007) and of WCF. Polio (2012) offers a very synthesized perspective of the application of this theory to L2 writing. When learners are exposed to a certain type of WCF, they are presented with a form of input; ideally, they would apply these corrections to their revised piece of text, which is a form of output. Furthermore, directing learners' attention to form might result in learning given the permanent nature of writing, and especially owing to the sometimes explicit nature of WCF. Long (1996) claimed that this explicitness might lead learners to noticing the WCF in addition to the amount of time to revise at their disposal in contrast to oral CF. Finally, Polio (2012) stresses out the relevance of individual cognitive factors as variables mediating learning, holding similar facilitative and inhibitive effects on the process of learning while writing to those in the oral domain. However, Bitchener (2012) adds to this debate that some individual cognitive factors might have a greater impact than others (e.g. working memory when engaging in oral communication).

2.3.6. The role of WCF in Leow's model of ISLA

Leow (2015) theorized his model of L2 learning process, which aimed at bringing together the different phases of L2 development and learning. In his model, input and output are regarded as external products in a learning continuum involving products and processes. Situated in "the instructed setting" (Leow, 2020, p. 103), his model takes the concept of attention as central, the introduction of another SLA construct, *depth of processing*, is pivotal throughout the different processing stages of the L2 learning process.

Learning as a process is internal, and comprises the following stages (Leow, 2015): (i) input processing, which is the initial stage, and encompasses apperceiving, detecting, as well as noticing the L2 information, (ii) intake processing, involving the creation of form-meaning, and hypotheses-testing functions such as modifying, rejecting or confirming them, and (iii) L2 knowledge processing, where the learner analyses and integrates the L2 internal representations, leading to potential learning as a result of the manipulation of the L2 developing system and the subsequent output. In Leow's (2020) words, "the knowledge processing stage is most pertinent to WCF" (p. 103) since the learner may retrieve the L2 information by returning to the early input processing stage as a result of the feedback provided. But most importantly, Leow (2020) alludes to how deep feedback is processed as pivotal for the restructuring of inaccurate knowledge. Nonetheless, the retrieval of this linguistic information is subject to the role of attention, which is crucial, along with the depth of processing and level of awareness. Interestingly, Leow (2020) proposes a feedback processing framework entirely based upon his ISLA model of L2 learning process, where the different stages of these processes are characterized with reference to the provision and processing of feedback. He distinguishes five processes within this framework (Leow, 2020, pp. 104-105): (1) *feedback* on the learners' output production, on which learners have to pay minimal attention, that is, detecting, noticing, and higher depth of processing so that feedback intake becomes part of the working memory, (2) *feedback processing* is a fully cognitive stage which alludes to the manner in which learners process feedback as regards their interlanguage or prior knowledge, (3) *internal system*, which is a stage recursively alternated with the feedback processing stage. If the learner is able to process the feedback either with a low or high degree of depth of processing or awareness, such linguistic information would reinforce prior knowledge or restructure the inaccurate knowledge in the internal as a result of the feedback provided. Following Leow (2020), the new restructured linguistic information would replace or become part of the knowledge present in the internal system. Then, (4) *knowledge processing* makes use of this restructured information which, depending on the degree of attention or awareness, would then become part of the L2 knowledge, and (5) the previous stage may give rise to the co-existence of old, inaccurate and new, modified output. According to Leow (2020), the presence of old output may be the direct cause of the low or absent depth of processing of feedback. Conversely, the new or modified output indicates that the learner has assumed the restructured L2 knowledge, which might be semi-permanent or temporary. In fact, the use of this restructured L2 knowledge points to the full internalization of the feedback and an accurate restructuring. The use of an old or inaccurate structure in the output may be related to a temporary or immediate restructuring, suggesting that learners did not deeply process the linguistic information in the feedback.

Contributions of these theories to WCF and writing as a site for L2 learning

As reviewed above, these theories (Krashen's Theory, Skill Acquisition Theory, The Noticing Hypothesis, The Interactional Hypothesis, and the ISLA model of L2 learning process) have something to say about the relevance of WCF and L2 learning. In summary, the overarching role of attention as a key cognitive operation in the process of writing and the processing of WCF has been central to the Noticing Hypothesis. In fact, when students engage in the revision of their texts with the WCF provided, several cognitive processes may be activated. In this sense, the processes of noticing-the-gap or the attention paid to the noticing process and the output have proved essential for WCF. This role of noticing but also of attention were highlighted by Leow (2015, 2020) as the stronger these cognitive processes are, the more learning will occur. This is particularly relevant in the case of the many forms of WCF, since learners compare their written output with the feedback provided to observe a potential mismatch. Such a mismatch is more likely to happen if feedback is at least minimally noticed (Leow, 2020), but most importantly, attended to. The off-line nature of both writing and WCF allows learners to examine their output and the feedback provided, and to pay more attention to different language aspects. The Output Hypothesis paves the way for L2 development in a similar fashion, since learners are encouraged to produce their written output, testing their hypothesis, but more importantly, by creating interest in learning new structures. Likewise, it is not clear whether the Skill Acquisition Theory is supportive of the usefulness of WCF for learning. In other words, the explicit nature of WCF may help that declarative knowledge becomes procedural knowledge as long as enough opportunities to practice are provided (DeKeyser, 2003). Yet, the role of this theory for WCF is inconclusive (Leow, 2020).

Feedback may help proceduralize L2 explicit knowledge present in WCF through systematic practice (Leow, 2020; Ellis, 2010). Finally, Leow's (2015) model sees a relevant role for WCF to advance L2 learning and acquisition. The opportunities to practice are, in fact, mirrored in the provision of WCF and the certain degree of attention to this WCF. As mentioned previously, this attention may be attended, detected or noticed, at different levels of processing. Leow's (2015, 2020) model allude to the differential effect of WCF depending on how it was processed, and how much it was attended to. Very importantly, his model states that, even if there is low depth of processing, as might be the case in young EFL learners, it is still likely that the new form is incorporated along with old, inaccurate output.

Bearing in mind the above discussion on the theories pointing to the LLP of WCF, it appears that the provision of WCF might result in varied effects depending on how WCF is provided, and the subsequent revision of the text both from a short-term and long-term perspective. Most of these claims have been tested with a focus on the final product, and the feedback processing involved during the uptake. Nevertheless, to our knowledge, no study has concentrated on the potential effects of these theories applied to the learner's strategic behavior during writing, i.e. writing processes, and how the provision of WCF might affect the subsequent writing process in a revised text. Hence the relevance of the present PhD dissertation with its focus on the effects of the learners' use of WCF on the writing processes and pausological behavior.

2.4. Learner-internal and learner-external variables mediating the use of WCF

A wide array of variables that place internal and external factors on focus may have an impact on L2 learners' writing processes. These factors have been traditionally divided into cognitive, affective, and personality-related ones, making a further distinction between cognition, motivation, and emotion (Dörnyei, 2010). The mediation of these factors may be subsequently divided into *internal* and *external* ones. As for the former – within the cognitive, motivational, and affective spectrum – it comprises factors such as L2 learners' working memory, their processing capacity, the language learning aptitude, goals and interest, and attitudes and beliefs. Concerning external factors, these encompass macro and micro contexts, type of corrective feedback, the role of instruction, and the type and modality of tasks. Given the nature of the participants of the present dissertation, i.e. young EFL learners, the most relevant aspects of learner-factors (both internal and external) will be discussed below by focusing mainly on the influence of *age* on each of them. Although this dissertation does not deal with these variables directly, in what follows, different mediating variables will be considered with regards to their purported influence on WCF processing. We consider it relevant to review the role of some of these individual differences in relation to the age factor in the first place (i.e. learner-internal variables), to examine the potential influence that might condition the processing of WCF and, as a result, the effect on writing processes.

2.4.1. Learner-internal factors. Age as an influencing variable

As expressed above, several are the factors mediating the outcome of the writing processes as well as the feedback provided. Following Bitchener & Storch's (2016) classification, cognitive factors include, as mentioned previously, (1) L2 learners' working memory and their processing capacity; (2) L2 learning aptitude. For the present review, only factors related to (1) are reviewed. I will look into the role of age in terms of working memory and processing capacity given the relevance of young EFL learners for this study, as announced before.

L2 learners' working memory holds a central role among the factors having an impact upon learning and feedback processing. Learners' capacity to internalize what has been noticed in either input or output is accountable to their inner ability to process the information and its ensuing storing in their working memory. Conversely, Baddeley (1986) realized that STM was not an umbrella term that could encompass all the components within the temporary store. Thus, he started to use the concept of *working memory* (Baddeley & Hitch, 1974) which expands upon a dynamic view of the theory of memory systems. As understood by Baddeley (2003), working memory is regarded as a combination of storage of information and its subsequent processing and manipulation, which is unquestionably tied together with cognitively demanding activities involved in the process of learning, and thus writing. Nevertheless, working memory is considered as a limited-capacity processor, and thus highly dependent upon the amount of information that can be processed. As such, the Limited Attention Capacity (LAC) model (Skehan, 1998; Skehan & Foster, 2001) showcases that only learners with a larger working memory capacity will be successful in attending to and processing the input received. Following this explanation, this model suggests that lower proficiency learners, such as children, face a great deal of difficulty if their attention is concurrently directed towards manifold aspects, resulting in cognitive overload. This occurs since L2 learners do have to attend to several aspects such as meaning and form, the different types of noticing (holes and gaps), and establish parallelisms with their output and previous knowledge. As the working memory is the responsible device for the coordination of attentional resources and the temporary storage of the information and its processing, the WM capacity - either large or reduced - will be a very relevant predictor of an ensuing success in numerous complex cognitive operations, such as writing or reasoning (Engle et al., 1999). Given these theoretical grounds, children as L2 learners are limited capacity processors for a number of cognitive and development factors (McLaughlin, Rossman & McLeod, 1983). As such, they are bound to experience some cognitive load as a result of the necessity to attend to semantic and contextual clues, hence leading to a reduction of the amount of attention devoted to form (Izumi, 2013). As Kormos et al., (2019) point out, the young learners' writing skills are still developing in their L1, and hence their cognitive ability. The ongoing development of cognitive abilities might indicate that the role of the WM might change across grades, echoing other voices (e.g. Gathercole & Alloway, 2004) stating its increase during middle childhood years.

As a consequence of the above, younger L2 learners are not able to engage in deeper processing of input as this requires greater attention, awareness, and cognitive effort (Leow, 2015) which goes in line with the idea that low-proficient learners should be guided through in a much more consciously controlled manner in addition to the enormous amount of information that shall be looked into and processed (Bitchener & Storch, 2016). Engaging in writing entails numerous operations, either concurrently or not, and thus the coordination of the different aspects concerned within the working memory is necessary for cognitive processing to be effective. In this context, after having produced their written output, L2 learners are provided with WCF, when the working memory is fully active to detect these inconsistencies between the output and CF through the noticing and output mechanisms claimed in the previously mentioned theories (Schmidt, 1995; 2001; Schmidt & Frota, 1986; Swain, 1998, 2005). When the linguistic information is eventually proceduralized, the load of the working memory is alleviated, thus facilitating its access both in terms of time and reactivity (Bitchener & Storch, 2016).

At this age, children are governed by the *primacy of meaning* principle (Van Patten, 1993, 2004) which stresses out the learners' predisposition for processing meaning as derived from input prior to the processing of form. Such a principle suggests that communicative-oriented L2 features are hence more likely to undergo processing through a three-stage system, whose organization is as follows: (a) learners' focus is on lexis rather than on the form itself, (b) prevalence of lexis to processing meaning rather than on analogous grammatical structures, and finally (c) there is a prevalence of semantic encodings over focal attention to forms, in which the former are processed in the first instance. According to Uggen (2012), learners may not succeed in noticing or attending to a particular structure or form if, when processing the L2 input, their focus is primarily on meaning, which might, in turn, exhaust their resources.

When interacting with WCF, young EFL learners usually identify points of grammatical and semantic disparity between their text and the feedback. However, these points of disparity largely depend on the type of WCF provided. In this case, more explicit feedback will very likely suggest changes in terms of formal aspects of language, while more implicit ones, such as indirect WCF or model texts, might encourage learners to focus on lexical aspects and phrases rather than on more specific grammatical aspects such as the verb tense (Coyle & Roca de Larios, 2021).

2.4.2. Learner-external factors mediating WCF

In contrast to learner-internal factors, L2 learners are further conditioned by a broad array of external factors which include the influence of the environment and context. Similarly, learners' L2 development – particularly in L2 writing – depends on learners' motivation, anxiety, attitude, and engagement. These external factors hold a mediating role in furthering L2 development as they conflate to influence L2 learners' cognitive processing. In fact, these variables might make a difference when related to other learner-internal variables such as the ones presented in the previous section. Thus, WCF is a facilitator factor furthering L2 development, but it also conditions the way in which children might engage in WCF depending on implementation variables such as the type of feedback.

► Type of feedback

The type of feedback has a great influence on the learners and their effectiveness at applying it. Bitchener & Storch (2016) indicate two main distinctions in the type of WCF: *focused* or selective, when the correction is targeted at specific linguistic features, and *unfocused* or comprehensive, which addresses every single error or mistake made. This dichotomy between focused and unfocused CF has received scholarly attention as regards effectiveness, reflecting a variety of views on the relevance of focused CF. Learners are thought to respond better to corrections when these are targeted at only one linguistic item (Ellis, Sheen, Murakami & Takashima, 2008), prompting scholarly debates around the relevance of noticing (Schmidt, 1994) in detecting and processing this WCF (see Bitchener, 2008; Sheen, 2007). Conversely, unfocused WCF is thought to overload learners cognitively, thus nullifying the potential effect of WCF.

Along with it, a further distinction can be easily made regarding "the learner's involvement in the correction process" (VanBeuningen, 2010: 11) and the degree of explicitness of WCF (see Chandler, 2003; Ellis, 2009; Ferris, 2006; Suzuki et al., 2019; Robb, Ross & Shortreed, 1986; VanBeuningen, 2010): direct WCF, which entails providing learners with the correct linguistic item in an explicit way, and crossing out or marking the wrong word or set of words, and indirect WCF which means indicating the location of the error without providing the correct answer through underlining the error. Finally, metalinguistic CF includes providing an explanation or a rule and it may entail correcting errors comprehensively or selectively targeting specific linguistic features. The degree of explicitness of WCF has been a matter of contention, and as a concept, feedback explicitness is a key variable for the effectiveness of WCF. Such explicitness refers to the degree of explicit information that feedback might provide learners with, and the manner of correcting these erroneous linguistic structures (Nassaji, 2009, 2015).

As with the focused-unfocused WCF dichotomy, the debate between 'direct-indirect-metalinguistic' WCF revolves around the effectiveness of each feedback type. Scientific literature has persistently advocated for direct and indirect feedback (VanBeuningen, 2010). In this vein, proponents of direct feedback argue that this type of corrective feedback has proved to be more useful given that learners may tap into the correct language items – favoring clarity over confusion – while at the same time fostering learners' noticing of other linguistic aspects such as propositional complexity, syntax or vocabulary choice (Allaw, 2019; Bitchener & Knoch, 2010; Chandler, 2003, Vasylets et al., 2019). By providing the correct language form, learners are pushed forward to internalizing the form with certainty whilst reducing their cognitive load owing to the explicitness of the correction. Furthermore, it is suggested that direct WCF is more effective when addressed to low-proficiency students (Ferris & Roberts, 2001) without neglecting the fact that it is less cognitively demanding as minimal processing is required, thus hindering its storage into the long-term memory (R. Ellis, 2009). Conversely, advocates of indirect WCF claim that learners will find it more advantageous in that they are supposed to engage in deeper thought-processing when analyzing their L2 errors. In other words, L2 learners are involved in problem-solving as reflection is needed so that the ultimate goal – that of correcting L2 errors and giving rise to L2 acquisition – is attained (Bitchener & Knoch, 2008, Ferris, 1995). Although several studies (Amelohina et al,. 2019; Lalande, 1982; Park et al., 2016) have shed light upon the effectiveness of indirect WCF in the short and long term, others have proved the opposite (Ferris & Roberts, 2001) as their findings were relatively equal and no statistically significant differences were found among both types of feedback. Beyond comparing direct and indirect CF in terms of productivity and efficiency, the debate is scrutinized at even more micro-contextual levels in the case of indirect WCF: that is, indirect WCF indicating where errors are located (i.e. explicit indirect WCF) in contrast to indirect WCF with no indication (i.e. implicit indirect WCF). In light of this, studies (see Lee, 1997) have endeavored to elucidate the effectiveness of each type of indirect WCF, and findings revealed that explicit indirect WCF had a more effective impact on learning than implicit indirect WCF.

Metalinguistic feedback constitutes another type of explicit feedback (Suzuki et al., 2019) in which rules and explanations are provided along with examples displaying the use of the correct linguistic form. Advocators of metalinguistic feedback believe that it facilitates L2 learning because of its explicitness (Bitchener, 2012) as well as fostering "guided learning and problem-solving" (Lalande, 1982, p. 143). Metalinguistic corrective feedback may be implemented in a number of ways, all of which put forward a specific strategy on the learners' part (R. Ellis, 2009). To illustrate this, metalinguistic feedback might be used with direct feedback, thus decreasing the degree of noticing on the learners' part (Schmidt, 1995). This type of feedback involves using labels under each error to specify the type. In other words, it would be a type of indirect feedback with metalinguistic information. Furthermore, metalinguistic explanations could be provided in the margins with several variations of the targeted or erroneous linguistic structures.

At a higher level in this debate, the long searched question has traditionally revolved around the long-term effectiveness of written CF, and scientific attention has been placed upon it to search for the most beneficial way to carry out WCF (VanBeuningen, 2010; R. Ellis, 2009). Few studies (e.g. Fathman & Whalley, 1990; Polio, et al., 1998; Ashwell, 2000; Nicolás-Conesa, Manchón, Cerezo, 2019) researching on corrective feedback have had recourse to a control group to observe the comparison between the provision or absence of WCF (Bitchener, 2008). Some of the arguments against have been put forward by Ferris (2006), who has stressed out the ethical concerns with the presence of a control group as they are not receivers of WCF at the expense of another group who are. Nonetheless, there is a common agreement among scholars that a true control group should be present in a WCF study (Bitchener, 2008; Ferris, 2006; Truscott, 2004). Although the present PhD does not look into the efficacy of WCF in the final written product, the inclusion of a control group in our study responds to a methodological requirement for the sake of internal validity.

Apart from the debate on the effectiveness of each type of CF, a great deal of attention has shifted towards the investigation of different types of CF as well as towards the control of external variables with certain influence over its effectiveness (Guenette, 2007). Despite the vast amount of research exploring the empirical contribution of direct or indirect WCF for language development, other types of CF have been explored in recent years to account for distinct alternatives in the provision of WCF and their impact on L2 learners' writing and L2 development (see Adams, 2003; Coyle et al., 2018; Hanaoka, 2006; Hanaoka & Izumi, 2012; Roca de Larios & Martínez, 2010; Sachs & Polio, 2007; Qi & Lapkin, 2001). The main focus of this research has centred on reformulations, which "straddles the boundary between direct and indirect CF since it provokes cognitive conflict but it does, in fact, offer a target like alternative just as direct feedback does" (Lázaro, 2013, p. 32), as well as model texts serving as WCF. Unlike traditional corrective feedback techniques, where errors are itemized and indicated, models as WCF include whole pieces of texts which are thought to support L2 learners in noticing and incorporating aspects of these models into a revised version of their text.

On this basis, Levenson (1978, as cited in Qi & Lapkin, 2001) defines *reformulation* as an L2 learner's text which is rewritten in a native-like manner but maintaining the original content and devoiding it of any inaccuracy on whichever level (i.e. rhetorical, ambiguity, coherence, vocabulary choice, grammar awkwardness). Thus, reformulation provides the learner with an opportunity to establish a thorough comparison between his original written text and the rewritten version by the native speaker, fostering the cognitive strategies backed by the Noticing Hypothesis (Schmidt, 1990, 2001) as well as the Output Hypothesis (1985, 1995; Swain & Lapkin, 1995). As expressed by Lázaro (2013), reformulation involves a dual objective since learners can benefit from the advantages of direct feedback and the presence of cognitive conflict (see Doughty & Williams, 1998). Nevertheless, the effectiveness of reformulations in terms of error correction and their influence over L2 development has been questioned by scholars (Allwright, Woodley & Allwright, 1988) on the grounds that reformulations are riddled with the constraints imposed by the original text, thus not allowing for a neat rewriting tailored to the standards of the L2. As a consequence of this constraint, vocabulary and grammar structures may be enhanced but such an improvement may not provide new lexis and morpho-syntactical additions. In line with this, R. Ellis (2009) understands reformulation as an extension of direct feedback correction plus revision, but they are not equal to each other as reformulation *per se* applies to the whole text by pushing learners' to read it and compare it with their original texts so as to detect changes and relevant modifications.

In response to the constraints imposed by reformulations insomuch as they do not allow for flexibility in terms of language development other than what is portrayed in the original text, another more recent type of feedback known as "model texts" has made its appearance in different research works (Coyle, Férez et al., 2020; Coyle, Cánovas et al., 2018; Coyle & Roca de Larios, 2014; Hanaoka, 2007; Martínez Esteban & Roca de Larios, 2010). Model texts are native or native-like texts following "the content and genre of the target text-type" (Coyle et al, 2018, p. 2) which are presented to the student so as to serve as a guide to enhance his own original text. For this thesis, this type of feedback will be thoroughly discussed in the next section.

2.5. The use of models as a WCF technique

As anticipated in the previous section, alternative techniques to traditional WCF forms include reformulations and model texts, both of which are different in their features but share in common the common purpose of providing learners with whole texts or chunks of texts instead of error coding or merely signaling the mistakes on text (Coyle & Cánovas, 2019). The use of model texts is a WCF technique whereby a text is modeled according to the task presented – which reflects a specific genre – and more importantly, adapted to learners' age and their L2 level (Coyle & Roca de Larios, 2014; Hanaoka, 2007; Qi & Lapkin, 2001). The use of models potentially fosters learners' language awareness that may not arise when other WCF techniques are resorted to, thus driving them to "engage in deeper processing" (Coyle & Roca de Larios, 2014).

2.5.1. Models as a type of WCF

Models have been traditionally provided as texts to which the writer may resort before initiating the act of writing. With this feedback technique, learners have the opportunity to observe the conventions of a particular writing genre, which serves as a guide. This is particularly visible in EFL textbooks, in which model texts are usually presented in the writing reference section whereby students are exposed to the conventions of the writing genre, as well as presented with expressions taken out from the very same model in order to serve as a basis for learners. Nevertheless, models as written corrective feedback are purposefully targeted at providing learners with an alternative version of a text following the same instructions of the task. In the early 70s and 80s, models served as a resource used *within* the writing process to find solutions to L2 issues (as firstly stated by Eschholz, 1980 and Paulston, 1972). More recently, models have been reported to furnish learners with the opportunity to match up their original text to the model provided in order to distinguish the most apparent differences on linguistic or ideational terms to enhance their L2. Such a comparison is highly related to the concept of noticing the gap (Schmidt, 1990, 1993) and cognitive comparison (Ellis, 1995) insomuch as learners are pushed to identify similarities and differences between the text produced and the input (model) given. Thus, models set the ideal stage for learners to detect the putative mismatch or gap between their output and the conventions of L2 writing (Schmidt, 2001).

The comparison of output performance with a model text contributes to attending to several L2 structures, which learners may not be familiar with and which could be integrated into their interlanguage (Hyland, 2004), thus allowing for the internalization of these expressions and for L2 development. Similarly, learners may not have fully internalized the writing genre after their being taught about its conventions, thus favoring the creation of a mental model of such writing type (Crinon & Legros, 2002).

Among the many advantages of models, it must be noted that errors or mistakes are not explicitly indicated, which indirectly encourages learners to identify errors, hence pushing them to engage in deeper processing (Adams, 2003; Sachs and Polio, 2007; as cited in Martínez & Roca de Larios, 2010). This is suggestive of the beneficial side of models as feedback since deeper reflection is fostered as a way of furthering L2 learning by incorporating or consolidating linguistic structures.

2.5.2. Previous research on the use of models from a product-oriented perspective

Research studies on the role of models as WCF has experienced considerable growth in recent times not long after Hanaoka's (2006) original study, which explored the function of models as noticing-triggering resources to linguistic problems aroused during the process of composition in individual writing. The research design entailed a three-stage procedure whereby learners, at stage 1, received their prompt (a picture-based story) and an additional sheet where they had to write the problems noticed during the composition stage. Subsequently, at a stage 2, learners received their original text and two different model texts of their picture-based story, which they had to compare and note down the differences detected. Ultimately, Stage 3 comprised rewriting the text with solely the use of the prompt, and Stage 4 was a post-test identical to the Stage 1. In his pioneering study, Hanaoka (2006) found out that adult learners noticed the holes in their interlanguage (Swain, 1998) as a function of the model texts, to solve the problems encountered throughout the feedback processing stage and to eventually incorporate these solutions to the rewritten stage. Findings revealed that learners tended to focus on lexicon and prepositional features to solve problems with the help of models and these partially helped to recover features from their original output. As regards L2 proficiency, less-proficient learners displayed greater difficulty in noticing linguistic features than high-proficient learners, whose attention to form was much more elevated. It was after Hanaoka's (2006) study that numerous attempts were made to delve into the language learning potential of models as a WCF technique. As can be seen in Table 1, this led to an increase in research studies analyzing from different angles the potential displayed by models in a number of ways and with a number of variables contributing to shedding light on L2 writing research agendas (Manchón, 2011). Table 1 presents a summary of the most important aspects of studies dealing with models. All of these studies will be synthesized in the ensuing paragraphs.

Table 1.
Studies investigating the effect of models on L2 writing.

Study	Participants	L2 Proficiency	Focus	Type of writing condition	Type of task	Setting	Main findings
Hanaoka (2006)	37 Japanese university students (17 students from Class A; 20 students from Class B).	Not mentioned.	The usefulness of native-speaker models in L2 writing.	Individual writing.	Two-frame picture prompt.	Pen-and-paper (laboratory study).	Lexically-oriented focus in the comparison task. Conscious recognition of holes, and subsequent noticing of these holes. Avoidance was present, interestingly with the use of the L1.
Abe (2008)	14 Japanese ESL adult students (7 advanced + 7 intermediate).	Approximately IELTS 6-8 (range)	Effectiveness of models as tools for L2 development	Individual writing.	Descriptive and argumentative essay.	Pen-and-paper	Noticing aspects of language. Focusing on lexicon, form and discourse, in this order. Less proficient learners are less likely to notice.
Martínez & Roca de Larios (2010)	17 EFL Spanish secondary students	Low-intermediate English.	Relation of noticing to composing and feedback processing.	Individual & Collaborative	Short story in response to a picture.	Pen-and-paper.	Noticing of lexical features (more individual than pairs) and ideas. Attentional focus on lexis as a result of the lack of knowledge.
Yang & Zhang (2010)	10 EFL Chinese university students (organized into 5 pairs).	Low-intermediate, intermediate and high-intermediate.	The effectiveness of reformulations and model texts in EFL writing.	Collaborative.	Narrative in response to a picture prompt (based on Lapkin et al., 2002).	Pen-and-paper.	Both reformulation and model texts fostered noticing the hole. More focus on the lexical and formal aspects than on organization or ideational content. The model text was not fully analyzed beyond the lexical level.
Coyle & Roca de	23 EFL children students pairs (aged 10-12).	Low, medium and high proficiency levels. Specific level is not mentioned.	The role played by error correction and model as WCF.	Collaborative.	Four-frame picture story prompt.	Pen-and-paper (4-week period classroom experimental research)	Self-initiated noticing, mainly lexically-driven. Noticing of lexis and language chunks. Grammatical revision is less fostered by models.

Larios (2014)							
Kim (2015)	52 EFL learners (aged 10-11 year olds).	Low, medium and high proficiency level. Specific level is not mentioned.	Children L2 learners' proccesing and use of error correction and models as WCF.	Collaborative.	Four-frame picture story.	Pen-and-paper (classroom-based experimental research)	Noticing was self-initiated by lexis. Young learners focused on lexical and sentence-related features. High proficiency pairs reported problems on grammar. Grammatical features were not noticed with models, but effective in noticing ideational content.
García Mayo & Labandibar (2017)	60 Basque-Spanish bilinguals (30 were around 13 years old and 30 were around 16 years old). EFL students.	Elementary and lower-intermediate proficiency level.	Noticing when writing a composition. How noticing and feedback processing affects revisions.	Individual writing.	Visual prompt to write a text.	Pen-and-paper (laboratory study).	Learners noticed gaps when writing and processing feedback. This noticing was mainly lexically-driven. Lack of vocabulary knowledge or lack of knowing the potential solutions were found in the feedback comparison stage. The feedback group noticed and solved two-thirds of the solutions available. Models provided learners with alternative vocabulary, phrases and content.
Coyle et al. (2018)	16 EFL children learners (10-11 years old).	Low level of L2 proficiency.	LLP of WCF processing and the trajectories in writing.	Collaborative	Six-frame picture story prompt.	Pen-and-paper (classroom-based research; 6-week instructional period).	Guided instruction on models impacted positively noticing processes. Procedural knowledge facilitated processing demands of the feedback comparison stage. Less useful trajectories led to more limited opportunities for knowledge development. Occasional attention to syntactic and discursive level: verb forms, gerunds, articles, or personal pronouns.
Coyle & Roca de Larios (2020)	16 EFL children student pairs (10-11 and 9-10-year-olds) in an EFL and CLIL class.	Low level of L2 proficiency.	WCF processing and uptake with model texts.	Collaborative.	Six-frame picture story prompt.	Pen-and-paper (multi-stage task).	Noticing of lexis in both EFL and CLIL pairs. Noticing surface-level differences with the model (especially, CLIL pairs). Multi-word expressions and morphological features were only partially noticed.

Luquin & García Mayo (2021)	38 EFL children learners (aged 11-12).	A2 (Elementary/Pre-intermediate).	What EFL primary school children notice and incorporate.	Collaborative.	Response to a picture prompt.	Pen-and-paper.	Identification of the content and language-related problems, and notice gaps during the output. More attention to form primarily, and subsequently to lexis and mechanics. The feedback group noticed more items than the control group, and lexical aspects were noticed in the feedback comparison stage in the feedback group. Models fostered the incorporation of discursive, mechanical, and formal features, mainly spelling and stylistics.
Kang (2020)	40 EFL South Korean secondary students (aged 17-18).	B1 (intermediate).	The effect of models for the improvement of L2 writers' output.	Individual	Argumentative essay.	Pen-and-paper.	Lexical problems were noticed at the moment of writing, and more noticing of grammar-related problems. The feedback group noticed more lexical aspects from the models, and participants were concerned about content-related issues. Models served as a trigger to obtain more ideational content.

The vast majority of empirical research on models has been concerned with assessing its effectiveness with a series of differential modalities of writing (Hanaoka, 2006; Abe, 2008; Martínez & Roca de Larios, 2010; Abbuhl, 2011; García Mayo & Labandibar, 2017; Martínez & Roca de Larios, 2010; Yang & Zhang, 2010; Coyle & Roca de Larios, 2014; Cánovas, Coyle & Roca de Larios, 2015; Kim, 2015; Coyle, Cánovas & Roca de Larios, 2018): individual or collaborative writing, the participants' age, the focus of the study (i.e. centered on the product or the process), and the context.

The question of whether models as a WCF technique should be conducted collaboratively or individually has caused much debate in the scholarly literature. Much of the research on writing collaboratively has shown benefits for accuracy against the effect of individual writing (Bueno-Alastuey & Martínez, 2017; Hidalgo & Lázaro-Ibarrola, 2020; Storch & Wigglesworth, 2007). However, empirical evidence has shown that, even if findings in terms of noticing are similar for both collaborative and individual writing, individual writing with models as WCF has been reported to foster more lexically-oriented noticing (e.g. Martínez & Roca de Larios, 2010).

Likewise, most of the research has supported that model texts are a valuable WCF technique, and empirical evidence has been mainly concentrated on adults' or teenage students' writing either in isolation (García Mayo & Labandibar, 2017; Martínez & Roca de Larios, 2010; Kang, 2020) or contrast with other types of WCF (see Hanaoka, 2006, 2007; Abe, 2008; Kim, 2015; Yang & Zhang, 2010, among others). As announced before, these studies have revealed that learners focus a great deal on lexical forms, and to a lesser extent, on grammar-related aspects. Initial research attempted to shed light on these beneficial effects by contrasting the use of models as WCF with other feedback techniques. For instance, Hanaoka (2006) and Yang and Zhang (2010) explored the role of reformulation and models with university students, revealing that the main focus of the learners using models was on noticing aspects of language, particularly those related to the lexicon. This goes in line with Abe (2008), whose study with EFL adult students with different proficiency levels revealed equal benefits for noticing language aspects. However, there were notable differences in noticing features of the language with less proficient learners

(Abe, 2008), who might have been more cognitively overloaded to pay attention to all differences (Schmidt, 2001).

The attentional focus on lexical features has been acknowledged in recent studies, such as Martínez & Roca de Larios's (2010) empirical research with teenage EFL students, where a comparison was established between individual and collaborative writing with models as WCF. In stark contrast with Abe's (2008) and Yang & Zhang's (2010) studies, whose participants focused more on lexical features, their study found participants in collaborative writing noticed more features related to ideational content in the comparison of their texts with models. Nonetheless, all the previous studies overall account for incorporations of several features (either language or content-related ones) into the revisions, supporting the view that models are effective WCF techniques. Similarly, García Mayo & Labandibar (2017) study reported that teenage EFL students, when writing individually, were more able to incorporate content features into their rewritten text when noticing was guided, an important variable which Martínez & Roca de Larios (2010) refer to as a potential contributor to increasing noticing of lexical concerns and ideational content. Interestingly, García Mayo & Labandibar (2017) revealed that when noticing was unguided, participants focused on lexis rather than on content features.

A common methodological feature shared by studies with model texts has been the use of note-taking as a noticing measure. For instance, García Mayo & Labandibar (2017), Hanaoka (2006), Kang (2020), Kim (2015), and Martínez & Roca de Larios (2010) provided their participants with note-taking sheets where they wrote the problems noticed in stage 1. This procedure was repeated in the comparison stage, where learners compared their written output with the model text provided. Studies, however, made use of different forms of note-taking. For instance, García Mayo & Labandibar (2017) used both guided and unguided noticing note-taking sheets whereas Kang (2020) used guided noticing note-taking sheets, where participants had to specify the type of problem encountered (e.g. "I wrote X, but I am not sure if this is correct" [Kang, 2020, p. 4]).

Although the previously reviewed research was fully conducted in adult or high-school settings, several studies have also investigated the effectiveness of model texts as WCF with younger EFL learners, i.e. Primary students. In this vein, Coyle & Roca de Larios (2014), Cánovas et al. (2015) and Kim (2015) reported that noticing was self-initiated by lexis, and young learners focused primarily on lexical and sentence-related features. In a similar fashion, these studies did not show any grammatical features as mediated by model texts. In the case of Kim's (2015) study, noticing was also reported in the differences between ideational content in parallel with Martínez & Roca de Larios (2014), even if this last one was not a young learners' based study. Interestingly, Coyle & Roca de Larios (2014) revealed that reflection on the output and potential errors was a cognitive operation fostered by models rather than error correction.

Coyle et al. (2018) also targeted the use of models by young EFL learners in the collaborative modality, and even if the focus was mainly on the trajectories followed by children, their findings support the previous empirical evidence. These indicated that young learners were more prone to allocating attentional resources to meaning over form and reproducing exemplars of language. Luquín & García Mayo's (2021) recent study with young EFL writers found that noticing was lexically-oriented at the comparison stage between the model and their written output, and in contrast with previous studies, content-related features were noticed as well. The findings reported in these studies, however, have to be cautiously considered given some factors: (i) all of them have been carried out in a collaborative setting given the purported language learning affordances for accuracy, especially for young EFL learners, (ii) some of these studies included training phases or instructional periods (e.g. Coyle et al. 2018; Luquín & García Mayo, 2021) in contrast to experimental research on model with no prior instruction, (iii) the use of procedures for WCF processing (e.g. guided or unguided noticing) is still inconclusive and more robust findings are required, and (iv) the vast amount of research on models as WCF both with children, adolescent and adult informants have been conducted in pen-and-paper environments, but not in digital ones. Therefore, more empirical evidence has to be accumulated regarding the use of computer-mediated writing since, to the best of our knowledge, no studies have addressed this issue thus far. Additionally, research on models as a WCF technique has focused on either the product as evidence in the studies reviewed, or the processing of WCF, but none of the studies reviewed thus far has focused on what writing processes young EFL learners are using when composing. Also, the act of writing collaboratively induces learners to engage in processes differently from individual writing. While writing collaboratively, learners interact with each other and activate scaffolding processes, and fostering co-constructing linguistic knowledge (Bueno & Martínez Lizarrondo, 2017). In contrast to this, individual writing entails that learners make use of their attentional resources, and hence use their own individual linguistic capacities.

Conclusion

These studies lend support to the use of models as an effective and genuine WCF technique to promote L2 language learning, but more importantly, to foster the noticing of certain language-related features such as lexicon, and/or despite the contradicting views, some ideas and expressions. Nevertheless, a number of questions await an answer in future research. Firstly, the abundant research on models with young EFL learners has relied on two axes: (i) collaborative writing, and (ii) product-oriented perspectives, with the recent emergence of WCF processing studies. To our knowledge, no studies have addressed the role of writing processes (i.e. planning, formulation, or revision) before and after the provision of WCF. Furthermore, the individual nature of the writing process leads us to consider the relevance of individual writing as essential for a number of reasons. On the one hand, the lack of individual writing studies on model texts with young EFL learners, and on the other hand, the fact that writing involves "individual choices of the writer in interpreting the task, managing her and/or his goals and purposes activating the knowledge linked to the task" (Spelman Miller, 2005, p. 303). As mentioned previously, the different product-oriented research has proved that model texts are beneficial, at least in collaborative writing with young EFL learners. In this vein, the effect of this type of WCF on a revised version of a text might be noticeable in the case of writing processes, since the repetition of a task after the provision of the model text may free up cognitive resources (Cánovas et al., 2015), and thus some related cognitive processes might be affected. The exploration of these cognitive processes has been manifold in L2 writing research tasks (for research on writing processes, see Roca de Larios, Marín & Murphy, 2001; Manchón & Roca de Larios, 2007; Roca de Larios, Murphy, Manchón & Marín, 2008), relying mainly on pen-and-paper environments, but the advent of the keystroke-logging software has broadened the scope. This type of software allows for getting an in-depth perspective of how the writing process is being developed thanks to the available research opportunities that it offers (see Van Waes & Leijten, 2005; Leijten & Van Waes, 2013) to observe and infer the underlying cognitive processes before and after the provision of WCF. Finally, to our knowledge, no studies have addressed the use of L2 writing by young EFL learners in a digital environment, and the potential effect of WCF on each of the three stages within the writing process, and on pausological behavior.

In the next section, we will present the theoretical foundations of the different models of writing processes in L1, and more specifically, in L2. This will be followed by a fine-grained synthesis of the current concepts and tenets of each of the macro and micro-processes of writing. This will include an overview of the different logging tools used to look into writing unobtrusively along with the relevance of using *Inputlog* (Leijten & Van Waes, 2013) for our research. Finally, the last section will explore the main concepts of pausological behavior research and a synthesis of different relevant variables receiving attention in the empirical literature, with a special focus on the role of keystroke-logging software.

2.6. Research on writing processes

The processes involved in writing a text – either in an L1 or L2 – are extremely invisible courses of action which give very detailed accounts of the writing performance and the writer's behavior. As complex a cognitive process as writing is, these courses of action are regarded as sub-processes being influenced by the working memory, and subsequently having an impact on the written output itself (Hayes, 2012; Hayes & Flower, 1980; Roca de Larios et al., 2006; Xu, 2017, 2018). When writing a text in an L2, writers encounter a series of difficulties that are inherent to the process itself: while producing their output, writers have to think thoroughly about the generation of ideas together with the cognitive effort of shaping them according to a series of grammar rules (Kellogg et al., 2013). These ideas materialize into the written output but are not ultimate since they undergo further revisions and editing on the writer's part.

The temporal distribution of writing processes is relevant since the process itself provides useful information abut the writer's strategic behavior. For instance, it allows for both observing and understanding the time-distributed management of attentional resources while writing in an L2, which may constitute an impediment, particularly for low-proficient learners. Also, observing writing processes purposefully points to the major issues encountered by the writer, for instance, if those problems are related to micro or macro levels within certain textual boundaries.

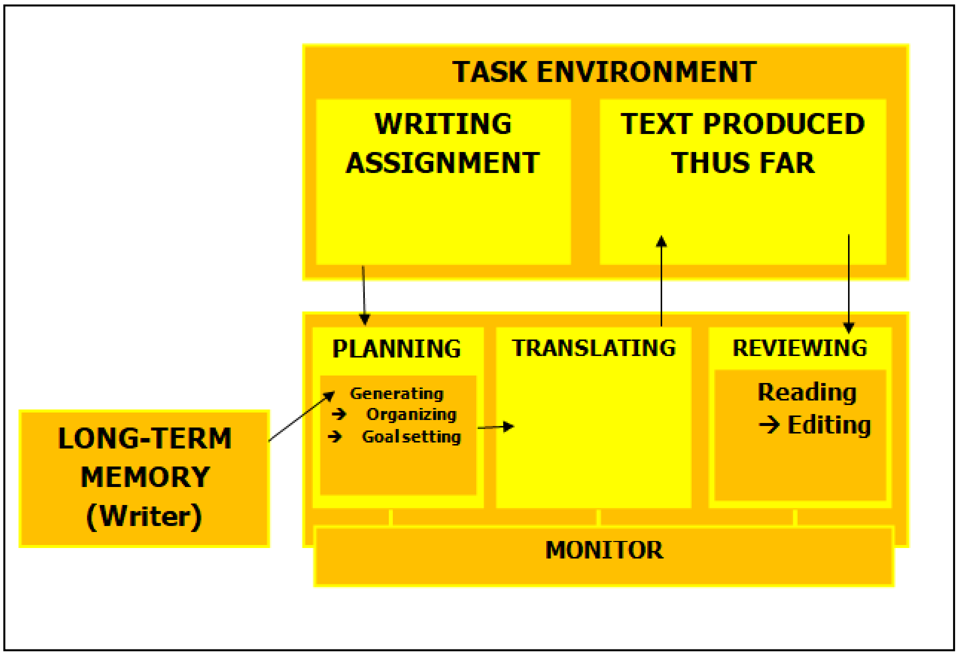
In the sections that follow, we first present relevant theoretical positions on the psycholinguistic processes during writing. Thus, the characterization of writing will be discussed through the different models of L1 and L2 writing. Subsequently, we examine the different macro-processes (i.e. planning, formulation, and revision) to gain a better understanding of what each of these processes entails, and we critically assess what research has suggested thus far.

2.6.1. Models of L1 and L2 writing processes

The temporal dimension of writing processes has been subject to study in recent decades as a result of the interest in looking into how learners approach writing and the way this writing process is addressed (Manchón, Roca de Larios & Murphy, 2009). This investigation trend began back in the 1980s when researchers first endeavored to break down the different processes associated with composing (Stapleton, 2010). Those studies were aimed at uncovering the processes in an L1 environment in an attempt to identify what processes take place in the writers' minds. Much later, attempts at observing how writing processes developed while engaging in L2 writing started to generate interest in the scholarly community. This is certainly relevant for the present study, as ours is an attempt to observe whether and what processes are altered as a result of the provision of WCF during L2 writing.

Inspired by the area of problem-solving research in psychology, the conceptual categorization of mental processes, methodological perspectives to the issue, and previous research on problem-solving aspects led to the emergence of seminal studies such as those of Hayes & Flower (1980) or Bereiter & Scardamalia (1987). These studies set the stage for a whole body of research addressing writing processes in an L1. Hayes & Flower (1980) developed their influential model with a case study in which one participant reflected upon his mental processes during composing. This seminal model, which was further theorized in Flower & Hayes (1981), consisted of three differential parts: the task environment, the writer's long-term memory, and the writing process as a whole (see Figure 1 below). The first of these parts, the task environment, comprises the writing assignment, which includes the topic, the audience the output is addressed to, the motivating cues, and the text produced thus far. Secondly, the writer's long-term memory comprises the knowledge of the topic, knowledge of the audience (applying the appropriate linguistic conventions to the addressees), declarative knowledge, stored writing plans (applying the text conventions appropriately to the text type), and text standards. Eventually, the writing process as a whole comprises three main processes: planning (generating ideas, followed by organizing and goal setting), translating, and ultimately reviewing (including reading and editing). Flower & Hayes's (1981) model theorized that writing was a recursive process against the product-based perspective of writing as a linear process (see Galbraith, 2009, for a review). These processes are reported to be highly dependent on the monitor, conditioning the performance and continuum of the composing.

Figure 1.
Original Hayes & Flower's (1981) model of written production in the L1.

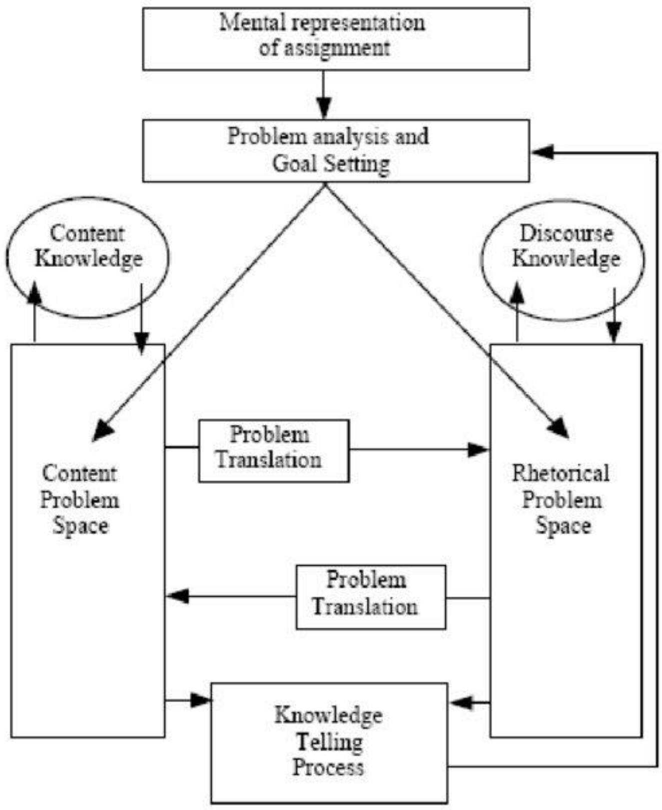


In view of that, the following three main stages were undertaken by L2 learners: planning, translation (we will assimilate this process to the concept of formulation), and revision. It is of paramount importance to note that when Flower & Hayes (1981) refer to *translating*, they do not allude to translating from an L1 to an L2, but rather to the transition of ideas generation into sentences. The temporal dimension of the writing process is certainly longer in an L2 seeing that writers have to shape their thoughts into the L2 before these ideas are eventually materialized. These processes interplay with each other in a cyclical fashion, which is precisely how texts are constructed (Manchón et al., 2009), hence assuming the importance of time allocation to each of the stages involved and their influence over the cognitive load on the writers' part. Most of the models of written production – be it in an L1 or L2 – have lent support to the claim that producing a text is predominantly a dynamic multioperation, as numerous processes interplay with each other in a cyclical and not a linear way.

As referred to previously, most of the research on L2 writing processes has leaned on what L1 research on the composing process has suggested (Myles, 2002). Far from being ideal models depicting the process itself, both models (Bereiter & Scardamalia, 1987; Flower & Hayes,

1981) have neglected some important features of L2 writing, such as the constraints of individual differences (e.g. L2 proficiency). Flower & Hayes (1981) have looked into the research potential of the problem-solving nature of writing as an activity, such as the rhetorical situation and the writer's goals, which ultimately shape their final output (Myles, 2002). Accordingly, Flower & Hayes's (1981) model has been regarded as excessively individualistic given that it does not take into consideration other features involved in the process such as the social dimension (Swales, 1990).

Figure 2.
Bereiter & Scardamalia's (1987) model of writing (in Galbraith, 2009, p. 10).



Bereiter & Scardamalia's (1987) model (see Figure 2 above) takes a cognitive-developmental perspective (Ranalli et al., 2019). Their model is composed of three distinctive models. Nonetheless, each will be understood as part of the whole piece proposed by the authors. These three types of knowledge are as follows: (i) *knowledge-telling*, in which writers do simply retrieve information from their memory and tell about it exactly as they perceive it, notwithstanding the text conventions, (ii) *knowledge-transforming*, wherein the writer established connections between their own mental representation of ideas and their “separate mental representation of the text” (Ranalli et al., 2019, p. 2), (iii) *knowledge-crafting*, which Kellogg (2008) added to this model in order to depict the particularities of professional writers whose task is beyond the mere entrenchment of words and whose efforts are pushed to additional mental representations. Neither this model nor Flower & Hayes (1981) bear in mind the importance held by cross-cultural references and sociocultural variation (Kern, 2000; as cited in Myles, 2002). In this respect, those models have failed to account for the implications of L2 writing. For instance, writing conventions are very much ingrained in native speakers’ minds as opposed to L2 learners, who are still developing their language skills and moderately acquiring the conventions of each writing genre. Additionally, both models neglect that language learning is continuous, which is in development throughout the whole process.

Nevertheless, these models have undergone several revisions in the scholarly community as a result of the several drawbacks encountered (Hayes, 1996, 2012; Kellogg, 1996). Ultimately, Hayes (2012) re-conceptualized his view on the model on the basis of his more recent research studies. This renewed conceptualization comprises a *control level* (motivation, goal setting, and current plan or writing schemas), a *general process level*, with *writing processes* (proposer → evaluation and/or translator → transcriber) and *task environment* (collaborators & critics, transcribing technology, task materials, written plans, and text written thus far), and a *resource level* (attention, working memory, long-term memory, and reading). If compared to the seminal model and subsequent revisions (Hayes & Flower, 1980; Flower & Hayes, 1981), several changes have been applied: the transcription process and the motivational aspect have been inserted, both of which are newly conceived for this model (Hayes, 2012), as well as the disappearance of the planning and reviewing processes. In contrast, Hayes (2012) regards the planning process as a completely different writing process "that produces a text designed to aid the author of the plan in producing another text" (p. 376) (see section 2.6.2. for a review of the planning process). In other words, Hayes (2012) has considered these processes as duplications of the writing process itself, especially when it involves a written plan. He also admits that "writers create many plans that they do not write down" (p. 376), and hence these processes are complete writing processes themselves, rather than part of the written output *per se*. Hayes (2012) accommodates the view that they are an integrative part of the own writing process. planning and revision processes are both essential to writing, and their characteristics do certainly aid in shedding light on what writers allocate time in.

2.6.2. L2 writing processes

Research on the process of composing in L2 derives from the major interest in how learners made use of their resources to write in an L1 (Stapleton, 2010), most of which took as starting reference the seminal work by Flower & Hayes (1981) addressed at L1 writers, as mentioned in the previous section. In what follows, a synthesis of the empirical research in the different macro-writing processes will be performed, which will help us gain a deeper insight into the intricacies of the temporal distribution of the writing process. The relevance of this review of recent research along with a clear definition of these macro-writing processes is rooted in what Galbraith et al. (2007, as cited in Manchón, 2014) refer to as a two-way interaction between the writing processes, given the cyclical nature of the writing activity. On this basis, it should be borne in mind that writers, when engaging in a writing task, generally manipulate the phases to suit their needs. As a result, none of these shall be understood as linear or separate processes. The engagement of writers with each of the macro-writing processes is largely dependent on the writers' needs and goals, as these will very likely condition the interaction between each of these processes. Given that WCF shapes writers' goals by encouraging them to improve the first version of their written text, the role of WCF is central given its potential effect on the cognitive demands and attentional resources in one or other stages of the writing process.

2.6.2.1. Writing process: planning

As said before, the first of the phases as initially proposed by Flower & Hayes (1981) referred to *planning*, which is regarded as an anticipatory process for any activity (Manchón & Roca de

Larios, 2007), and which Hayes (2012) later reconceptualized as a complete independent writing process itself prior to the production of another text. Planning is a course of action presenting a series of guidelines leading the writer towards the successful completion of a written task. The planning process as an operation within the writing process of a single text has been widely examined in the last three decades (Hayes & Nash, 1996; Johnson, 2020; Manchón & Roca de Larios, 2007; Roca de Larios et al., 2006; Roca de Larios et al., 2008; Roca de Larios, Murphy & Manchón, 1999; Ong, 2014). Regarded as the initial operation executed by writers, research has cautiously demonstrated that planning diminishes as the composing process progresses (Ong, 2014). Planning shall be regarded as an essential operation whereby learners anticipate their ideas so as to materialize them coherently since “the previous discourse or the text produced so far act as constraints on what has yet to be spoken or written” (Manchón & Roca de Larios, 2007, p. 552). This major necessity of interweaving the text produced and the most immediate upcoming output implies that great cognitive load is placed upon the central executive (Kellogg et al., 2013), thus constraining the amount of time that may be spent on other operations. In this regard, Hayes & Nash (1996) posit the idea that pre-task planning is advantageous given the fact that it facilitates the mental representation of the task and how it is thought to be executed, hence requiring less cognitive load than when writers face planning whilst composing. In this vein, writing itself is considered as a complex task in nature (Manchón, 2009) which learners are reported to complete by engaging in a possible route by which the task is subdivided into subsequent subtasks (Manchón & Roca de Larios, 2007) taking planning as the initial drive for it. In spite of the previous reference to writing as a route, it shall not be taken in the literal sense of the word seeing that the writing process itself is “non-linear and recursive” (Ong, 2014, p. 18).

Temporality in the process of planning, and any of the ensuing phases (formulation and revision) is especially uneven given that the time allotted to each phase fluctuates (Manchón & Roca de Larios, 2007). The temporal dimension of writing is highly related to the working memory, in which the studies by Kellogg (1990) shall be underscored. Ong (2014) reviewed the theorizings which look into the time allotted to planning, where two hypotheses were exponentially described vis-à-vis the concept of time and its place as a task environment factor. The Interaction Hypothesis (Long, 1983, 1989) and the Overload Hypothesis (Kellogg, 1990) are especially related to the planning process, and both hypotheses are opposed to each other. The former posits several ideas such as that text quality will decrease as a function of the planning condition and that writing lacks linearity and is by and large recursive. As for the Overload Hypothesis, its claims were divergent from the Interaction Hypothesis in that planning operated to enhance text quality, as this writing process reduced the amount of cognitive demand, acting as a working memory liberator, enabling the writer to focus on other processes. Advantageously, what the Overload Hypothesis proposes is a shift from planning to focusing on other cognitive operations which require a higher degree of cognition. Bearing this in mind, L2 writers often encounter greater difficulties than L1 writers as they "require additional attentional resources to notice gaps in their knowledge" (Ong, 2014, p. 19), but namely, because they are likely to fall upon linguistic and content-related obstacles – similar to the ones L1 writers encounter – plus the ones associated with the L2.

Of special interest for this PhD is the relevance of the expertise in writing (Sasaki, 2002), considering that more skilled L2 writers are reported to take a longer time at the planning stage than novice writers. Thus, novice writers, as young EFL learners are regarded, will not devote initially as much time to planning as their skilled peers. This finding is at some point contradictory if it is assumed that young EFL learners will need more planning on account of their novelty as writers. Yet, many other factors contribute to rejecting this assumption given that young writers address their attentional focus towards ideas generation, especially when being provided with WCF (e.g. Coyle et al. 2019; Luquín & García Mayo, 2021), and how to express them accurately in L2. With reference to the aforementioned ideas, it seems convenient to point out what Cummins (1989) referred to as advanced planners and emergent planners. Whilst advanced planners generally opt for outlining a plan for their writing task, emergent planners begin composing directly in order to "find[ing] their focus, and appeared to create plans during the course of composing" (Manchón & Roca de Larios, 2007, p. 578).

Nevertheless, and in addition to the writer's focus on planning their writing, there are a number of similarities between L1 and L2 writers in terms of this process, as was reviewed by Manchón & Roca de Larios (2007), who conveniently pointed out the existence of several connections and disparities. Through the planning process, L2 writers have been reported to generate ideas (Sasaki & Hirose, 1996; Uzawa, 1996), goal-oriented organization (Skibniewski, 1988), and modes of composing which shift from intensive to reflective and vice versa (Pennington & So, 1993). Also, special attention is paid to text organization before writing as well as simultaneously during the composing process (Sasaki & Hirose, 1996). All of these are reported as similarities between the L1 and L2 as regards the planning process. Differences, as reviewed by Manchón & Roca de Larios (2007), are related to the number of goals, which have been reported to be lower in L2 writing since the attention focus is directed toward lexicality and morphosyntax (Skibniewski, 1988; Whalen & Ménard, 1995). In the same line, Jones & Tetroe (1987) found that L2 writers devote much less time to the planning process due to what they are able to retrieve from their L2 repertoire, resulting in a reduction of the amount of planning.

In their study, Manchón & Roca de Larios (2007) found that L2 writers devoted the first third of the composition process to planning. Interestingly, data revealed that L2 writers were constrained when the time was a variable in the task, thus altering the writing process. In the same vein, they pinpoint that owing to the very non-linear nature of the composing process, these processes may occur in a variable way depending upon the task environment. More recent studies (see Ellis & Yuan, 2004), have accounted for the effects that planning may exert over linguistic operations. For instance, it has been found that L2 writers undergoing a pre-task planning phase displayed a vast focus on fluency and syntactical variety. Conversely, online planning brought about accuracy, and no planning did have negative consequences over the composing process and the final written output (Ong, 2014). These findings suggest an advantage in the attention to a planning phase but are not conclusive as to its relevance as a cognitive operation (see Johnson, 2020, for a recent review). Nevertheless, it has also been a matter of contention whether L2 writers focused their planning time chiefly on text organization or, conversely, they continued placing their attention on “content generation and organization of ideas” (Ong, 2014, p. 20) along the process of writing. While Ellis & Yuan’s (2004) study evidenced a focus on text organization, Ong & Zhang’s (2013) findings revealed that it was not a linear, clearly separate operation but rather a dynamic one.

The dynamism of writing may be explained through the interplay of processes that learners, as main executors, control to their needs and adapt to this said varying situation (Van den Bergh & Rijlaarsdam, 2001). As a result, it might be concluded that the planning phase and its associated operations (generating ideas, organizing them, connecting them, and setting goals) (Ranalli et al., 2019) are reported to be part of the composing process depending on the task environment (Roca de Larios et al., 2008; van den Bergh & Rijlaarsdam, 2007) or task conditions (Ong, 2014). From a temporal perspective, subprocesses of writing, mainly generating and organizing ideas (Kellogg et al., 2013), have been reported to have a differential effect on working memory (Hayes & Flower, 1980). Kellogg’s (1996) model, however, suggested that the load on the working memory was roughly balanced.

The studies on planning reviewed in this section reveal that low proficient L2 writers, as is the case of young EFL learners, tend to resort to planning in several ways: (i) emergent (Bereiter & Scardamalia, 1987) or online planning, since they do not reflect upon what they are going to write prior to the act of writing, and (ii) what they focus on when they are planning, e.g. ideas, lexis. In sum, the findings of these studies have to be considered with caution since, on the one hand, scarce studies examining the planning behavior of children have been conducted and, on the other, the tendencies shown in adult writers might not reflect the reality of young writers.

Importantly, given the complexity of the act of writing, the study of macro processes, such as planning, needs to be complemented with the microprocesses involved in each one of them so as to obtain a fuller picture of the writing composing phenomenon. The following synthesis will provide an overview of how previous theoretical positions and empirical endeavors have conceptualized planning at different levels. The limited attentional capacity model (Foster

& Skehan, 1996; Skehan & Foster, 2001) and Robinson's (2003) Cognition Hypothesis operationalize planning in two subprocesses: (i) online or within-task planning, and (ii) pre-task planning. Drawing on this distinction, Sasaki (2000, 2002, 2018) distinguishes between *global planning* and *local planning* within online or within-task planning. For this dissertation, we opted for Sasaki's conceptualization of the subprocesses of planning. The reason for this decision is based on two premises central to the purposes of our research study: children are considered as limited capacity processors (Leow, 2015; McLaughlin et al., 1983) and emergent planners (Bereiter & Scardamalia, 1987), along with their consideration as low proficient language learners and unskilled writers. This has led us to merge several sub-categories of planning (namely, the organization of the text, ideational content, or procedural aspects) from previous empirical attempts with adults or teenagers (e.g. Manchón & Roca de Larios, 2007; Roca de Larios et al., 2008) into these broader categories.

Regarding *global planning*, Sasaki (2018) associates it with "the overall planning of the content of the text to be written" (p. 3). This global planning procedure involves interruptions, which, from the perspective of the time-distributed nature of the writing process, are generally expressed in terms of pauses. This characterization is very similar to Spelman Miller's (2006b) view that longer pauses might be indicative of macroplanning processes such as paragraph organization. This also goes in accordance with Hayes & Nash (1996), whose research indicates that longer pauses reflect more planning. Hayes & Nash's (1996) operational definition of planning involved process planning and textual planning. Within this last category, textual planning is subdivided into (i) abstract text, and (ii) language planning. Thus, our conceptualization of global planning would fit in both subdivisions given the generation of content, and the importance of formulation of content in the writing process (Baaijen et al., 2012).

The inference of cognitive processes behind longer pauses preceding long textual units (sentences or paragraphs) has been present in countless studies (e.g. Barkaoui, 2019). This conceptualization of global planning is in line with Sasaki's (2002) study where he identified low proficient writers to pertain to his "Type C" writing style, that is, those who wrote according to rough global planning (i.e. thematic planning) together with consecutive planning episodes. In this sense, this type of planning behavior corresponds to Bereiter & Scardamalia's (1987) knowledge-telling model. Thus, the conceptualization of global planning assumes the act of planning the ideas or content for the text. As mentioned before, three aspects are key to unveiling global planning: (1) the long duration and (2) location of a pause between paragraphs and sentences. This includes some isolated global planning episodes preceding words, and finally, (3) the surrounding environment, since according to Spelman Miller (2006b), the text produced (thus, the formulation process) and the revision behavior surrounding the pause help to understand whether the writer was planning globally or not.

As for *local planning*, our view fits Sasaki's (2018) definition that this type of planning involves "planning what to write next without any sense of the overall picture of the composition to be formed" (p. 3). Hayes & Nash's (1996) conceptualization of language planning partially fits within local planning since it refers to the mental planning activity of grammatical text for potential inclusion in the text. Similarly, Bereiter & Scardamalia's (1987) knowledge-telling model is a hallmark of this consecutive local planning procedure since text generation is produced without a general overall plan. Thus, if the writer plans at a local level, pauses might be (1) shorter in length and (2) situated at lower-level units such as words or phrases, rather than entire clauses or paragraphs (Spelman Miller, 2006b). From this assumption, the operationalization of local planning involves largely shorter pauses which reflect microplanning processes, mainly grammatical and lexical choices (Barkaoui, 2019).

The focus of the studies reviewed has been almost exclusively placed on an adult or teenage populations. Our study is once again an attempt to both characterize young EFL learners' planning behaviors and observe whether or not the provision of model texts as WCF affects planning and its subprocesses. In light of this, model texts have been reported to engage writers in the generation of both ideas and lexicon (Luquín & García Mayo, 2021; Martínez & Roca de Larios, 2010), which leads us to consider that model texts might affect the degree of global and local planning.

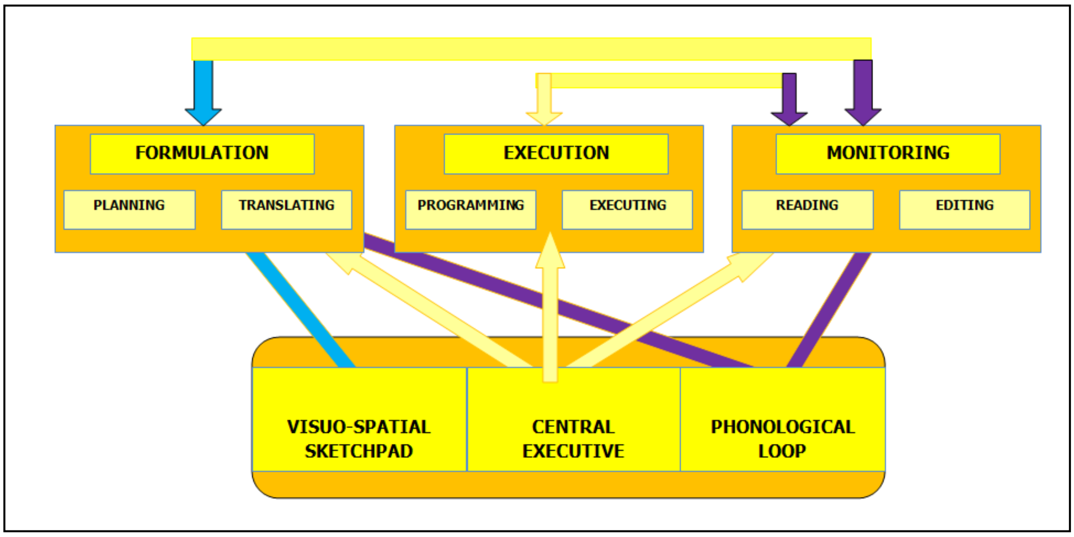
2.6.2.2. Writing process: formulation

L2 writing has been proved to be particularly time-consuming, requiring arduous effort and too much focus on form (Silva, 1992), which has at times made accuracy prevail over the textual organization. This concentration on the form has given rise to diverting the attention away from goal generation (Skibniewski, 1998). For all these reasons, L2 is thought to restrain not only the formulation of ideas (Roca de Larios et al., 2006) but also L2 writing on its whole.

The second of these cognitive operations in the writing process is *formulation* (Roca de Larios et al., 2006; Roca de Larios et al., 2008), which was early described by Flower & Hayes (1981) as *translation*. When learners engage in the writing task, they undergo a series of processes that cyclically interplay with each other. In this cyclical process, formulation occupies a central role since it is reported to occur in the middle of the composing course and is thought to drag along as the longest operation (Manchón et al., 2009; Roca de Larios et al., 2006). From an L1 perspective, Flower & Hayes (1981) regard the formulation process as a translation or transcription of ideas. On the L2 counterpart, Roca de Larios et al. (2006) define formulation as the process of converting thoughts into the language since writers have to enquire themselves about the type of information they are going to produce and the way it is going to come out. Added to the inherent difficulty of materializing our mental ideas into written words, L2 writers struggle much more in the task of formulating these ideas and shaping them according to their needs.

Attending to Kellogg's (1996) working memory model of composing (see Figure 3), he distinguished three systems within the working memory, on which the basic writing processes hinge. Visual-spatial sketchpad, central executive, and phonological loop are the components of the working memory used by the six basic processes of writing (Kellogg et al., 2013; Kellogg, 1996). these six processes are planning, translating, programming, executing, reading, and editing. Nevertheless, these refer to the working memory in written composition particularly in the L1, which does not exempt it from being transposed to L2 settings (Kellogg, 2008; Kellogg et al., 2013; Révész et al., 2019). Of interest in this regard is the association of the working memory to each of the cognitive operations in the composing process given the temporal nature that they display. As Kellogg et al. (2013) point out, those cognitive operations are at some point dependent upon the execution of the central executive as a demand to the working memory.

Figure 3.
Kellogg's (1996) working memory model of composing.



Most of the models of composing (see Bereiter & Scardamalia, 1987; Flower & Hayes, 1980, 1981) have considered the thought-process as a separate entity from the cognitive operation of translating them into words. Nevertheless, this view should be taken very cautiously as regards the role of thoughts that might be included in the text production process rather than separately. Interestingly and very conveniently, Roca de Larios et al. (2006) identify the process of formulation as the hierarchical blending of goals and ideas encompassing two types of limitations, i.e. one which entails the morpho-syntactical features of the L2, and a second one, which is the execution of these features into sentences accurately incorporated into the writing thread (Roca de Larios et al., 1999). The sub-process associated with sentence construction has been reported to take two-thirds of the time devoted to composing (Wang & Wen, 2002), which is allegedly the most important operation to which writers resort. In line with this finding is Roca de Larios et al's. (2008), who detected that lower-level writers – for instance, young EFL learners – devoted approximately 81% of the composing time to formulation. This finding goes hand in hand with a previous study (Roca de Larios et al., 2006) where they identified that L2 writers devoted more time to solving formulation problems than in their L1 owing to the compensatory nature of their problem-solving skills. Accordingly, L2 writers take a substantial amount of effort in formulating their ideas and materializing them to reflect their writing goals. In his case study, Stapleton (2010) revealed that the participant devoted 33% of the time to compose. Nonetheless, Stapleton's study allowed the participant to use resources other than her linguistic knowledge, which alters the variables affecting the composing process in comparison with other studies. Nevertheless, these findings have to be considered with caution since, to date, scarce studies have explored writing processes, and more importantly, the process of formulation in young EFL learners. The exception was López et al.'s (2019) study on collaborative writing, which reported that upper-primary EFL learners devoted nearly 52% to formulation processes. Again, caution must be taken when considering these findings since our study only deals with individual writing.

The formulation process implies, as seen, several problem-solving operations that writers must engage in to attain their prime goal, i.e. that of materializing their ideas into a set of paragraphs as part of a whole macrostructure as a text is. In this vein, problem-solving is highly linked to the construct of fluency, which has been dissimilarly operationalized in the scientific literature. In this sense, "fluency" has been defined from a cognitive perspective as involving efficient access to linguistic knowledge and retrieval of linguistic form (van Gelderen & Oostdam, 2002). Important data in the writing process such as word count in final but also linear text are particularly revealing for this aspect. Similarly, studies have made use of other measures such as word per minute (Van Waes & Leijten, 2015), characters per minute (Palvaiainen et al., 2012) or have used bursts as the referential unit of measurement, i.e. words per P-Burst (Révész et al., 2017). Though occurring in a different fashion in L1, L2 learners have been found to produce fewer words in their final written output as well as pause a great deal more as a result of engaging in a thought-process which undoubtedly hinders the production of words in contrast to L1 writing (Roca de Larios et al., 2006).

The number of pauses is equally relevant for our research since they give us an explicit overview of the necessary effort that is involved, and its close connection to the formulation process. As a consequence of this cognitive effort, L2 writers tend to pause in an erratic way (Silva, 1993; Ransdell & Levy, 1998). Nonetheless, Roca de Larios et al. (2006) conveniently pointed out that the main hindrance of research in this area has been that fluency is an invisible process, only attestable through transcription. The advent of new technological software has allowed for widening the scope of this research towards paths of clarity where this invisible process has been at least partially unveiled (such software will be scrutinized in section 2.7.). For introductory research, see Van Waes & Leijten, 2005, 2006). Thus, our research intends to connect the empirical need to accumulate evidence regarding the process of formulation in young EFL learners with the recent efforts to tap into writing processes through this type of software. As previously mentioned, the formulation process entails a series of problem-solving operations which may constraint the act of writing through a series of interruptions as well as thought-process operations. One of the endeavors of this Ph.D. is to shed light on a missing gap in the literature: how the formulation process changes as a consequence of processing WCF and, more specifically, the differential effects of WCF on the potential cognitive effort regarding the translation of goals and ideas into the text as revealed by pauses.

2.6.2.3. Writing process: revision

The third cognitive operation referred to by models of written production (Flower & Hayes, 1980, 1981; Bereiter & Scardamalia, 1987) is *revision*, about which several other models have been proposed throughout the three decades of research on writing processes (see Faigley & Witte, 1981; Barlett, 1982). To define it, revision involves revisioning or reseeing the text (Barkaoui, 2007), which is an evaluation of what has been written up to that moment, leading to changes if this process of self-reflection of their output requires it. Along the same lines, revision entails the suspension or termination of text production in order to evaluate the produced text in an attempt to transform it (Van den Bergh et al., 1994; as cited in Xu, 2018). In general terms, revision entails a complex activity where the writer performs a series of changes from reading, evaluating, and altering the text representation (Alamargot & Chanquoy, 2001; Chanquoy, 2001, 2009; Stevenson et al., 2006).

To provide a clear picture of revision, an initial model (Faigley & Witte, 1981) distinguished among *surface revisions* and *text-based revisions*. Whilst the former meant changes at the accuracy level, including orthographic ones, the latter entails changes at a contextual level, where meaning is altered (Barkaoui, 2016). Surface revisions do include “meaning-preserving changes” (Barkaoui, 2007, p. 96), for example, addition, deletion, or substitution. In contrast, text-based revisions may be classified into macrostructure and microstructure changes. Respectively, one modifies the overall meaning or focus of the text, while the second involves a minor overall revision not impacting the global understanding of the text. The degree of attention allocated to the process of revision has been deemed as recurrent in the case of skilled writers, who have operated it throughout the composing process. Significantly enough, revision has been proved to be especially useful to provide a reformulation of thoughts and ideas previously materialized for the sake of meaning-reconstruction and text enhancement (Faigley & White, 1981; Reynolds & Bonk, 1996).

By its very nature, revision has been taxonomically reviewed in the models proposed in the scholarly community (Lindgren & Sullivan, 2006; Stevenson, Schoonen & de Glopper, 2006), for instance, Lindgren & Sullivan’s (2006) model put forward a dual view of categorizing revisions concerning their position and its purported effect on text production (Barkaoui, 2016). Just as Faigle & Witte (1981) suggested, this model also includes a pair of revisions at an internal and external level: *internal revisions* are generally framed within the scope of conceptuality and planning as a means of restructuring the planned thoughts without visibly reporting them into the written output. Very importantly, internal revisions are invisible in the composing process except at the temporal level, where they might be interpreted as pauses (Barkaoui, 2016). Following Lindgren & Sullivan's (2006a, 2006b) taxonomy, *external revisions* are visible changes in the writing process insofar as the changes undergone are transferred and applied to the written output. A further subdivision is made (as seen in Table 2): *precontextual revisions* imply modifications at the final point of the text produced up to that moment, which means that this type of revision can be preceded but not followed by text. Precontextual revisions can be either conceptual or centered on the form. The second subtype, c*ontextual revisions*, entails moving away from the point of inscription to apply changes (insertions, deletions, substitutions, or rearrangements) to the already written text.

Table 2.
Taxonomies on revision.

Study	**Revision taxonomy**
Faigley & White (1981)	• Surface revision ○ Meaning-preserving changes • Text-based revision

Lindgren & Sullivan (2006)	▪ Internal revision
	  ○ Prelinguistic revision
	  ○ Pretext revision
	▪ External revision
	  ○ Precontextual revision
	  ○ Contextual revision
	    ▪ Form revisions
	    ▪ Conceptual revisions

Each of the macro-writing processes reviewed so far has displayed a number of sub-processes that account for the complexity in their conceptualization. The coordination of such cognitive operations and their own sub-processes is required for writing to occur. The concurrence of all these processes entails a limitation in the use of cognitive resources (Hayes, 2006; Kellogg, 1996; Torrance & Galbraith, 2006; cited in Barkaoui, 2016) which hinders to some extent their degree of performance. This limitation imposes a cognitive load on writers that restrain them from allocating attention to other operations. As precisely pointed out by several scholars (e.g. Alamargot & Chanquoy, 2001; Chanquoy, 2009), providing there is not any automatization in some operations (for instance, typewriting or handwriting in the case of young learners), writing processes may hold the cognitive resources at the writer's disposal. This means that their demands may be increasing and dynamic, i.e. L2 writers may resort to planning in the middle of the writing process, and immediately revise a chunk of the text belonging to the very beginning (Broekkamp & van den Bergh, 1996).

The number of operations in the revision process is highly varied depending on models put forward by scholars. Scardamalia & Bereiter (1987) proposed a three-stage model of mental representation of the revision process, whose operations were:

- Establishing a comparison between the intended text which is thought to be produced and the written texts stored in the long-term memory.
- In case a disagreement between the mental model of the written text in LTM and the intended written text appears, the writer undergoes a diagnosis to detect where the disagreement lies.
- After this diagnosis, the discord is solved by modifying or altering the text, and possibly, the goals. Very importantly, such changes may occur in the written output or as a mental representation (Barkaoui, 2007).

Besides this model, Barkaoui (2007) reviews another model by Hayes, Flower, Schriver, Stratman & Carey (1987), where revision was divided into four basic process-oriented categories. Nonetheless, both models (Hayes et al., 1987; Scardamalia & Bereiter, 1987) should not be taken as fixed operations followed by the writer as these are highly dependent on individual differences such as writers' aptitude to writing, particularly in an L2, where constraints and difficulties are much more recurrent. Along this line, L2 writers with limited working memory capacities might be constrained by the revision process on its whole. Such a situation entails that writers are more focused on providing a solution to linguistic errors (surface-level revision) than to content or ideational ones (meaning-level revision), possibly as a result of the lower cognitive load required by surface-level revisions (Hacker, 1994). It is convenient to point out that the amount and type of revision in L2 writing vary depending on variables such as the writing mode, i.e. pen-and-paper versus digital writing. Along this line, Barkaoui (2016), in a study with digital writing with L2 adult writers, identifies very precisely several factors at both an individual and contextual level which are reported to affect the way in which these revisions are undertaken, the specific time at which these are carried out and their main focus. Those individual and contextual factors comprise (i) L2 proficiency, which shall be regarded as the core factor influencing the composing process, (ii) writing expertise, which shapes the style and conventions of the text by enhancing the style and content. This factor is highly dependent on L2 proficiency, (iii) task type, which conditions this expertise by establishing the limits imposed in terms of text conventions (e.g. if it is a formal or informal text) and the audience, (iv) time constraints, given the fact that L2 writers are by and large under classroom conditions where the writing task must be finished at a specific timing, and (v) writing mode.

In particular, the first individual factor, L2 proficiency, has been the main constraint imposed on L2 writers. Under this premise, children would be assimilated as low L2 proficiency and unskilled writers. As such, it has been reported that low L2 proficiency writers devote very little time to revision throughout the composing process, most of which is concentrated at the end of the whole writing task composition (Alamargot & Chanquoy, 2001; Chanquoy, 2001, 2009). Among the operations such unskilled writers perform within the revision process, they are generally surface-level, such as word changes or deletions, and less likely insertions. High-level revisions (content alteration or text organization) are less probable for a number of reasons, being the main one the high cognitive load which is demanded in order to perform these revisions on the part of unskilled writers (Broekkamp & van den Bergh, 1996; McCutchen, 1996). Aside from these particular difficulties, low L2 proficiency writers are expected to find themselves in conflicting situations regarding their attentional focus, thus increasing the number of linguistic problems. Low-level L2 writers are reported to allocate their attentional resources to forms rather than meaning-level aspects such as ideational content (Porte, 1996). Such focus on form has led to their neglection of other textual aspects such as the generation of ideas (Whalen & Ménard, 1995), an aspect which goes hand in hand with Myles's (2012) claim that learners revise more at a superficial level and devote little reflection time to their written text. Thus, should revision take place, it would be directed towards form. Other distracting aspects to which lower level L2 writers may not attend are (a) the complexity required in the writing task, i.e. if it exceeds the learner's ability to apply the L2 to the conventions of the text, and (b) the writing environment or medium, that is, the use of traditional settings such as pen-and-paper or the use of computers which entail typewriting skills (Barkaoui, 2016). Nonetheless, the question of whether other underrepresented populations in this kind of studies, such as young EFL learners, would revise at the same levels as teenage and adult learners still await an answer.

Given the importance of computers as a mediating medium for writing in this PhD, research on the revision process using this writing environment (see Chan et al., 2017) has also been a matter of concern in the last decades. While a more detailed account of these studies as regards the whole writing process will be provided in section 2.7.1., a general overview of the influence of computers on the revision process is to be presented in what follows.

Research on the revision process under this computer environment has brought to light several interesting findings. In the advent of keystroke logging software for writing, Bridwell et al. (1985) conducted a case study of advanced undergraduate students on writing memos in an L2 business writing course. Their study shed light on the potential effects of computers on revising, revealing a focus on surface-level revisions and the overall structure of the text. Another study by Stevenson et al. (2006) with 253 EFL Secondary students explored the role of online revision using keystroke-logging software. Findings revealed that L2 writers revise a great deal more than their L1 counterparts, which coincides with Van Waes & Schellens's (2003) study. In their study, 60 experienced writers – university faculty and graduate students – wrote using a word processor, capturing keystrokes, while 20 others wrote in pen & paper. The findings revealed that L2 learners using a word-processor software revised in a more extensive fashion in the initial stages of the composing process, but disregarded the role of revisions when the written product was completed.

Barkaoui (2016) examined the role of time constraints for the revision process in L2 writing, in which 54 undergraduate participants belonging to low and high proficiency levels and keyboarding skills. Using Inputlog, they wrote two timed writing tasks (one independent, lasting 30 minutes, and the other integrated, in which they had to listen to a lecture, reading a text about the topic for five minutes and write a summary of both activities in 20 minutes). Findings revealed that writers tend to focus their attention on form rather than content revisions, as well as the importance of L2 proficiency for revisions, which seemingly had some effect on it. Using Lindgren & Sullivan's (2006) taxonomy, L2 writers were reported to resort to precontextual revisions at the expense of contextual ones, and linguistic revisions were more recurrent than content ones. Very importantly for the temporal dimension, participants tended to allocate their precontextual revisions to two-thirds of the composing process whereas the last third was allocated to contextual revisions.

For instance, in Xu's (2018) study, 57 Chinese EFL university learners took part in a study where, after a short instruction period on the conventions of the text type of the task, they had to write a text in a computer environment using a keystroke logging program called Inputlog 6.0. (for an extensive review, see section 2.7.1.). This allowed the researcher to check upon online revisions undertaken by the participants, and findings revealed that low proficiency L2 learners are more likely to revise on a frequent basis, in reduced scopes, and without engaging in end revisions. Interestingly, revisions at an immediate time were found to be particularly important for reformulation at the expense of planning or revision.

In a more recent study, Kim (2020) explored the incidence of undergraduates' writing profiles on the distribution of different writing strategies. Using a text processor software and Camtasia, 260 Korean university students wrote an initial and final draft of an opinion essay in an L2. They were allowed to use several internet resources in order to plan their writing content, and have recourse to search for any type of information or queries on the Internet. In spite of the efforts to maintain a naturalistic setting of writing, participants were encouraged to proceed in a very specific way: planning had to be produced on a separate page, and revision was subdivided into two drafts. Findings revealed that among the four writing profiles which were distinguished (revision-based, plan-based, search-based and correction-based), correction-based writers were predominant while revision-based profiles displayed higher text quality compared to the other profiles. Kim used several measures to account for planning and revision operations, of which the ones concerning the latter are particularly relevant for this PhD. Basing his decision on other studies (see Bridwell, 1980; Daiute, 1986; Faigley & White, 1981, Kehagia & Cox, 1997), the author was inclined to use a linguistic-based taxonomy encompassing features of proposals as the ones presented in this section: *microscopic* and *macroscopic* revisions. As Kim (2020) himself states, both of these types of revisions are very similar in nature to the ones proposed by Bridwell (1980). *Microscopic revisions* are applied when writers resort to modifying their written compositions at a surface, lexical or phrase level. In contrast, *macroscopic revisions* are framed within modifications at clause, sentence or text level. This type of revision is aligned with young learners' characteristics as limited working memory processors, supposedly leading them to concentrate their effort on linguistic errors, that is, surface-level revisions (see Hacker, 1994).

As seen throughout this section, several theoretical underpinnings on the conceptualization of revision have been outlined on the basis of previous empirical endeavors. Much of it has concentrated on the location and nature of the revision process (Adams et al., 2010; Bridwell, 1980; Faigley & White, 1981, Kehagia & Cox, 1997) and others have included abstract thoughts as part of the revision process (Lindgren & Sullivan, 2006), which are invisible to the researcher's eye. Our focus on the revision process would take a more linguistic and textual approach by following the taxonomy proposed by Kim (2020). Table 3 summarizes these taxonomies by synthesizing where they belong according to Kim's (2020) taxonomy.

Table 3.
Correspondence of taxonomies in previous studies with ours as based on Kim's (2020).

Our categorization based on Kim (2020)	Studies	Revision taxonomy
Microscopic revisions	Faigley & Witte (1981)	Surface changes (Formal and Meaning-preserving changes) Text-based changes (Microstructure changes).
	Lindgren & Sullivan (2006)	External revision (Contextual revision: form revision [typograhpy, spelling, grammar, punctuation and format, meaning-preserving]; precontextual revisions).
Macroscopic revisions	Faigley & Witte (1981)	Surface changes (Formal changes **above sentence level** and meaning-preserving changes) Text-based changes (Macrostructure changes).
	Lindgren & Sullivan (2006)	External revision (Contextual revision: form revision [typograhpy, spelling, grammar, punctuation and format, meaning-preserving]; precontextual revisions).

2.7. Logging tools to study digital writing processes

With the advent of computers and word processors, the use of computer-mediated writing has exponentially increased at the expense of traditional writing, i.e. pen-and-paper (Leijten & Van Waes, 2005). The great difference existing between the latter and the former is that computers –

and subsequently, mobile phones and Tablet PCs – have widened the scope of possibilities in which writing as a mode might be used (Ho & Savignon, 2007), whereas in pen-and-paper settings writing has traditionally been relegated to the usual types of communication, i.e. letter writing, academic writing, and personal writing as self-reflection aspects. Along these lines, computer-mediated writing has expanded the said scope by providing writers (and hence, learners) with countless tools whose main functionality is giving writing output a relevant communicative role, for instance, through blogs or social networks. The use of computers has allowed for the emergence of other web applications as is the case of wikis or Google docs, thus responding to calls for collaborative settings (e.g. learners writing collaboratively or correcting their peers' works). In addition to that, the former, Google docs serve well to explore the writing processes (Li, 2018), especially in collaborative settings.

The use of computers for students' learning is a potential motivational factor thanks to which some studies have reported improvements in L2 writing given that learners are socially immersed in computer-mediated environments (Kelm, 1995; Warschauer, 1996). Nevertheless, writing in computer-mediated environments (henceforth, CME) must be cautiously regarded since the learners' use of these resources is highly influenced by external sources, i.e. if a learner is writing an essay on a computer at home, s/he is very likely to resort to dictionaries or any other useful sources to complete the assignment. Such great access to numerous sources must be considered from the perspective of autonomous writing and the writing setting. The advent of CME has paved the way for its use as a very beneficial writing tool since writers are able to compose recursively and revise their writing on a more frequent basis. More interestingly, it becomes even more evident that writing is not a linear process given that writers might start a sentence but modify it right afterward at a precontextual and contextual level (Eklundh & Kollberg, 1996). As has been mentioned before, precontextual revisions are modifications to the text at the point of inscription, that is, the revision is preceded by text but not after it. Contextual revisions, however, are modifications to the text away from the point of inscription, which means that the writer goes back to another part of the text and edits it (Lindgren & Sullivan, 2006). Along the same lines, writing on a word processor is regarded as more comfortable for the writer since the constraints of pen-and-paper writing – involving modifications on the text, taking more time and effort – vanish, and revisions in computer writing are more easily carried out (Van Waes & Schellens, 2003).

With reference to the working memory and within the framework of difficulties in computer writing, it should be noted that handwriting does not entail the same cognitive load as typing. While learners might be reasonably skilled at handwriting, typing may be an obstacle for them in case they are not accustomed to it. For instance, young learners may encounter great difficulties at typing if their skill is not acceptably developed (Bourdin & Fayol, 1994). Furthermore, their cognitive load is much higher than that of skilled writers since their central executive might be fully exhausted, and thus less attention is placed on the basic processes of writing (planning, formulating or revision) given the insufficient capacity of the central executive to attend to both typing or handwriting mechanics and spelling (Kellogg et al., 2013). As seen, the constraints imposed on young learners at the time of handwriting or typing entails limitations to the processes of composing (Swanson & Berninger, 1996), an aspect which must be kept at very close hand when analyzing writing from young EFL learners, the processes performed and the resources allocated. All things considered, the non-automatization of keyboard use is a distraction for L2 writers, affecting their attentional focus, and the writing process itself, as they might devote pauses to high-level cognitive operations and their execution pauses to typing (Alves et al., 2007).

Just as it was mentioned previously, writing processes are generally invisible and cannot be examined unless one resorts to online procedures (e.g. think-aloud protocols), which has prevented studies from delving into the subject unobtrusively (Barkaoui, 2016; Leijten & Van Waes, 2005, 2006). The advent of computers and technological advances has not only permitted learners to enhance their L2 writing and linguistic development (see Li et al., 2017, for a review on language-learning tools), but it has also allowed researchers to gain deeper insight into how writing processes may be observed, and the available tools to unveil the underlying cognitive processes unobtrusively. Ultimately, such drive for non-obtrusive methods has led to the development of software that registers keystrokes and mouse movements to be further processed, rendering research on written processes ecologically valid (Van Waes, Leijten & van Weijen, 2009). Keystroke logging programs are defined as recording a writing activity whilst writers are constructing their texts on the computer (Miller & Sullivan, 2006).

The need for a consistent methodology for researching writing processes has been steady (Révész & Michel, 2019) with methods ranging from on-site observation to think-aloud protocols (Hyland, 2016; Smagorinsky, 1994). These data-gathering techniques have been commonplace across studies on L1 and L2 writing processes, but as precisely indicated by Spelman Miller (2005), their application has not been exempt from criticism on account of the reliability of the data with reference to the processes involved (Russo et al., 1989). Such controversy has been highlighted by Leow (2015), who expressed his views on the appropriateness of think-aloud protocols in contrast to off-line measures. In summary, the research tools which have preceded the advent of KLS, and whose application is still in vogue are: (i) online: on-site observation of the writing process, for which note-taking is deemed to be essential (for example, see Matsuhashi, 1981, 1987); verbal reports, for instance, think-aloud protocols (TAP), which is one of the most used techniques (Révész & Michel, 2019), but whose reliability has dwindled as a result of its obtrusive nature in the process.; and (ii) offline: questionnaires, whose main purpose has pushed writers to reflect upon their writing process, not without the variable which might influence such reflection by diverging the reality of the process itself. Also, stimulated recalls and retrospective interviews are included under this category. These procedures will be detailed later.

Keystroke logging software (henceforth, KLS) has broadened the scope of writing research (Sullivan & Lindgren, 2006) since “these new technologies [have] also created new possibilities for writing researchers to investigate writing as it unfolds in real-time” (Van Waes, Leijten, Wengelin & Lindgren, 2011, p. 507). Furthermore, they have made it possible to analyze processes which are invisible to the researcher’s eye. As reviewed by Van Waes et al., (2011), KLS are particularly functional for research purposes as the data collected is much more finetuned. Such a detailed account is rendered possible as KLS stores or *logs* keystrokes and the time stamp (either in seconds or milliseconds), indicating when each key was pressed and released. Remarkably, this KLS is "non-obtrusive" (Van Waes, Leijten & van Weijen, 2009, p. 41) as the data are collected without having to disrupt the writer to reflect upon the writing process or the difficulties encountered (e.g. in the case of observations, video analysis or think-aloud protocol), and thus making our research purposes a bit more visible (Plakans, 2009; Xu & Ding, 2014; as cited in Conjin et al., 2019). KLS has been reported to cover a wide range of research foci (Conjin et al., 2019), for example, to identify writer profiles (Karnan, Akila & Kirshnaraj, 2011), text quality, writing fluency (Van Waes & Leijten, 2015), emotional states (Salmeron-Majadas, Santos & Boticario, 2014), task complexity (Grabowski, 2008; Révész et al., 2017), motor functionality, linguistic features (Leijten, Macken, Hoste, van Horenbeeck & van Waes, 2012), writing processes (Ranalli et al., 2019; Van Waes, van Weijen & Leijten, 2014; Xu, 2018) and temporal aspects of writing (Latif, 2008).

The aforementioned applications of KLS programs point to their usefulness for this study. More specifically, this PhD intends to contribute to assessing the usefulness of KLS as applied to young EFL writers, an underrepresented population in this research field, and (ii) thus far, KLS has proved to be particularly practical in uncovering the time allocation of writing processes in isolated writing tasks, but its potential has not been utilized for revealing changes in time allocation of writing processes when learners are provided with feedback. Our study intends to fill this research gap.

2.7.1. KLS programs: an overview

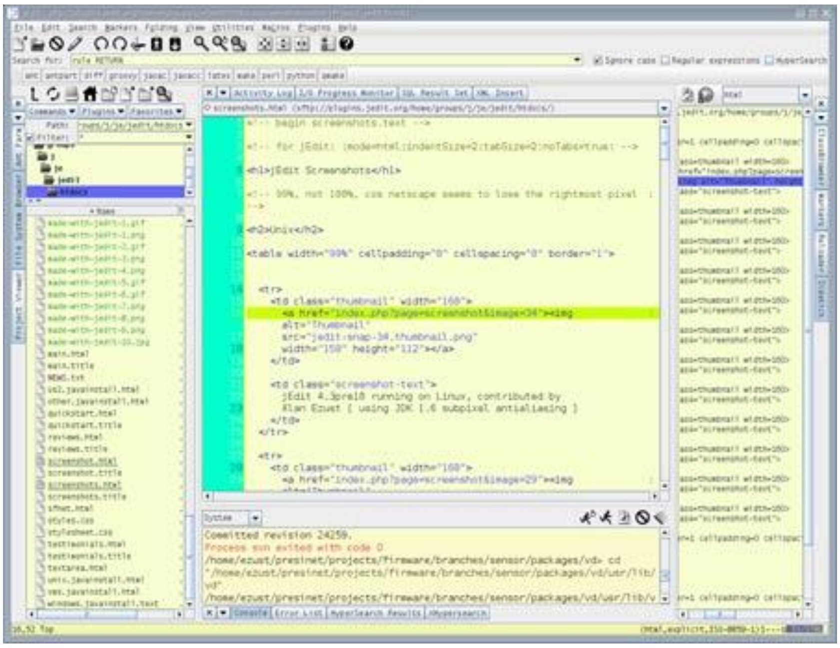
In what follows, a detailed account of each KLS, their purpose, and the scope of implementation is presented. Given the existing number of KLS, and the different applications in the scientific literature, the purpose of this review is to characterize each of them, since, as noted by Van Waes et al. (2011), KLS programs might "have overlapping capabilities, but all of them also have very specific characteristics making them complementary in their use" (p. 508). Likewise, the review will help us discern the reasons that justify why *Inputlog* was selected for the present study.

An overview of programs other than Inputlog

(1) JEdit and Trace-it *environment*. Both programs, though perfectly functional as separate entities are supplementing each other, which makes them part of a set of keystroke logging programs. It was developed at KTH by researchers (Erklundh, 1994; Eklundh & Kollberg, 1992; Kollberg, 1998; Severinsson Eklundh, 1996) who pioneered the use of technological tools such as computers for writing research. Originally intended for Macintosh, it has not been developed thus far so that its use could be spread to Windows settings. Nonetheless, its research potential amounts to a series of characteristics that were seminal for other KLS programs (e.g. Inputlog). JEdit allows researchers to obtain basic statistics about the writing sessions on the basis of data about pauses, keystrokes pressed and released, input devices used (for instance, at the time, diskettes).

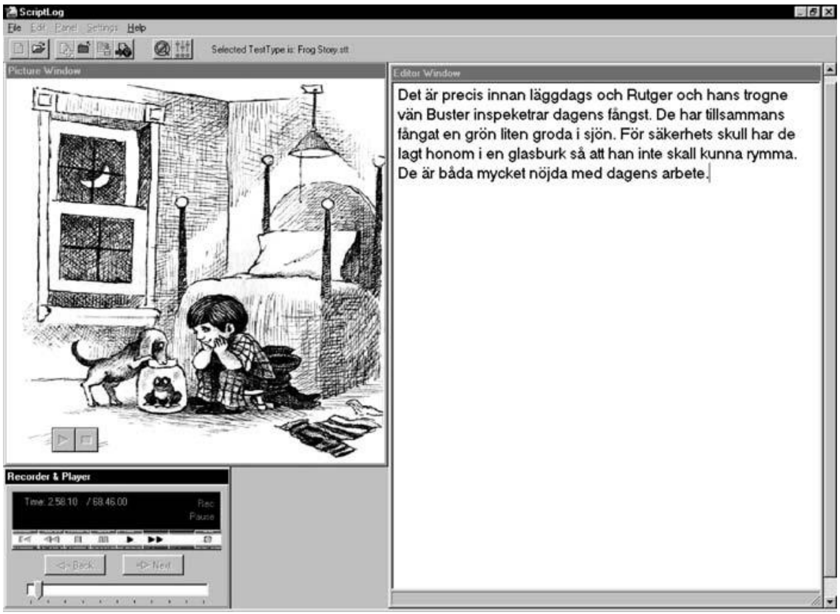
As reviewed by Leijten & van Waes (2005), JEdit is basic word processor software that undoubtedly depends on the functionality of Trace-it which, as aforementioned, works separately by obtaining the log file from JEdit. In doing so, JEdit and Trace-it combined to display the number of revisions performed throughout a whole computer writing session. Interestingly enough, JEdit and Trace-it allow for a replay of the composing process and various interactive analyses (Van Waes et al., 2011). Learners compose in the word processor (JEdit) which logs the writing session into an MID file, which stands for movement, insertion, and deletion (Leijten & van Waes, 2005) and which contains the data on elementary operations in the writing process associated with their time stamps. This MID file is generated by JEdit but only readable by Trace-it, which allows researchers to analyze the data with S-notation (Kollberg, 1998) by placing the focus on the revision process. In this vein, such revisions are presented in order of occurrence but without altering the internal structure (Leijten & van Waes, 2005). In spite of its importance and influence on the development of KLS, JEdit and Trace-it have not been developed further from 1990 onwards, and their characteristics and interface remain the same. Nonetheless, both tools are at any researcher’s disposal to be downloaded (http://www.nada.kth.se/iplab/trace-it).

Figure 4.
Screenshot of jEdit.



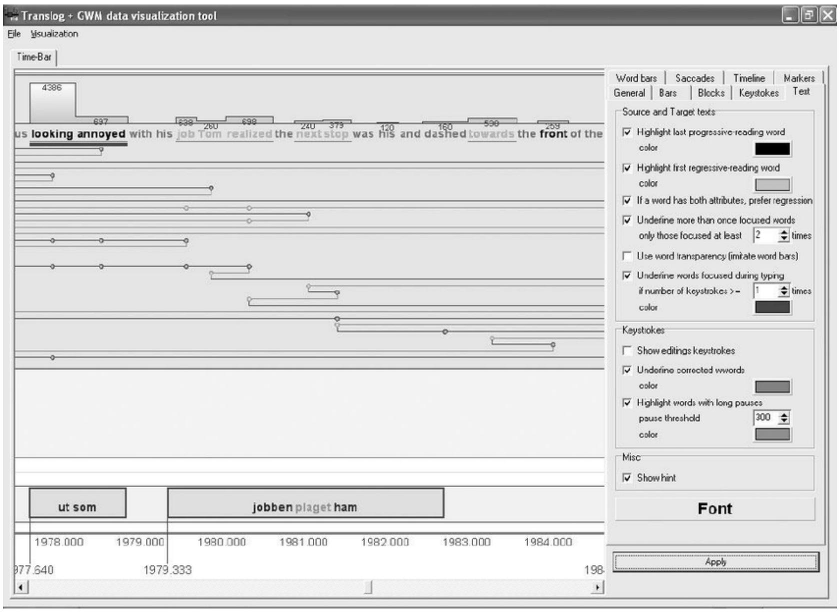
(2) ScriptLog is a keystroke logging program originally developed by Strömqvist & Malmsten (1997) as part of a research project. The seminal version was initially conceived for Macintosh, but a Windows version was later developed including enhanced functions (Andersson et al., 2006; Wengelin et al., 2009; as cited in Van Waes et al., 2011). Similar to JEdit, ScriptLog allows for logging keyboard operations during the composing process, with very useful information on temporal distribution of processes, position, mouse action and key press and release. Despite its similarity to the previous KLS program, ScriptLog software environment is composed of a text editor and an additional window frame, allowing for the inclusion of input – for instance, task prompts, such as pictures, texts or sound files – working as eliciting tools (see Figure 5). Much as other KLS programs (see (6) Inputlog), ScriptLog contains a module that allows recording these events on the keyboard and any other function or underlying operation associated with it. Interestingly, ScriptLog has a wide degree of flexibility in order to customize the analyzes derived from the data, i.e. the researcher might modify time intervals to fine-grain the data. Another noticeable functionality of ScriptLog is that it has included the possibility of eye-tracking data, which favors the study of monitoring (reading). As JEdit and Trace-it, ScriptLog allows the researcher to replay the writing session (Van Waes et al., 2011).

Figure 5.
The interface of ScriptLog with a picture prompt and the text editor window on the left.



(3) Translog is another KLS program whose main focus is placed not only on writing process, but especially on translation processes (see Figure 6). Originally developed by Jakobsen (2006), Translog is composed of two frames or functionalities: (i) a component acting as supervisor to record and log all the keystroke data, and (ii) a user component, which intends to provide the translator with a source text and a separate frame where the translation shall be written (Van Waes et al., 2011). Translog seems an interesting tool for writing processes associated with translating from an L1 to L2, but its relevance for newly-produced texts in an L2 goes beyond its purposes.

Figure 6.
Screenshot of TransLog.



(4) uLog is a more technical KLS program which, although not being developed by researchers on writing, enables researchers to “observe, describe, and code behavior” (Van Waes et al., 2011: 511). uLog allows for logging all keyboard and window events as well as keystrokes. Despite its easiness at setting it up, which exempts it from being installed on the computer, uLog does not offer a keystroke analysis functionality.

(5) EyeWrite was one of the first attempts to integrate KLS functions with eye-tracking records. Developed by Simpson & Torrance (2007) and Wengelin et al., (2009, as cited in Van Waes et al., 2011), EyeWrite has two different components: (i) basic text editor with simple functions, and (ii) analysis tool. However, this KLS program captures eye movements with a head-mounted eye-tracker (Van Waes et al., 2011), permitting the identification of the word on which the writer is fixating his eye. By way of the analysis tool, keystroke logging information is conveniently combined with the eye-tracking data file, thus providing a general file with both data and the possibility to execute a replay of the writing session. It should be noted that one of the most attractive aspects of EyeWrite lies within its potential to display where the eyes are fixated at the time of replaying.

An overview of Inputlog

(6) InputLog is a KLS program developed in 2003 at the University of Antwerp, and its main purpose goes in line with that of the KLS programs previously described, i.e. to study writing processes in detail from an online perspective (Leijten & Van Waes, 2006, 2011, 2013; Van Waes, Leijten & van Weijen, 2009; Van Waes et al., 2011). Originally, Inputlog was developed taking as a starting point a study on the influence of speech recognition on the writing processes (Leijten & Van Waes, 2003).

The KLS programs presented above share a common characteristic: the text editors were integrated into the software. However, Leijten & Van Waes (2005) discerned that the reality of writing on a computer entailed the use of different sources and word processors, such as ordinary word processors in a Windows environment (Microsoft Word). Most KLS tools are conceived for the writing process to occur in their text editors, and none of them is accurately adaptable to commercial word processors. Out of this necessity, Inputlog was originally devised so as to provide the research setting with a more natural writing environment where writers did not feel odd or out of their comfort zone (Leijten & Van Waes, 2005).

Apart from the aforementioned KLS resources, Inputlog is also distinguishable from other KLS programs in several exclusive characteristics: its functionality to be used in other text editor environments, the XML format of the output, the parsing technology and the speech recognition functionality (Van Waes, Leijten & van Weijen, 2009). Such possibilities allow the researcher to delve into the composing process from different angles given the number of analyses that Inputlog offers (all of which will be detailed afterward).

Inputlog has a number of features that are at the researchers' disposal to facilitate their research task: (a) the logging tool allows the researcher to record any Windows-based word processor; in other words, it captures any keypress, cursor and mouse movement, (b) after one Inputlog session, you can generate XML files from a wide array of analyses which are performed by the KLS program itself, (c) several programs may be combined with Inputlog data in order to expand the scope of the analysis, such as eyetracking features or speech recognition, and (d) as it was mentioned previously, Inputlog allows for a replay of the composing process, an aspect which might be used for the benefit of a more fine-grained analysis. For instance, it may be used in retrospective interviews (Leijten & Van Waes, 2011).

Along the same lines, the potential of Inputlog for investigating writing processes has proved to be quite convenient from a technical point of view as a result of (1) its accessibility for use under Windows-based word processor programs, unlike other KLS tools which have their editor. This allows for greater flexibility in the implementation of several data collection procedures, and it reduces the scope of obtrusiveness into the composing process itself as the environment within which it is performed is adapted to the reality of the participants, (2) its straightforward interface facilitates the setup of parameters according to our research interests, and (3) such setting of parameters provides the researcher with a great degree of flexibility to adapt the analyses to certain theoretical foundations even after having gathered the data.

To gain a deeper insight into the functionality of Inputlog, the output analyses available to researchers are going described in the following paragraphs. Firstly, it must be noted that Inputlog offers two different types of analysis: general analysis and advanced analysis. Building on previous theoretical and empirical findings in the scholarly community (Leijten & Van Waes, 2014), *General analysis* is purely based on the logging data obtained from Inputlog; advanced analyses process the raw logging data algorithmically.

- *General Analysis* might be regarded as the core analytic procedure as it provides information about all keystrokes, mouse movements, and the time stamps in a linear fashion, among other important data (Van Waes & Leijten, 2015). Additionally, this analysis includes information on revision in the XML file.

Figure 7.
General Analysis output data

#Id	Event Type	Output	Position	DocLength	Character Production	StartTime	StartClock	EndTime	EndClock	ActionTime	PauseTime	PauseLocation	IntervalFixedSize	IntervalFixedNumber	X	Y	Re
0	focus	Inputlog 8.0.0.6			0	4781	00:00:04	4781	00:00:04	0	0	CHANGE	1	1			
1	mouse	Movement			0	4781	00:00:04	5515	00:00:05	734	0	INITIAL	1	1	803	206	
2	focus	TASKBAR			0	5515	00:00:05	5515	00:00:05	0	0	CHANGE	1	1			
3	focus	WordLog_HES4P_20190225132731.docx			0	5578	00:00:05	5578	00:00:05	0	0	CHANGE	1	1			
4	mouse	LEFT Click			0	5578	00:00:05	5672	00:00:05	94	0	INITIAL	1	1	803	206	
5	mouse	Movement			0	5734	00:00:05	5953	00:00:05	219	62	MOUSE	1	1	800	214	
6	keyboard	LSHIFT	0	1	1	54734	00:00:54	54810	00:00:54	76	48781	COMBINATION KEY	1	1			
7	keyboard	t	0	1	1	55130	00:00:55	55194	00:00:55	64	396	BEFORE SENTENCES	1	1			
8	keyboard	BACK	1	2	2	56554	00:00:56	56612	00:00:56	58	1424	REVISION	1	1			
9	keyboard	CAPS LOCK	0	1	2	57250	00:00:57	57322	00:00:57	72	696	UNKNOWN	1	1			
10	keyboard	T	0	1	2	57970	00:00:57	58042	00:00:58	72	720	BEFORE WORDS	1	1			
11	keyboard	H	1	2	3	61066	00:01:01	61146	00:01:01	80	3096	WITHIN WORDS	1	1			
12	keyboard	E	2	3	4	61506	00:01:01	61586	00:01:01	80	440	WITHIN WORDS	1	1			
13	keyboard	SPACE	3	4	5	62278	00:01:02	62322	00:01:02	44	772	AFTER WORDS	1	1			
14	keyboard	M	4	5	6	64066	00:01:04	64234	00:01:04	168	1788	BEFORE WORDS	1	1			
15	keyboard	A	5	6	7	64866	00:01:04	64946	00:01:04	80	800	WITHIN WORDS	1	1			
16	keyboard	L	6	7	8	68666	00:01:08	68770	00:01:08	104	3800	WITHIN WORDS	2	1			
17	keyboard	E	7	8	9	69298	00:01:09	69362	00:01:09	64	632	WITHIN WORDS	2	1			
18	keyboard	F	8	9	10	70963	00:01:10	71066	00:01:11	103	1665	WITHIN WORDS	2	1			
19	keyboard	I	9	10	11	71794	00:01:11	71906	00:01:11	112	831	WITHIN WORDS	2	1			
20	keyboard	C	10	11	12	72498	00:01:12	72578	00:01:12	80	704	WITHIN WORDS	2	1			
21	keyboard	SPACE	11	12	13	72914	00:01:12	72994	00:01:12	80	416	AFTER WORDS	2	1			
22	keyboard	D	12	13	14	73330	00:01:13	73411	00:01:13	81	416	BEFORE WORDS	2	1			
23	keyboard	O	13	14	15	73778	00:01:13	73842	00:01:13	64	448	WITHIN WORDS	2	1			
24	keyboard	C	14	15	16	74130	00:01:14	74210	00:01:14	80	352	WITHIN WORDS	2	1			
25	keyboard	T	15	16	17	75042	00:01:15	75114	00:01:15	72	912	WITHIN WORDS	2	1			

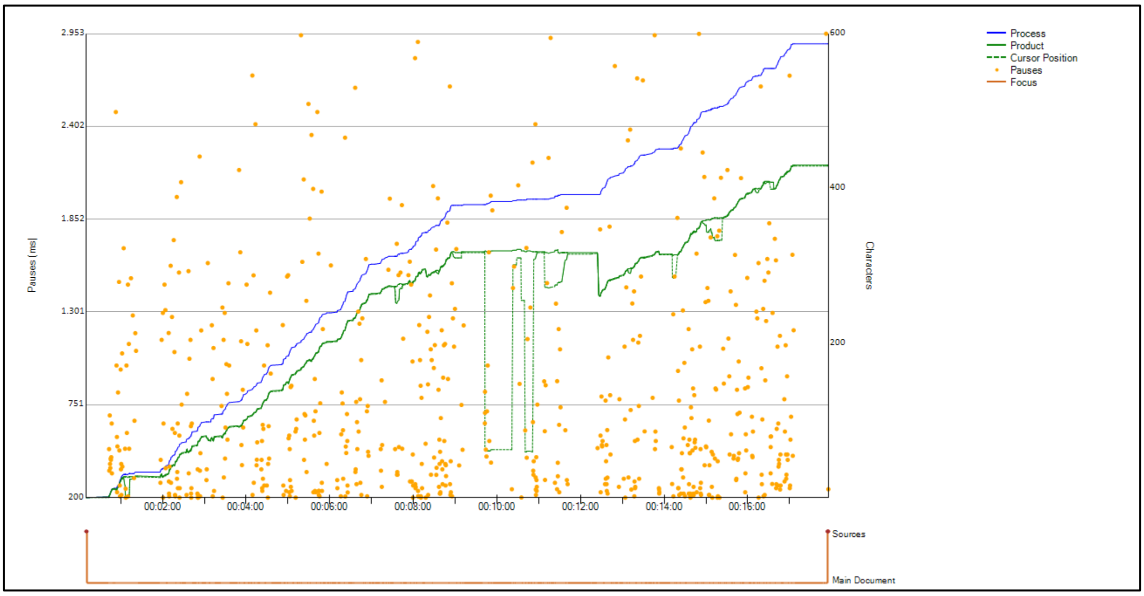
- *Linear Analysis* takes into consideration the text as a linear entity by organizing the cursor movements and keystrokes in order of occurrence. The parameters on which it is based include pause threshold in milliseconds (which has to be set to the researcher's interests), linear analysis type (fixed number of intervals, focus-based, revision-based, or condensed), and the length of intervals in seconds (for instance, 20, 30, 60 o 120 seconds).

Figure 8.
Example of revision-based linear analysis.

Interval	Output
0: INSERT	{12854}[Movement][LEFT Click][Movement][Movement]{2777}[M
1: DELETE	[BACK][BACK][BACK][BACK]{8768}
0: INSERT	tif·{20616}drink·a·meid
2: DELETE	[BACK][BACK]
0: INSERT	dicine·{11632},·{23755}
3: INSERT	[Movement][LEFT Click][Movement][Movement]ther
4: DELETE	[HOME][BACK][Movement][LEFT Click][Movement][BACK][BACK][
5: INSERT	[CAPS LOCK]T[CAPS LOCK]here·{3264}are·a·dog·and·a·{4736}sc
6: DELETE	[BACK]
7: INSERT	entf·
8: DELETE	[BACK][BACK]
9: INSERT	·{3790}[Movement].
10: DELETE	[BACK][BACK]
11: INSERT	

- *Process Graph* is generated from the data gathered, and it displays the progression of the writing process (x-axis) and the number of characters produced (right y-axis), while the left y-axis showcases the length of pauses employing orange dots. In addition to this, Figure 9 represents the number of transitions made between different sources according to the source analysis data (Leijten & Van Waes, 2014).

Figure 9.
Sample of Process Graph with Inputlog 8.0.



- *Summary Analysis* deals with aggregated data, which is represented by total process time, total pausing time, nº of pauses and their mean, the number, and length of P-bursts. As Van Waes & Leijten (2015) established as a parameter in a study, the pause threshold may be modified for the sake of the research study interests.

Figure 10.
Summary analysis showing information processing in Inputlog.

Process Information	
Keystrokes Produced in This Session	
Total Keystrokes incl. Inserted and Replaced Characters in Main Document	305
- Total Non-Character Keys	0
- Characters Inserted	0
- Characters Replaced	0
- Total Typed (incl.spaces)	305
- Per Minute (incl. spaces)	23.722
- Total Typed (excl.spaces)	236
- Per Minute (excl.spaces)	18.356
Words	
Total Words in Main Document	59
Per Minute	4.589
Mean Word Length	3.814
Median Word Length	3
Standard Deviation Word Length	2.266

- *Pause Analysis* gives a detailed account of the pausing processes and behavior by presenting data on the number of pauses, the mean pause length, where pauses were located in the text (e.g. within or between words), and the phase (Van Waes & Leijten, 2005). The pause threshold might be tailored to the research study interest (from 0 to 5000 milliseconds). The manner of presenting the data is variable, and might be chosen according to the researcher's needs: it may be generated on a general or a more specific interval level – in which case the whole composing session is divided into 10 time slots.

Figure 11.
Pause Analysis showing the section "General Information".

General Information	
Overview	
Total Process Time	00:12:51
Total Pause Time	00:08:34
Total Active Writing Time	00:04:17
Total Process Time (s)	771.425
Total Pause Time (s)	514.412
Total Active Writing Time (s)	257.013
Proportion of Pause Time	66.683 %
General	
Total Number of Pauses	43
Arithmetic Mean of Pauses (s)	11.963
Median Pause Time (s)	6.688
Geometric Mean of Pauses (s)	7.468
95% CI Log-Transformed - Low Boundary (s)	5.545
95% CI Log-Transformed - High Boundary (s)	10.057

- *Source Analysis* is another function whereby Inputlog obtains data relating to window events, including statistics on the total time spent on each event, the number of keystrokes performed, a window transition statistics, which displays the transitions from window to window undergone by the writer, and finally a detailed overview of intervals of each window event (Leijten & Van Waes, 2014).

Figure 12.
Source analysis output.

Window Title	Total Time (s)	Total Time (relative)	Total Keystrokes
* Time before the first window *	12.854	0.017	0
Inputlog 8.0.0.6	0.874	0.001	0
TASKBAR	1.029	0.001	0
WordLog_Noa Carrillo Pérez_20190527133105.docx	756.668	0.981	430
Total	771.425	1.000	430

Window Transition Statistics

From Window Title	To Window Title
Inputlog 8.0.0.6	TASKBAR
TASKBAR	WordLog_Noa Carrillo Pérez_20190527133105.docx
WordLog_Noa Carrillo Pérez_20190527133105.docx	TASKBAR
Total	

- *Revision Analysis*, which is further divided into a Revision Matrix and S-Notation (Kollberg, 1998).

The first analysis output presents a summary of the type of revisions, i.e. normal production, deletion, insertion, production + revisions, all of which are specified as regards their number, edits[1], duration, length, characters (with and without space), and the number of words. Furthermore, a statistical analysis of R-Bursts is provided. Apart from this summary, a chart is offered to the researcher with a detailed account of each type of revision, together with more information on the time (for instance, when it started and ended, duration, the location of each revision, and the position of the cursor). For researchers, this Revision Matrix data output is a clear asset as revisions may be looked into in more detail to fine-grain the research data to our study needs.

Figure 13.
Revision Matrix output data.

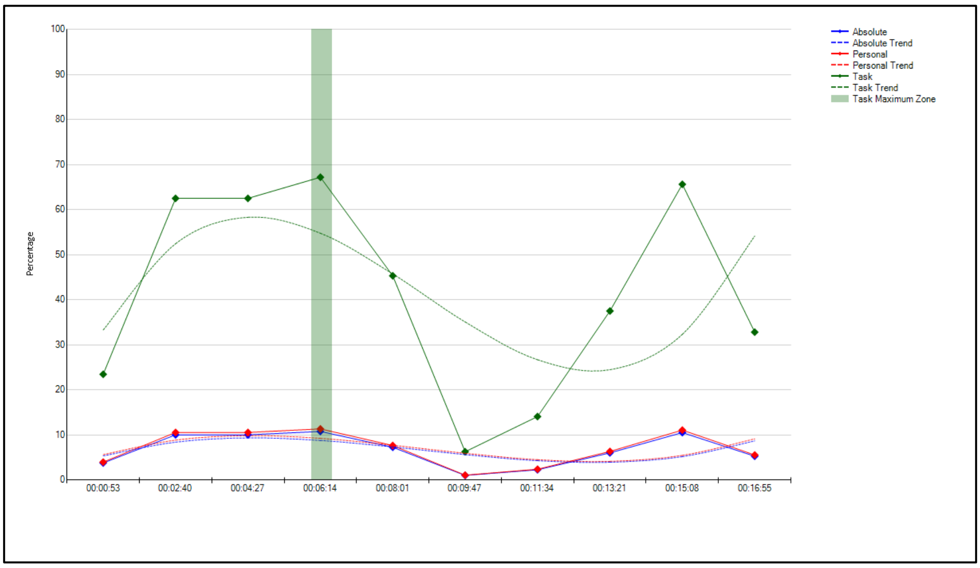
#Revision	Type	Content	Edits	Start	End	Duration	BeginPos	EndPos	Length	Chars	Chars without space	Words
0	Normal Production	A·cientif·	19	01:04.904	01:10.128	00:05.224	0	10	10	8	6	2
1	Deletion	·fit	4	01:15.863	01:17.903	00:02.040	10	6	4	3	2	1
0	Normal Production	tif·drink·a·meid	16	01:26.583	01:53.495	00:26.912	6	22	16	13	10	4
2	Deletion	di	2	01:54.575	01:54.839	00:00.264	22	20	2	2	2	1
0	Normal Production	dicine·,·	9	01:55.503	02:09.503	00:14.000	20	29	9	6	4	1
3	Insertion	ther	9	02:38.782	02:39.478	00:00.696	0	4	4	4	4	1
4	Deletion	reht	11	02:40.470	02:46.878	00:06.408	4	0	4	4	4	1
5	Insertion	There·are·a·dog·and·a·scin	28	02:47.422	03:06.158	00:18.736	0	26	26	20	14	7

The second type within the Revision analysis framework is S-Notation (Kollberg, 1998), which depicts a linear display of the text whilst indicating the breaks in the text and number (i.e. pauses). It also displays where insertions were located, and which parts of the text were deleted. All of these writing events are marked with their timestamp or number of pauses, and a symbol is assigned to both insertions (curled brackets) and deletions (square brackets). Such organization makes the process at ease for the researcher to be able to reconstruct the revision operations in the composition process.

[1] Edits are "the number of actions that the writer needed to complete the action. This also includes shift key, cursor movements, cursor clicks, among others" (Leijten & Van Waes, 2015-2019, p. 89).

- *Token Analysis* is a tool that carries out an analysis of specific words under orthographical constraints. Once the token is presented, which might be a full word or lemma, the analyzer searches for words sharing fully or partially the orthographical components of the targeted word. Interestingly, should any of the words found have undergone a revision operation, this is indicated in the chart generated with an indication of when these revisions occurred and their type.
- *Fluency Analysis* is another function whereby Inputlog organizes the data as per the information obtained on the keypresses, pauses, and the time interval, which may be parameterized to the researcher's interests. Inputlog generates a graph displaying all the information, which very well serves to explain what analyses are performed under this function: (i) the average strokes per minute and their duration, as well as the total number, (ii) absolute maximum, personal maximum and task maximum, all of which are dependent upon the task itself, our maximum regarding other tasks performed in Inputlog, and eventually the task absolute according to the parameter set up by the researcher, and (iii) interval of strokes, for instance, 10 seconds; the total number of strokes and strokes per minute in that interval (Leijten & Van Waes, 2014). Such statistical measures lead to the generation of a graph which intends to represent Fluency:

Figure 14.
Sample of a Fluency graph generated by Inputlog.



- *Bigram Analysis* is a complementary analysis to pause analysis as its main purpose is to provide more fine-grained information on the process. This type of analysis generates statistical data on sequences of bigrams (or two consecutive keystrokes), combined with information on the latency of other keystrokes (Leijten & Van Waes, 2014). As for the measures generated by this analysis, we find the number of observations, mean length of interkey pause, its standard deviation, median length, minimum length, and maximum length, and 90% of interval low and high. Such an analysis consists of a series of categories: bigram categories (12 different), bigram speed, bigram frequency, and an alphabetic and non-alphabetic list of bigrams.
- *Linguistic Analysis* provides the researcher with a very valuable tool that technically extracts the logging file and performs a W-Notation, that is, "a time and revision enriched word aggregated process file at the word level" (Leijten & Van Waes, 2014: 101). In this vein, this Linguistic analysis offers an overview of the final product, S-Notation, W-Notation, the text reconstructed, and insertions and deletions in context. Nonetheless, this type of analysis is not

integrated within Inputlog as it has to be dispatched to a web server, and a further log-in to the page is required. In any event, the linguistic analysis is out of service owing to a server migration issue (Dr. Luuk Van Waes, personal communication, February 19th 2019).

In addition to the analyses presented above, Inputlog offers other types of possibilities, some of which go beyond the scope of this PhD (for a review on how to use Inputlog and other functionalities, see Leijten & Van Waes, 2005, 2013; Van Waes, Leijten & van Weijen, 2009). For instance, one might set the pause threshold at a certain specific parameter or might conduct the different analyses from a different perspective (e.g. revision-based or pause-based). Nevertheless, Inputlog is not devoid of a few drawbacks which, while not downplaying the quality and potential of the KLS tool, are certainly influential over the research process if not carefully tested. For example, the Linguistic analysis, as specified previously, does not work at present for server migration. Likewise, technical problems with Inputlog are reduced to timely issues when generating output of a certain analysis – as it is the case of Linguistic analysis – or using other features such as replay which, on the basis of our personal experience, has not fulfilled its function on several devices whilst on others it has.

2.7.2. The research potential of KSL programs and Inputlog for the field of L2 writing

The debate on writing research has traditionally revolved around investigations undertaken within traditional settings (i.e. pen-and-paper studies), and little has been done on the potential displayed by KLS tools except for the recent decades. Tracing the relevance of writing back to its origins, psycholinguistics has dealt with language production from the perspective of orality at the expense of writing (Spelman Miller, 2005), leading to a wide array of studies looking into speech processes (Butterworth, 1980; Levelt, 1989). The study of writing has been approached through different lenses on the basis of the research focus, one of which has been second language writing. Such an interest in L2 writing has placed greater emphasis on process-oriented research (Spelman Miller, 2005), whose main purpose has been that of unveiling the underlying writing processes.

The advent of computers led to the creation and spread use of computer-based research tools which accounted for these processes more directly and reliably, avoiding any obtrusion into the writing process. KLS tools, framed within these research instruments, allowed the researcher to look into writing processes, which previously seemed invisible to the human eye. Tracing KLS back to its origins, in the 1980s, the first attempts at regarding KLS as a valuable tool for writing research were made (Bridwell & Duin, 1985). As reviewed previously, KLS tools are quite abundant in number despite their different focus for research. Despite that, Inputlog seems to be the tool that gathers the finest conditions under which our research may be conducted. Inputlog as a KLS program offers different writing environments and the necessary setting so that the collection procedure is undertaken with no obtrusion, for instance, when conducting a research study in an educational setting (Leijten & Van Waes, 2013). Precisely, gathering data with a KLS tool is theoretically justified on the grounds that its use may indicate traces of cognitive processes which underlie writing fluency and flow. In this vein, pause and revision behavior hold pivotal roles not only in the construction of an analytical perspective to writing as a non-linear process but also as measurable variables which hint at such underlying cognitive processes behind the composing process (Spelman Miller, 2005). It must not be forgotten, however, that pauses themselves are not to be attributed to a sole operation, as these might be related to different cognitive activities other than revision itself, for example, planning, rereading or inner speech (Baaijen et al., 2012; as cited in Révész, Michel & Lee, 2019). Hence, the importance of analyzing the surrounding textual context, and manually observing the data output of KLS.

Building on the advantageous functionality offered by KLS tools, research on writing processes with the use of this KLS-based methodology has been combined with the techniques traditionally used in an isolated manner (Hyland, 2016). Although the amount of research data that KLS tools gather is considerable, there is a tendency towards a multimethod approach (Leijten & Van Waes, 2013). The following figure (Leijten & Van Waes, 2013) describes the writing observation methods which might be used to analyze writing processes:

Figure 15.
Writing Observation Methods (Leijten & Van Waes, 2013, p. 4).

	Direct research methods	Indirect research methods
Synchronous	Concurrent think aloud protocols Prompted pauses	Keystroke logging Video observation Double task method Eyetracking EVP or fMRI
Asynchronous	Retrospective protocols	Text analysis Versioning

As can be seen above, there are four intertwined and discernable dimensions: (1) synchronous methods are observational techniques that entail data gathering on underlying cognitive processes within the composing process, (2) asynchronous methods are techniques whereby data are collected after the writing session, (3) direct research methods encompass data gathering techniques involving conscious participants' verbalizations towards given sources, (4) indirect research methods refer to inferential data gathering techniques whereby the researcher is expected to draw conclusions upon the information (Leijten & Van Waes, 2013). Such observation methods for the writing process are not without their weaknesses.

In what follows, a combination of KLS with some of these techniques is intended to reveal positive and negative sides of their joint use: KLS and TAP, KLS and Eye-Tracking, and KLS and retrospective protocols. In this way, we will gain deeper insight into the possibilities that the use of both instruments offers for L2 writing research.

(a) KLS and TAP (Think-aloud protocols). The use of TAP as a single tool to obtain data concerning cognitive operations during the composing process has been widely accepted (Beauvais, Olive & Passerault, 2011; Llosa, Beck & Zhao, 2011; Van Weijen, 2008; as cited in Leijten & Van Waes, 2013), but the advent of KLS programs has allowed the possibility of exploiting the supplementary potential on data gathering that the combination of both techniques offers. To trace TAP to its origins, this verbal report technique stemmed from the claims in the field of psychology (Ericsson & Simon, 1993), which suffered an adaptation to its use in writing research (Dam-Jensen & Heine, 2009). TAP is a concurrent verbal technique whose research potential lies in the rich data from the participants' thought-process in an online fashion during the writing process. Nonetheless, TAP has received criticism from three angles: the obtrusiveness into the writing process, its accessibility, and the incomplete data obtained (Bowles, 2018; Dam-Jensen & Heine, 2009). Firstly, the obtrusiveness of TAP alludes to the interference with the writing process by diverting the attentional focus and influencing the sequence of thoughts. Such apparent interference has been questioned in the scholarly community as several studies (see Ericcson & Simon, 1993, Kowal & O'Connell, 1987 or Toury, 1995) have attempted to prove it on the grounds that the human brain is not able to perform two tasks simultaneously.

In addition to the inconsistency of TAP as a fully valid research technique, opponents (e.g. Toury, 1995) have justified it on the grounds that TAP is delaying (Bowles, 2018). Second, as regards its accessibility, tapping into the writer's cognitive operations is fairly indirect as it is not possible to obtain data straight out of the writer's mind. The third criticism is leveled against the incompleteness of the data obtained which, as pointed out by Hansen (2005), depends on the writer's decision as to what to verbalize. It has also been contended that subconscious processes will not be part of such verbalization as only processed operations are voiced (Jääskeläinen, 2000; Hansen, 2005; as cited in Dam-Jensen & Heine, 2009).

All in all, the main suggested drawback of TAP lies in the fact that "the act of thinking aloud may lead to reactivity and thus jeopardize construct validity" (Révész & Michel, 2019, p. 495). Nevertheless, studies such as Roca de Larios et al.'s (2008) acknowledged that such an influence might involve a reactive issue, which in their case, they controlled by giving instructions in Spanish and running a trial, endeavoring to minimize the risks. Research, however, has not yet revealed clearly whether or not the writing process is influenced by TAP.

Research studies where TAP and KLS have been effectively combined have made an appearance in recent decades. By way of example, Schrijver et al. (2012) explored the role of transediting strategies with translators, where participants had to complete a translation assignment of a medicine-related information leaflet into English, a task to be accomplished with concurrent TAP while their writing operations were recorded with Inputlog. It was reported that TAPs and KLS (Inputlog) were successful tools with supplementary aspects to each other, which led to the compensation of data gathering shortcomings. For instance, low-level and high-level processes, which were difficult to unveil by examining the logged data, could be counterbalanced with the information obtained from TAP. In their study, Schrijver et al. (2012) realized that pauses posing difficulty to be interpreted were easily neutralized by the effect of TAP, which gave a clear picture of what the participant was thinking about (word choice, translation...). An earlier attempt at combining both techniques was made by Stevenson et al. (2006), who explored the role held by TAPs together with KLS in order to shed light upon the type of revisions. The use of TAPs was of help for their research since revisions could be scrutinized to the extent of detecting what participants were thinking at the moment. As reviewed by Révész et al. (2019), other attempts were made to combine KLS tools and other techniques to favor construct validity. For example, Révész, Kourtali & Mazgutova (2017) made use of several techniques, including KLS and retrospective interviews, which will be dealt with in more detail in the ensuing paragraphs.

In summary, combining KLS and TAP entails benefits to research as a result of their complementation. Whereas KLS allows the researcher to tap the pausing and revision behavior on the part of the participants, as well as a representation of the process *per se*, TAP adds the contents of participants' reflections to the information holes which have not been fully extracted by KLS data (Leijten & Van Waes, 2013).

(b) KLS and eye tracking. Much as it happened with the advent of computers, which sparked the researchers' interest in developing new tools such as KLS to scrutinize and mine data from writing (Leijten & Van Waes, 2013; Van Waes, Leijten & Neuwirth, 2006), the shortcomings perceived in some of these techniques have given rise to the use of alternative instruments (Latif, 2008), all of which are regarded as cutting-edge. In this vein, eyetracking is a type of technology that records eye movements, measures them according to their location, the sequence of occurrence, and the length of gaze. The advent of eye-tracking acquired relevance in the field of reading research and vocabulary combined with reading research (see Pellicer-Sánchez & Siyanova-Chanturia, 2018). It has been a seemingly useful research tool for many decades (Rayner, 1975), but the use of eyetracking in other fields such as L2 writing is a newly conceived trend that needs even more testing (for example, see Andersson et al., 2006; Chukharev-Hudilainen et al., 2019; Gánem-Gutiérrez & Gilmore, 2018; Wengelin et al., 2009; Wengelin, Leijten & Van Waes, 2010). Along this line, the combination of KLS and eyetracking has provided a new sense of perspective in interpreting the fixation of the eye while writing a test (Leijten & Van Waes, 2013).

We will proceed to analyze the advantages and disadvantages of eyetracking techniques. First, eyetracking shares a common advantage with KLS programs, which is that it is unobtrusive - though physically obtrusive - and "support ecological validity in L2 writing research by not interfering with the composing process" (Gánem-Guitérrez & Gilmore, 2018, p. 7). This non-interfering feature allows for a greater fine-tuning of the data by being able to analyze the occurrence of events with respect to the data processed with KLS. Second, it allows for analyzing "the interaction between reading and writing" (Leijten & Van Waes, 2013, p. 13), certainly aiding the identification of pausing and revision behavior.

Eyetracking entails a major drawback to the study of composing processes, which is associated with its technical functionality. Under this assumption, eyetracking analysis software is precisely targeted at reading and visual perception studies. To put this more clearly, this eyetracking software performs analyses whose main focus is placed on so-called areas of interest, chiefly on the basis of static images. Nevertheless, such an analysis is not functional in writing research given the dynamic nature of the composing process, which progresses as the writer types (Chukharev-Hudilainen et al., 2019). This drawback was highlighted very conveniently by Leijten & Van Waes (2013), who reviewed the possibilities afforded by Inputlog in combination with eyetracking technology. In this vein, it was reported that, while reading has predominantly featured small units of text statically, scarce research had been conducted on global text processing. One reason for this may be the controversy revolving around the measures used to tap into such a process (Hyöna & Lorch, 2004). Another drawback is the high-priced cost of eyetracking software and hardware, which not only hinders the development of research studies using such data collection procedures. It also reduces the scope of application by constraining it to laboratory research at the expense of naturalistic environments (Chukharev-Hudilainen et al., 2019).

From the research point of view, eyetracking may be regarded as a supplementary tool if used with KLS for several reasons. For instance, KLS allows for a more visual view of the writing process by signaling those parts of the text production dynamism that have to do with both the temporal dimension and the writing process itself, along with the behavior exhibited in terms of cognitive processes (Révész & Michel, 2019). Despite the drawbacks exposed in the previous paragraph, research has overcome the difficulties and developments are being fostered (Wengelin et al., 2009), where the merging of KLS and eyetracking technology are under scrutiny as a way of unveiling the processes of reading and rereading (Leijten & Van Waes, 2013).

Nonetheless, most of the eyetracking-based studies conducted thus far have directed their attention toward the relevance of reading in the translation process rather than in the writing process itself (see for example Dem-Jensen & Heine, 2013; Dragsted & Carl, 2013; Jakobsen, 2011). The research focus has not permanently remained in translation studies, and its scope has been extended to the field of L1 writing with pioneering studies such as de Smet et al. (2018), in which they explored the role of reading as a process during writing in the hope of revealing whether fluent or nonfluent writing exerted any influence over reading behavior. In another recent study by Révész et al. (2017), the researchers looked into the pause behavior within word, clause, sentence, or paragraph level and the writers' eye fixation during these pauses. Similarly, they combined eyetracking with KLS and retrospective interviews, following the research design premises under which later studies would be conducted. Gánem-Gutiérrez & Gilmore's (2018) latest study also sheds light upon how the composing process evolves in real-time with 22 Japanese university L2 learners, with a data collection procedure that involved digital screen capture and eyetracking technology, but no KLS was used. More recently, L2 writing studies have relied upon eyetracking as a joint technique with KLS (see Révész & Michel, 2019 for a state-of-the-art), among which studies are dealing with the cognitive processes according to pause behavior and the levels of revision such as Révész et al. (2019). They used KLS in combination with eyetracking and retrospective interviews with 30 advanced Chinese L2 English learners with a writing task. The use of eyetracking favored the understanding of pauses during revisions in consonance with the act of revisions themselves, indicating that writers showed a tendency towards revising or gazing at the text before applying any modifications.

All things considered, the combined use of eyetracking with KLS is a tool that might provide with the information holes that the latter data gathering technique does not guarantee. In light of this, eyetracking is advantageous in that it offers direct information of the position to which the writer's eyes were fixated, thus paving the way for a deeper analysis into the pause behavior in relation to revisions. By triangulating both procedures, the researcher can undoubtedly obtain greater insight into the way in which the composing process was developed, and the relevance of time length at all levels (eye movements and key strokes) for the process itself and the different cognitive operations associated with it (planning, formulating, pausing and revising).

Notwithstanding the potential exhibited by eyetracking for research on writing processes, its implementation in the research design for this PhD has been disregarded for a number of reasons. First, eyetracking might allow for a broader overview of the writing process owing to the supplementary information that it furnishes. However, the use of eyetracking with children has been reportedly rejected insofar as young participants are unease as a consequence of wearing the eyetracking helmet which enables the online recording of eye fixations. Such physical discomfort is likely to divert young participants' attention away from the main purposes of the research, adding to this view the exclusion of children with eye problems and visual acuity (myopia, astigmatism, hyperopia, among others) and thus reducing the scope of the sample size (Dam-Jensen & Heine, 2009). At the technical level, the implementation of eyetracking raises concerns as regards the high cost of eyetracking devices, together with the impossibility of setting up the experimental tools at the school – in order to ensure ecological validity – which would force to locate the data collection procedure in an experimental laboratory.

(c) KLS and stimulated recalls (*retrospective verbal protocols*). Together with TAP, stimulated recall (also retrospective verbal protocols) are reported to be two of the most used data gathering methods used in research on L2 writing (Latif, 2009, 2018). In comparison with TAP, which is an online verbal report, retrospective interviews and stimulated recalls to take place at an offline time. That aspect contributes to adding complementary information to the operations in the writing process which are not visible to the researchers' eye by obtaining raw data (for instance, understanding whether a pause is related to the operation of planning, or it is rather part of a revision process). In this regard, retrospective interviews entail relating the process towards the product (final text) by recalling the different cognitive operations. Retrospectively, stimulated recall entices the participants to reflect upon their writing process by concentrating on specific features tailored to the research interests. Nonetheless, the use of this verbal report technique is not without criticism, as the roles of memory, recognition of actions, and retrieval are accounted for as essential (Hansen, 2005; as cited in Dam-Jensen & Heine, 2009). Probably, one of the major issues might be related to memory, as participants may not recall very vividly how the composing process underwent (Levy, Marek & Lea, 1996), an aspect which is to be counteracted if a retrospective interview or stimulated recall is conducted in the aftermath of writing the text with a maximum timelapse of 48 hours (Krings, 2005; as cited in Dam-Jensen & Heine, 2009). In fact, the length of the delay between the writing session and the stimulated recall session results in an increase of memory decay, hence the limitation to recalling the events occurred at that session (Gass & Mackey, 2000). Much as TAP, retrospective interviews are largely reliant on the participants' account, which increases the risk of over-reporting (Harwood, 2009), or even providing a distorting view of the process as a consequence of their ability (Smagorinsky, 1994). To counteract this, the use of KLS tool as a support for the stimulated recall aids in balancing the holes existing in the data. Similarly, it contributes to decreasing the effects of lack of memory or over-reporting (Révész et al., 2019) using a replay of the whole writing process. Such a playback is one of the main functionalities offered by KLS tools, for instance, Inputlog, which allows for fine-tuning the replay to our research needs (Leijten & Van Waes, 2011).

Notwithstanding the potential pitfalls which may derive from its application, there are many benefits associated with its use. By way of example, Lindgren & Sullivan (2003) concluded that the use of KLS in conjunction with stimulated recall favored noticing and language awareness as a drive for more text revision. Building on the authors' suggestion that "this case study [...] needs to be followed up with studies [...] before any general conclusion can be drawn" (Lindgren & Sullivan, 2003, p. 185) other studies have emerged on this account, where retrospective interviews were used as a data-gathering technique in an individual manner (see Khuder & Harwood, 2015; Révész, Kourtali et al. 2017; Révész, Michel et al., 2019).

Bearing in mind that KLS data are not precise regarding the reasons underlying pauses, that is, whether a pause or action time is related to planning, formulation, revision, or other related processes, their classification may only be achieved by engaging in a fine-grained analysis that takes into account the surrounding textual context. Such operations may only be unveiled on the grounds of pause thresholds (Van Waes & Leijten, 2005), but one may incur the risk of misunderstanding the real purpose behind such pauses. In order to pave the way for a greater comprehension of the writing mechanisms, stimulated recall strives to be an elucidating tool that fine-grains the data obtained to prevent information loss. Given the necessity of enquiring participants about the many operations and writing strategies used in the composing process, researchers can unveil those processes for which raw data do not provide sufficient tangible evidence of the process itself. Thus, the replay of the writing process allows for the exploitation of the affordances by both KLS data and stimulated recall (Révész, Michel & Lee, 2019). In fact, replaying the text on a computer screen draws the attention of the participants as well as motivates the purpose of the stimulated recall by energizing the struggle to seek memories associated with this specific writing event (Sullivan & Lindgren, 2002).

Stimulated recall in L2 writing has been used as a supplementary technique with screen recording, interviews, or pause analysis (Bosher, 1998). Likewise, stimulated recall is a technique that is not solely centered on showing the participant a replay of his writing process while awaiting a reflection or reaction on his behalf. It might also entail questioning the participants about specific behaviors at some points of the composing process, or encouraging them to reflect upon changes at letter, word, sentence, or paragraph level. There are several possibilities regarding the use of stimulated recall: (i) a replay of their writing process whilst leaving sufficient time to reflect upon it without the researcher's intervention, (ii) a replay of the writing process but focusing on very specific parts, allowing the researcher to ask questions about pauses, i.e. to ascertain if an accumulation of pauses prior to producing text is due to operations such as planning or, revision, for instance, (iii) the omission of a replay would somewhat distort the nature of stimulated recall *per se*. However, it is a meaningful choice depending on the characteristics of the research. For instance, not using a replay may require that the researcher deepens into what the *Inputlog* data indicates. This is a necessity for the stimulated recall insomuch as a series of very specific questions regarding the writing process have to be framed in an attempt to unveil reasons behind some pausological behavior linked to underyling cognitive processes. To illustrate this, participants might be asked for details about why they took numerous pauses at a certain time of the writing process (without necessarily having to resort to the replay), thus elucidating and fine-graining the data processed by KLS tools.

Validity issues are especially relevant in stimulated recall research despite its apparent neglect in literature reviews (Gass & Mackey, 2016). Memory decay is one of the most recurrent drawbacks of verbal report techniques, which might be partially neutralized in stimulated recall via stimulus provision (Bowles, 2018). Providing that this stimulated recall is combined with appropriate inquiries on the researcher's part, the issue of memory decay is minimized. Nonetheless, as was mentioned previously, it is recommended that the time-span between the stimulated recall and task completion does not lapse more than 48 hours (Bloom, 1954; Henderson & Tallman, 2006), together with the necessity of following carefully the guidelines on stimulated recall design (Bowles, 2018; Gass & Mackey, 2000). Along this line, this offline verbal report should be implemented after the first task completion session. This procedure would ideally provide a picture of the cognitive processes and writing strategies underlying the participant's writing process (Bitchener & Storch, 2016). Thus, this verbalization will result in a reactivity effect on the participant. Here the idea of adding questions to simulated recall is reinforced since it will prevent the participant from dwelling on aspects forcibly irrelevant to research. Scientific literature has explored the effect of recalls when these are provided immediately after the task completion, and when it is provided retrospectively (see Egi, 2007, 2008).

Several conclusions may be drawn from the previous review: (i) the invisibility of cognitive processes to the researcher's eye is dually neutralized: KLS tools such as Inputlog allow for a classification of these processes, and attempt to provide a clearer picture of how the composition process was developed; also, stimulated recall conceivably guarantees that the inaccessible information out of these processes is carefully unveiled by showing the process itself to the writers, and enquiring them about events whose classification is deemed unclear, (ii) stimulated recall entails the provision of a stimulus to the writer as a way of recalling events and providing reasons supporting them; in light of this, stimulated recall might be combined with guidelines from retrospective interviews as a way of addressing shortcomings of one type through the strengths of another (Gass & Mackey, 2016). This widens the scope of action by clarifying the reasons underlying some events in the composition process in which a clear conclusion may not have been reached with the sole data from KLS tools.

2.7.3. Pausological behavior as a trace of writing processes in KLS programs

Pauses are, as evidenced in the previous sections, the basis of the usefulness of KLS tools for the study of writing processes for several reasons: (1) pauses are at the core of the composition process, amounting to nearly three-quarters of the composition time (Alamargot et al., 2007); (2) pauses are quantifiable and observable, allowing the researcher to determine pause location in time and place, and the duration (Schilperoord, 1996); (3) they are thought to cover certain cognitive operations associated with writing processes for which L2 writing research has endeavored to set patterns for their interpretation (Barkaoui, 2019).

The relevance held by pauses in writing is directly related to speech production research (Spelman Miller, 2006a) which paved the way for the study of writing production and pauses. Speech production is filled with hesitance, silent pauses, and false starts, thus indicative of certain cognitive operations underlying each of these pauses (Barkaoui, 2019). In line with this, pauses in writing have allowed for examining the potential reasons which underlie such halts, being those associated with several pre-writing or writing operations, such as a conceptualization of the whole task, goal outlining, planning before writing, and revisions, including lexical retrieval (Schilperoord, 1996; Spelman Miller, 2005, 2006; Wengelin, 2007).

As dynamic an activity as writing is, pausing behavior and its relevance to study cognitive processes is suggested to be reported by different cognitive models of writing (see Hayes & Flower, 1981; Hayes, 2006, 2012; Kellogg, 1990, 1996, 2008; Torrance & Galbraith, 2006). When engaging in writing, pauses might occur as a consequence of the interplay between several operations (Kormos, 2012; Torrance & Galbraith, 2006) given that writing is a complex activity where several cognitive processes are coordinated. Thus, pauses may be the cause of transitioning from one operation (e.g. formulation) to another one (e.g. planning). This transition is supposed to create more cognitive demands on the writer. In other words, the activation of such processes leads to the deactivation of the concurrent ones in such a way that the learner has to accommodate the other by way of pausing (Xu & Qi, 2017). This course of action is particularly relevant in the case of planning and revision, as these both are operations where pauses remain as traces of higher cognitive processes. Pausing behavior is also associated with the writer's intentions since pausing behavior may be intentional or constrained by the writer's cognitive capacities (Alamargot et al., 2007). Intentional pauses might also be referred to as strategic pauses which writers opt for to halt the concurrent process and engage in another one; in the case of pauses as a direct cause of limited cognitive capacities, L2 writers may not be able to handle two operations simultaneously, i.e. planning and formulating at the same time. Such categorization suggests that pauses might as well be occurring due to factors closely related to internal variables. In this regard, the existence of pauses by a limited cognitive capacity is indicative of restrained cognitive use for several reasons, for instance, writers are young learners, and their demands on the central executive may be overloaded and beyond their available resources (Olive & Cislaru, 2015).

Nonetheless, controversy has aroused concerning the evidence provided by these pauses, all of which are reported to be indirect evidence of the development of the composing process, leading to inferring what operations might be underlying them (Spelman Miller, 2006; Wengelin, 2006; Matsuhashi, 1981). Despite the apparent difficulty in detecting what the real reasons behind pauses are, this has not prevented researchers from finding a pattern of behavior and cognitive capacities therein. According to Alamargot et al. (2007), four assumptions have been taken on average within empirical research on pausing behavior and their strategies to detect which cognitive processes are underlying. The first has to do with *pause duration*, as it undoubtedly reveals traces of how complex a cognitive process was. The second assumption entails *pause location*, which hints at where the pause took place, thus tightening the grip on the reasons behind such halts. Third, *pause behavior* relates specifically to the part of the text written afterward. Finally, writing pauses occur when the central executive is overloaded by its inability to juggle both demanding processes and graphomotor execution.

Building on the above, L1 and L2 writing research has tried to shed light on the implications of these concepts in the writing process (see Medimorec & Risko, 2017; Révész, Kourtali & Mazgutova, 2017; Van Waes & Leijten, 2015; Wengelin, 2006; Xu & Qi, 2017), and conclusions have been drawn from the fact that pauses are supposedly not arbitrary. Some pauses have been genuinely identified as fairly regular at some locations within the structure of the text (e.g. at sentence or paragraph boundaries). As mentioned above, pause variation has a strong influence over the development of the composing process, and this variation is highly dependent upon other variables – for instance, typing skills, text boundaries, and eventually text genre (Medimorec & Risko, 2017). The latter gains traction in the framework of this PhD since the writing task, i.e. a picture-based story, clearly defines the route writers have to follow. Thus, the more information given on what to write as well as the text genre *per se* demarcates the difficulty, and thus pause variation. To illustrate this idea, writing a formal letter does not require the same amount of attention as writing a picture-based story. This one may lead to pausing much less on account of the task conventions and task complexity (McArthur & Graham, 2016).

The conceptualization of pauses as indicators of cognitive processes has not been without controversy when setting up a threshold as to defining what a pause is and its duration (Wengelin, 2007), an aspect dealt with by several studies. Thus, defining pauses in writing is restricted to periods that transcend the necessary time for typing, and in which writing is not produced. A consensus has been adopted after several studies (e.g. Alves et al., 2007 or Wengelin et al., 2009) which have established the pause threshold at 2000 milliseconds. Van Waes & Leijten (2015) indicate that a smaller or longer pause threshold may indicate different types of cognitive processes. For instance, small pause thresholds (e.g. between 1000 and 2000 ms) may indicate low cognitive processes. Larger thresholds (over 2000 ms) may indicate higher cognitive processes and have been extensively used in the scholarly literature to capture cognitive processes such as planning, revision, or formulation, among others (Olive et al., 2009). Likewise, it has been the "most used threshold level in writing research to define the pauses" (Tiryakioglu et al., 2019, p. 218). Previous empirical research, however, has been conducted at variable pause thresholds, some of which were set at from 300 to 2000 milliseconds (e.g. Chukharev-Hudilainen, 2014; Medimorec & Risko, 2017; Olive & Kellogg, 2002). Thus, the use of a certain pause threshold is appropriately justified by literature (Wengelin, 2006; Chenu et al., 2004), as pauses below 2000 milliseconds are thought to pertain to "simple mechanisms of typing" (Medimorec & Risko, 2017: 4). Additionally, pauses below 200 ms are related to "motoric interkey-transitions" (Van Waes & Leijten, 2015, p. 86).

Once pause thresholds have been theoretically justified, I will review how pause studies have paved the way for analyzing pause frequency and duration. Very recently, KLS programs such as Inputlog (Van Weijen & Leijten, 2005) have widened the scope of research actions by providing data on these aspects without obstruction. In light of this, previous studies have accounted for the importance of pause frequency among writing processes and time allocation - which goes hand in hand with the writer's distribution of time across the task. As such, pauses are scattered throughout the time provided for the writing task. It must be noted, however, that studies have brought to light a number of variations in pause frequency and duration across the composing time (see Roca de Larios et al., 2008; Van Weijen et al., 2008), and sometimes other variables are held responsible for this variation. For instance, pause frequency and its allocation throughout the writing process have been reported to be steady in the case of low-proficient L2 writers while high-proficient L2 writers have shown more dynamism across the writing process (Roca de Larios et al., 2008). Empirical evidence then suggests that more focus should be placed on *when* these pauses occur since it helps uncover the dynamics behind the evolving and situation-changing writing process (Breetvelt et al., 1994; Roca de Larios et al., 2008; Van Weijen et al., 2008). In the same line, several studies have provided firm evidence that internal and contextual factors contribute to altering time allocation of writing processes, and hence pauses associated with them (Medimorec & Risko, 2017; Révész et al., 2017; Xu & Qi, 2017). To illustrate this, it has been proved that writers, as per their L2 proficiency, resort to pausing at some locations in a more frequent way than at others. As such, Dean & Zhang (2015) found out that low L2 proficient writers encountered greater difficulty in producing an efficient text, which appears to be motivated by constraints on spelling, lexical retrieval, morphosyntactic search, or ideas generation (i.e. planning operations). Within such a context, other studies (Sasaki, 2000; Xu & Qi, 2017) have reported that low-proficient L2 writers are likely to pause more frequently before "thematic episodic boundaries" (Barkaoui, 2019, p. 532), which is indicative of, for instance, local planning if they employ a “what next strategy”. Similarly, this suggestion is backed up by evidence (Xu & Qi, 2017) which supports a pausing pattern in the case of low-proficient L2 writers, whose pauses were shorter but more recurrent in the first interval of the composing process; thus, while pause frequency decreased in the second interval, pause duration escalated. Xu & Qi (2017) put forward that such a modification of the pausing pattern reveals a close association between longer pauses and formulation processes where ideas are materialized, and thus not restrained from difficulty or doubts such as word searches or the generation of ideational content.

It has been mentioned previously that task complexity might be regarded as one of the few variables influencing the writing process and the pausing behavior associated therewith (McArthur & Graham, 2016). In this regard, numerous studies have confirmed a correlation between task complexity and L2 writers' behavior (Révész et al., 2017; Spelman Miller, 2000), the findings of which have shed light on the use of content support (for instance, the information provided in task instructions). These studies have also shown that fewer pauses were produced at sentence boundaries and more revisions took place at word or morphosyntactic boundaries, pushing writers to focus their attention on enriching their linguistic structures.

Another variable affecting fluency, revision but more importantly, pausing behavior is the writing environment chosen for writing. This PhD is framed within the scope of studies in digital environments (or computer writing), and I will focus on research conducted under such conditions. Scarce have been the studies in L2 contexts in which the effects of pen-and-paper against computer writing were compared, or where the effects of typing skills were analyzed (cf. Barkaoui, 2015, 2016; Lee, 2004; Phinney & Khouri, 1993). Among the many findings, these studies reported significant results as regards pausing and revision behavior when writing on a computer. For instance, L2 writers' pauses were longer under this condition than when engaging in pen-and-paper writing, while little time was devoted to planning at the initial stages. Additionally, L2 writers' typing skills have also set the stage for other studies (more specifically, Barkaoui, 2016 and Phinney & Khouri, 1993) which provided evidence for the relevance of such computer writing skills for the allocation of pausing across the writing process. Related to that, studies analyzing typing skills in L1 (Alves et al., 2007) have also reported important findings as to pausing patterns between skilled and unskilled typists. The latter's texts were considerably shorter, and composing time was longer. A consequence of this may be attributed to the inefficacy of unskilled typists for producing text, and materializing it by typing on a concurrent basis. This causes the overload of the capacity of the central executive whilst mixing both mechanics and spelling, and L2 conventions (Kellogg et al., 2013). Interestingly, it was also found out that slow typists' texts were affected by ineffective typing skills in that content was not properly interwoven and thus planning operations were required more frequently. Such findings are supported by other studies (see Barkaoui, 2019) where results brought to light that low L2 proficiency writers with low typing skills displayed shorter execution periods with numerous pauses along the composing process. Barkauoi (2019) hypothesizes that the large number of pauses likely devoted to high-level writing processes, and execution periods (shorter pauses) to low-level processes, might be associated with L2 writing ability. In this respect, he alludes to a potential overlap between typing skills and L2 writing ability, whose development may occur simultaneously.

Finally, I cannot conclude this section without making appropriate reference to the challenges and difficulties encountered when analyzing pauses by using KLS software. As was mentioned previously, pauses in writing may only be analyzed by adopting an inferential perspective, and thus interpreting them is subject to the appropriateness of each empirical study (Barkaoui, 2019). Such controversy has stretched to the establishment of a pause threshold which, as previously anticipated, has been a matter of contention in recent studies (Van Waes & Leijten, 2015; Wengelin et al., 2009) with some even making use of several thresholds (Medimorec & Risko, 2017). Others have endeavored to provide strategical tools to define such pause threshold on the basis of the needs of our research study (Chenu et al., 2014) by presenting (1) a temporally-driven approach and (2) a linguistically-driven approach. The first is constrained to a pause threshold and requires observing pause frequency and location throughout the composing process. The second entails the division of the text into linguistic units at micro and macro levels (i.e. lexical, phrasal, sentential, clausal, paragraph, or textual level), and observing when and how long these pauses occurred. Such approaches and a pause threshold are hard cores of any research study framed within fluency, pausing and revision behavior, and the use of KLS as a data gathering tool. In this vein, the challenge of a pause threshold has been continually debated and effortfully responded to by setting the pause threshold from 1000 or 2000 milliseconds onwards as a valid measure choice (Van Waes & Leijten, 2015). Nonetheless, the variety of pause thresholds used in the scholarly literature has reduced the scope to the findings that may be compared. Ideally, using a linguistically-driven approach would use up much more resources on the part of the researcher as it would constitute a dull task to perform. Studies have suggested, however, that a mixed perspective is applied (Chenu et al., 2014), that is, adopting a temporally-driven approach that takes the best of the linguistically-driven proposal.

A troublesome challenge in pausing research (and thus research on writing processes) is closely tied to interpreting pauses and locating them under the adequate writing sub-process. While pause location does allow for inferences, the underlying cognitive operation results in an arduous task for the researcher, all of which might be owed to physical or motor, socio-psychological (presumably, learner's internal variables) or cognitive causes (Alves et al., 2007; Wengelin, 2006). Beyond this clarification, should pauses be related to cognitive processes, it is a complex task to determine or itemize the different writing subprocesses (for instance, whether a pause refers to online planning and it was a lexical retrieval, or it was precisely part of a formulation process). Also, a single pause may just refer to a foregoing activity (Wengelin, 2006; Chenu et al., 2014). To overcome such difficulty in interpreting pauses and appropriately allotting them the corresponding process or sub-process, Barkaoui (2019) reviewed other studies (cf. Baaijen et al., 2012; Spelman Miller, 2006; Alamargot et al., 2007; Hayes & Chenoweth, 2006; Leijten & Van Waes, 2013; Wengelin, 2006) and concluded that numerous strategies had been mapped out. He distinguishes three: (1) interpreting pauses based on text generation (i.e. text as a dynamic entity) and revision operations undertaken to edit it; (2) paying close attention to intervals between sequential pauses as well as variations in the pace of writing (i.e. looking at writing fluency hints at the existence or absence of processes occupying the central executive); (3) resorting to other data collection procedures which can be combined with KLS data gathering techniques, and whose main aim is to shed light on those pauses (and hence cognitive processes) which are not disclosable by merely inferring (see 2.6.1. for a detailed review of KLS in combination with other data gathering techniques).

2.8. Chapter summary

In this chapter, we presented the different theoretical underpinnings along with empirical research in the field of L2 writing, and subsequently WCF. Regarding the views on L2 writing in this dissertation, we have highlighted the research value of the WL strand, and the implications of writing as a site for language learning with children's writers as mediated by WCF. Children, as limited capacity processors, are more cognitively loaded when engaging in deeper processing and hence their attention to meaning is reduced. As reviewed, extensive empirical research (e.g. Coyle & Roca de Larios, 2021) has attempted to demonstrate this preconized learning value of WCF by testing its effectiveness on adult populations, and more scarcely on children informants. Some research endeavors with young writers have been recently focused on the learning potential of models as WCF (e.g. Cánovas et al., 2015; Coyle et al., 208; Coyle & Canovas, 2019), and its focus has been on product and process-oriented research with studies dealing with the trajectories or routes that children follow over a cycle of time, elucidating the linguistic modifications applied as mediated by WCF (e.g. Cánovas, 2017; Coyle et al., 2019). Among the main characteristics of empirical research with models, we have observed that they have highlighted the value of collaborative writing mostly in a pen-and-paper environment. In line with recent empirical studies on the use of models, the main findings of research with children have attested that the main affordances of this type of feedback are based on (i) the rich meaning-focused nature of models as input and (ii) the evidence that young writers deplete their attentional resources on lexical chunks and ideational content. Nonetheless, those studies have failed to account for the role of individual writing and model texts as WCF, and thus, more empirical evidence is needed in this regard since the extent to which young children may make use of model texts as a feedback technique is yet unknown. Similarly, L2 writing is an individual activity, especially in school settings where learners perform their writing tasks on their own (e.g. when they have to hand in a school assignment). Hence, the necessity to explore the potential effect of model texts on individual young learners' writing, and how these writing processes develop.

Likewise, our review has revealed that a bulk of research has focused on L2 writing processes, with variables including the type of task (e.g. Révész et al., 2017; Barkaoui, 2019), pausological behavior (e.g. Révész et al., 2019), L2 proficiency (e.g. Tiryakioglu et al., 2019), keyboarding skills (e.g. Barkaoui, 2016) or text quality (e.g. Medimorec & Risko, 2017), mainly with an adult population. Of special interest to us is that no studies to date have investigated the impact of feedback processing on the writing processes, especially with the influence of models as WCF. This is an important gap because the provision of feedback may condition the strategic allocation of writing processes, and the amount of time devoted to each of them. Furthermore, recent methodological trends on investigating writing processes have resorted to keystroke-logging tools, a computer-mediated software to tap into cognitive and writing processes. A considerable body of research has shown a series of research benefits regarding these tools such as its unobtrusiveness for looking into these processes and pausological behavior. Nonetheless, another important gap has been the scarcity of studies where KLS is used with an underrepresented population as children are.

Building on all of the above, our literature review has revealed that: (a) among the substantial amount of recent research on models as WCF and both previous and recent empirical studies on writing processes, there is a research niche in understanding how L2 writers make use of feedback and how certain writing processes (i.e. planning, formulation, and revision), as well as pausological behavior, are influenced by the provision of model texts as WCF, (b) despite the great amount of both cross-sectional and longitudinal studies on models discussed throughout this chapter, none of them has explored how model texts might contribute to L2 writing research with young EFL writers producing texts individually, (c) the language affordances of models as WCF are mirrored in how learners noticed and attended to certain language features and ideas. These concepts of noticing and attention are connected with what the different models of writing processes have explained (see Bereiter & Scardamalia, 1987; Cummings, 1989; Hayes, 2012; Kellogg et al., 2013), e.g. the strategic allocation of attentional resources, and hence the relevance of looking into how these processes are affected by the influence of feedback appropriation, and

(d) there is a need for new methodological calls to collect data on writing processes and pausological behavior, which include the use of KLS without resorting to obtrusive methods such as TAP.

CHAPTER 3. AIMS AND RESEARCH QUESTIONS

3.1. Ultimate aim of the thesis and focus of the study

This doctoral dissertation is motivated by two aims. Firstly, we intend to contribute to the renewed interest in the study of L2 writing processes from an SLA perspective (cf. Gánem-Gutiérrez & Gilmore, 2018; Révész & Michel, 2019). More specifically, we aspire to shed light on the temporal distribution of young writers' writing processes and pausological behavior. This is an important gap in the scientific literature since, to my knowledge, no studies have been conducted on the writing process, and pausological behavior with children, an underrepresented population in this type of studies. The rationale behind the study of pausological behavior is rooted in the fact that it would certainly be "interesting to explore [pausing behavior] given that these phenomena co-occur during the writing process" (Barkaoui, 2019, p. 628). Additionally, this study intends to shed light on and provide empirical evidence of how young writers' writing processes and pausological behavior develop in a digital environment. More and more studies are being conducted in a digital environment justifying the global recognition of writing in these settings (e.g. Asker-Árnason et al., 2009; Miller et al., 2008, among others), but few have focused on writing processes and pausological behavior (see Hafner et al., 2020). Thus, our study will add a new contribution to how writing is developed in a digital setting with a population scarcely studied under such conditions. Accordingly, we have intended to attend methodological calls (see Lindgren & Sullivan, 2019) to expand the range of data collection methods in the domain by using keystroke-logging software to collect the data produced in the writing process. The use of such a software tool is empirically appropriate to identify both the different subprocesses of writing and the pausological behavior involved (cf. Barkaoui, 2019; Leijten & Van Waes, 2013; Révész et al., 2019). The second primary research aim intends to observe whether and how young writers' writing processes and pausological behavior may change as a result of WCF processing and self-editing prior to the rewriting of the same text. Also, our focus is on the use of models as a specific written corrective feedback technique, a very common pedagogical practice, which has been widely utilized with both adult and young learners in pen-and-paper writing environments. However, their role in computer-mediated writing settings is virtually non-existent. The question whether and to what extent the writing processes and pausological behavior are altered by the effect of model texts as WCF still awaits an empirical answer. Additionally, model texts have been widely used with children, and they are said to foster noticing of lexically-driven features of texts along with ideas generation (see Cánovas et al., 2015; Coyle & Roca de Larios, 2020; García-Mayo & Labandibar, 2017; Hanaoka, 2007; Martínez Esteban & Roca de Larios, 2010; Young Kang, 2020). Model texts have been reported to promote deeper reflection, which may lead to more language learning in contrast to explicit WCF. They also constitute a more manageable resource for the teacher to use than traditional and time-consuming forms of feedback such as direct error correction. In conclusion, we intend to contribute to the renewed interest in the study of L2 writing processes from an SLA perspective (cf. Gánem-Gutiérrez & Gilmore, 2018; Révész & Michel, 2019).

3.2. Research questions guiding the study

RQ 1. To what extent are there differences in young EFL learners' temporal distribution of writing processes (i.e. planning, formulation, and revision) when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?

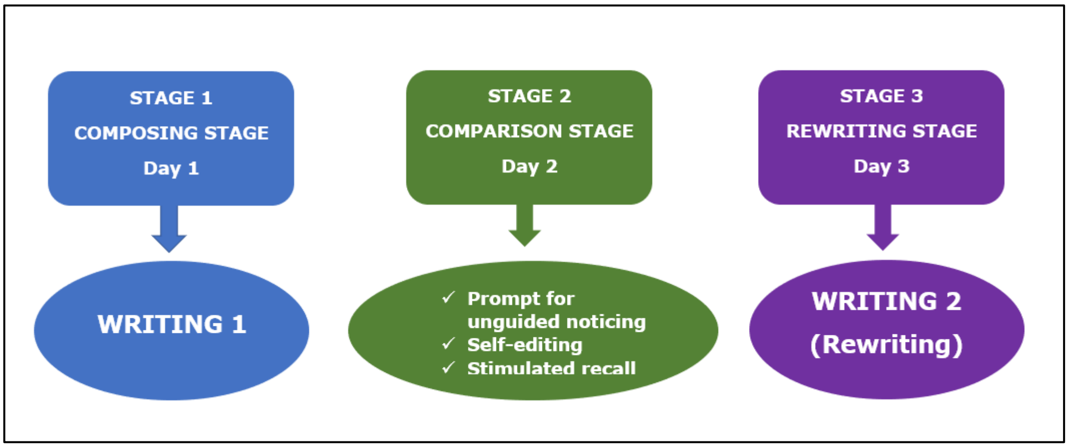
RQ 2. To what extent are there differences in young EFL learners' pausological behaviour when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?

Chapter 4. Method

4.1. Research design

Our study is an example of what DeKeyser and Prieto-Botana (2019) call classroom-based laboratory experimental studies, which they define as those conducted with "learners from regular classrooms who were taken to a laboratory context for a specific treatment that was separate from their regular classroom instruction" (p. 3). We opted for an experimental design as the most appropriate one to observe potential effects of our treatment (provision of WCF). Following convenience sampling, an intact group of ten-year-old children from a Primary school classroom participated in the study. They were randomly assigned to two different groups: (1) feedback group (experimental group. Henceforth, EG), and (2) no feedback group (Control group. Henceforth, CG). Random assignment ensured that each participant had equal chances of their inclusion in each condition, decreasing potential pre-treatment differences between or within-groups (Loewen & Plonsky, 2016).

Figure 16.
Research design.



Our study was composed of three stages with a series of instruments, as will be fully described below in "Data collection procedures".

As shown in Figure 1, in Stage 1 (Day 1) all groups wrote a story (Writing 1 in Figure 1) based on a picture-based prompt (see Appendix 1). The participants wrote their texts individually on a computer using Inputlog, a logging software that captured all their keystrokes, mouse movements and pausological behavior. Stage 2 (day 2) constitutes the treatment. The EG was provided with the texts they had written in Stage 1 and feedback in the form of models, whereas the CG received only their texts but no feedback. Additionally, the EG was given a prompt for feedback processing, whereas the CG group was given a prompt for self-editing. Finally, in Stage 3 (day 3), experimental and control groups rewrote their texts produced in Stage 1 (Writing 2 in Figure 16), reproducing the same writing conditions as in Stage 1.

Based on the above, and as will be further explained in "Data analysis procedures", our research design included one independent variable: 'WCF' with two levels 'absence' or 'presence'. The dependent variables are 'the temporal dimension of cognitive processes during the composition process' and 'the pausological behavior'. The first dependent variable is operationalized as engagement in the three major writing processes (i.e. planning, formulation and revision) and, when indicated, subprocesses. The second dependent variable is operationalized in terms of pausing measures as provided by the logging software, namely pause frequency, pause duration and pause location. We also collected information on the students' perceptions of their own typing skills as a potential confounding variable within our study. This was supplemented with the analysis of the participants' writing transcription fluency, that is, the strokes produced per minute and the interkeystroke interval in milliseconds (as done in studies such as Barkaoui, 2016, 2019 or Medimorec & Risko, 2017).

4.2. Context and participants

►Context

The present study was conducted in a semi-private school partially funded by the Regional government in a district close to the city center of Murcia, in Southeast Spain. Early Childhood, Primary, and Compulsory Secondary Education are taught at the school. The school became part of the Bilingual program implemented in the Region of Murcia. This program was implemented gradually throughout Primary and Secondary education stages. In Early Childhood Education, the bilingual program is absent, but a native speaker teaches English once a week during the official teaching hours.

As regards the socioeconomic level of the participants' families, they belong to a medium-low status – with numerous cases of monoparental families with hardly any income. As regards parental education, legal representatives hold basic and post-compulsory education studies.

Our participants belonged to the 5th Year of Primary Education, which means that they had started learning English at school when they were three years old. Furthermore, most of them had received private tuition in English in addition to their lessons at school. On average, they had been receiving 3-4 formal hours per week in Primary Education. There was, however, an additional lesson every week for all Primary grades as the school opted for implementing an intensification program of the English language rather than including non-linguistic subjects (e.g. Science or Maths) into the curriculum with English as the vehicular language.

As regards the methodology used by the teacher from the classes involved in this study, it should be noted that the EFL teacher resorted mainly to the English language textbook, as explained by the teacher herself. In this regard, the role of the listening and speaking skills is predominant, and the use of the L1 is restricted to instances where unknown or difficult vocabulary appears. Teaching emphasized both grammar and vocabulary learning via the use of the textbook, and other timely activities such as reading comprehension exercises. The teacher informed us that the children's writing practice was mainly based on several writing exercises from the textbook. These writing exercises involved short notes or informal letters. The instruction of writing did not go beyond the mere explanation of set phrases to use in these writing tasks. Feedback was mainly provided in an explicit manner. Additionally, the teacher herself admitted that the writing skill is regarded as one of the students' greatest weaknesses.

►Participants

Out of a convenience sampling, a total of 18 participants (10 female and 8 male) with an average age of 10 years old (M = 10.15) belonging to an EFL Primary class took part in the study. Participants were assigned randomly to each of the groups (EG or CG). All participants were provided with a consent form to be signed by their parents or legal representatives prior to the beginning of the study. Participants belonged to a 5th grade EFL Primary class, and their proficiency level was low according to their school grade, i.e. A2 level according to the Common European Framework of Reference, with regards to the Spanish education legislation. These participants are thus considered as low proficient.

Additional information about the participants was gathered through a demographic questionnaire which participants had to complete at the end of Stage 3. The questionnaire (see Appendix 2) asked about their tying skills and past language learning experience (such as the starting age of learning English and whether or not they had attended or were receiving extra EFL lessons, or the number of hours studying English outside school). Among the participants, 77.7% had started to study English before the age of 5, while 16.7% had started between the age of 5-7, and a 5.5% from the age of 8. As for their perception towards the English learning and achievements, 88.8% indicated that they liked the language, and regarding writing *per se*, 72.2%

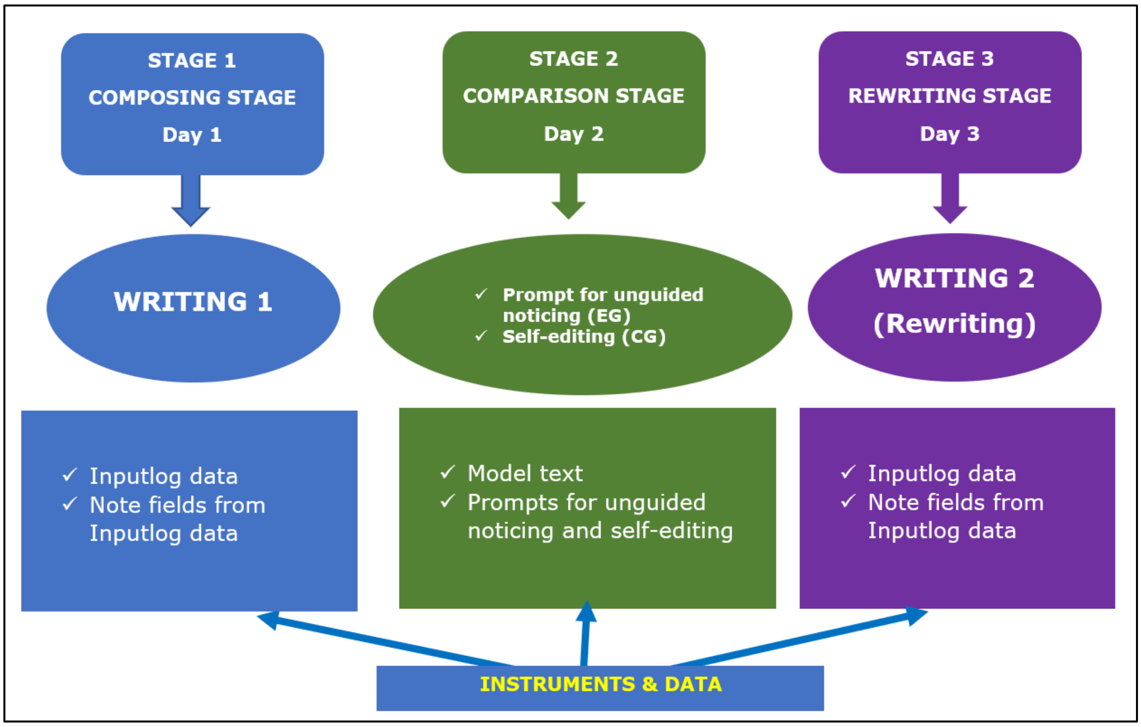
admitted they liked this skill. As to the degrees of difficulty during writing, 5.5% found it very difficult, 22.2% difficult, 55.5% easy, and 16.6% found it very easy. Regarding their own typing skills, 42% believe they did not write very well in contrast to a 31% who believe they did so. As regards their perceptions about their typing speed, 26% believed they wrote slowly while 52% considered they wrote at a normal speed. In section 4.3.2., objective data on the children's keystrokes per minute is provided by Inputlog.

4.3. Data collection instruments and procedures

4.3.1. Instruments

Several instruments were used in the present study for the collection and analysis of the data: (a) a keystroke logging tool (Inputlog) to record the participants' composing process, (b) note fields from Inputlog data generated, and (c) exit questionnaire. Additionally, participants received a model text as feedback and 3 sets of instructions: tasks instructions for Writing 1 and Writing 2, and prompts for the feedback processing/self-editing stage. The description of each of these instruments is provided below according to the stage in which they were used. Figure 17 shows these instruments as well as the data generated.

Figure 17.
Research design, instruments and data.



►The writing task

The writing task used for Writing 1 and Writing 2 was a picture description task previously used in Cánovas (2017) and Coyle et al., (2018). In most of the studies on models as WCF, the written task has consisted of visual prompts, namely representing a story or other situations. In the case of adolescents (García Mayo & Labandibar, 2017) and adult studies, picture prompts consisted either of two or three photographs depicting a very basic sequencing (Hanaoka, 2005, 2007). Conversely, children studies on the use of models have widely resorted to picture-based story prompts (e.g. Coyle & Cánovas, 2019; Coyle & Roca de Larios, 2014; Roca de Larios et al., 2015) for their writing tasks. The use of a numbered picture-frame sequence allows the child to follow a route in the story.

Figure 18.
The "Scientist" six-picture frame story task (Cánovas, 2017).

Observe the pictures below. Write a **brief story** which reflects what happens in the pictures. Please do not forget to write about every single picture.

As can be seen, the task involves writing a story based on a six-picture frame. Children were told to write about the sequence of events depicted in these picture frames. In order to do so, children were provided with a printed DIN A4 and were verbally told to write the story on their computers in 30 minutes. No indications were provided in regard to the use of specific linguistic structures or vocabulary.

► Model text

Children in the feedback group were provided with only one model text following previous research indicating that providing two model texts would be excessive for young learners to handle (Cánovas, 2017). In order to ensure the validity of the model text selected, we used the same text as in Cánovas' (2017) study with young EFL learners. In fact, this text complied with what Cánovas (2017) stated, following Ortega (2009), in terms of linguistic competence. The text was written in present simple but the children were challenged through the use of linguistic elements beyond their current level. Thus, the use of new structures may help to "destabilize internal interlanguage representations" (Ortega, 2009, p. 6, as cited in Cánovas, 2017) to promote L2 learning.

Figure 19.
Model text given to the treatment group (Cánovas, 2017).

THE SCIENTIST

One day, a scientist is in his laboratory. It's late and his dog is sleeping on the table next to him. After a lot of work, he finishes his new potion. He is very excited and decides to try it immediately. He drinks the whole potion. Suddenly, he feels very strange. There is a loud noise and a bright flash of light. Then, the scientist turns into a cat! Finally, the scientist's dog wakes up and attacks the cat.

As can be seen, the text contains familiar elements (e.g. dog, table, work, cat, drinks) combined with more challenging new vocabulary items and lexical chunks (e.g. potion, a bright flash of light, turn into, suddenly).

►Keystroke logging tool

As advanced above, the children were gathered in a computer room where Inputlog was installed (Leijten & Van Waes, 2013). The software was tested appropriately before any research procedure was undertaken. As was reviewed in section 2.7. (Chapter 2), *Inputlog* allowed capturing online processes during writing, especially keyboard presses, pause and revision behaviour. Inputlog has been used in countless research on writing processes and pausological behavior (e.g. Barkaoui, 2016, 2019; Révész et al., 2019; Xu, 2018), proving to be a reliable software tool for this research purpose. Additionally, we opted for Inputlog as our keystroke-logging tool given its integration with common text processor software such as MS Word (Leijten & Van Waes, 2005), and its research potential for studying L2 writing processes. In light of this, Inputlog generates a vast amount of data which may be set up according to certain empirically relevant parameters (e.g. establishing the pause threshold) and allows for replaying the writing process in its entirety.

►Note fields from Inputlog data generated

These notes correspond to annotations related to the writing processes and pausological behavior on the basis of Inputlog data output generated at both times. They provide a view of the participant's writing process, including text production, pausing behavior (which allows for inferring further associated processes such as planning) and revisions made. We gathered these data prior to delving into a more fine-grained analysis in an attempt to unveil any underlying reasons behind the different decisions made during the composing process. These note fields emerge from three different types of analyses (see Figures 20 to 22 below), as will be further detailed in "Data analyses procedures". We made use of these three types of Inputlog analyses in order to gain insight into the cognitive value of pauses as regards each of the writing processes, but fundamentally planning and revision. As seen in Figure 20, pause-based analyses provided the amount of pauses before the output in each row. Revision-based analysis presented the type of revision (insertion or deletion) and a quasi-linear view of the writing process together with the text production, pause location and duration, and cursor and keyboard action performed. Finally, the S-Notation analyses provide a fully linear perspective of how the text was written, including the revisions made, and the order in which such revisions were performed. Building on this brief overview of the three analyses, pauses and the overall text production were examined to indicate potential writing processes.

Figure 20.
Linear Analysis: pause-based.

PAUSE-BASED ANALYSIS

	Interval	Output
	Pause 0 (2422 ms)	
GPlann.	Pause 1 (7922 ms)	[Movement][LEFT Click][Movement]
	Pause 2 (2408 ms) ⊗	in·
	Pause 3 (2288 ms)	[BACK][BACK][BACK]
Form	Pause 4 (4328 ms)	in·a·laboratory·there·was·a·
	Pause 5 (4848 ms)	boy·
	Pause 6 (9336 ms)	[BACK],a·dog·h[BACK]and·his
Rev	Pause 7 (13544 ms)	[BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK]·and·his·dog·
Form	Pause 8 (6648 ms)	,the·dog·it
	Pause 9 (4536 ms)	[LCTRL][OEM_7][OEM_7][OEM_7][BACK]s[LCTRL][RALT]·bored·

Figure 21.
Linear Analysis: revision-based.

REVISION-BASED ANALYSIS

Interval	Output
0: INSERT	{2422}[Movement][LEFT Click][Movement]{7922}in·{2408}
1: DELETE	[BACK][BACK][BACK]{2288}
0: INSERT	in·a·laboratory·there·was·a·{4328}boy·{4848}
2: DELETE	[BACK]
0: INSERT	,a·dog·h
3: DELETE	[BACK]
0: INSERT	and·his{9336}
4: DELETE	[BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK]
0: INSERT	·and·his·dog·{13544},the·dog·it{6648}
5: DELETE	[LCTRL][OEM_7][OEM_7][OEM_7][BACK]

F Lexical

P Context

Typing P Context/ lex.

Figure 22.
S-Notation analysis.

S-NOTATION ANALYSIS

[[[in·]1|1in·a·laboratory·there·was·a·boy[[·]2|2,a·dog·[h]3|3and·his]4|4·and·his·dog·,the·dog·i[t]5|5s·bored·and·the·boy·has·go[y]6|6t·a·pocion.[the]7|7then·the·boy·drink·the·pocion··but[·,]8|8,the·pocion·i[[s·flse·]9|9t]10|10s·mlva[[de{vade}11]12|12|13{in·bad·}13]14|14|15{f[l]16|17{alse[·]18}17|18}15|16·,then·|11·the·pocion·exploted·,but·the·boy·evolution·a·cat[·and·fi]19|19.fin[n]20|20ally·the·dog·atack·[the·]21|21the·cat.]22|22in·a·laboratory·there·was·a·boy·and·his·dog·,the·dog·it´s·bored·and·the·boy·has·g

►Prompt for unguided noticing

Twenty-four hours after Stage 1, participants engaged in Stage 2. They were given a prompt for unguided noticing that encouraged them to review their original writing 1, available on screen, and compare it with a model text as WCF. They were not provided with any other instructions than the ones provided in this prompt.

Thus, participants were asked to note down differences or errors, as was done in previous studies with models (see Hanaoka, 2007; Roca de Larios et al., 2015). The use of this feedback technique is justified since young children need some type of engagement with the feedback given the absence of a cycle of instruction in WCF (as opposed to Coyle & Cánovas, 2019). Therefore, participants are required to put some effort into noticing the features and differences between their output and the model provided, activating a certain degree of cognitive comparison (Ellis, 1997), which potentially leads to noticing the gap and noticing the hole (Schmidt, 2001). In this sense, both of our RQs (see section 3.2., Chapter 3) aim to identify the extent to which the provision of model texts as WCF might impact writing processes and pausological behavior, respectively.

Figure 23.
Prompt for unguided noticing.

COMPARAR DIFERENCIAS ENTRE TEXTOS

Instrucciones: Revisa el texto que escribiste y compáralo con el modelo que se te da abajo para así mejorarlo lo máximo posible.

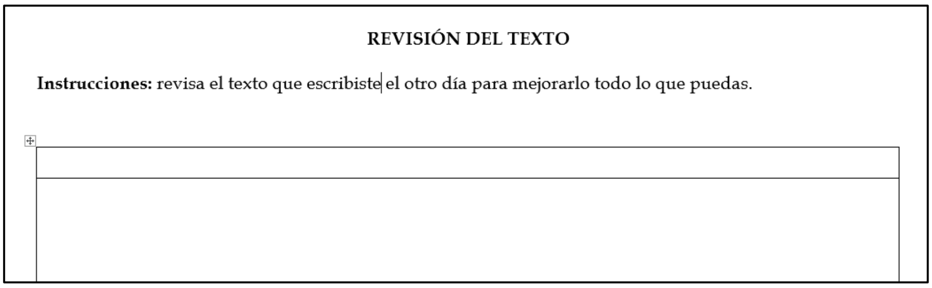
MODELO
One day, a scientist is in his laboratory. It's late and his dog is sleeping on the table next to him. After a lot of work, he finishes his new potion. He is very excited and decides to try it immediately. He drinks the whole potion. Suddenly, he feels very strange. There is a loud noise and a bright flash of light. Then, the scientist turns into a cat! Finally, the scientist's dog wakes up and attacks the cat.

DIFERENCIAS

►Prompt for self-editing

As can be seen in Figure 24, a self-editing table was also provided to participants in the non-feedback group during phase 1 in stage 2 twenty-four hours after stage 1. Self-editing has also been used in previous research with models, such as Luquin & García Mayo (2020, 2021). This self-editing table, unlike the prompt for unguided noticing for the EG, did not provide any model text, and thus children in the CG were invited to self-correct their texts.

Figure 24.
Prompt for self-editing.



REVISIÓN DEL TEXTO

Instrucciones: revisa el texto que escribiste el otro día para mejorarlo todo lo que puedas.

►Final questionnaires

There were two questionnaires which included both close-ended (Likert scale) and open-ended questions, which participants completed through *Google Forms*. All the participants completed a demographic questionnaire (see Appendix 2) which asked them about their sociodemographic data, past language learning experience and typing skills. The second questionnaire, an exit questionnaire (see Appendix 3), was only completed by the EG and its aim was to obtain information about their perceptions towards L2 writing and models as WCF (Marsden & Togerson, 2012).

4.3.2. Piloting

Once the task was selected, piloting was carried out for a number of reasons. Firstly, to test the validity of the task for young EFL learners as a way of supporting the use made by other scholars (Cánovas, 2017; Coyle et al., 2018). Secondly, piloting allowed us to assess whether the timing proposed for the task was adequate as well as the time span that the use of the computer required. Given that previous research had been carried mainly in traditional pen-and-paper settings (e.g.

Cánovas, 2017; Coyle & Cánovas, 2019; Coyle et al., 2018; Roca de Larios et al., 2015), the nature of this research in a digital environment required testing how children coped with computer writing. Therefore, I observed very carefully if using computers – in contrast to pen-and-paper – delayed in any way the participants' writing performance as a result of typing. Third, the instructions for the second phase of stage 2 (use of prompts for unguided noticing, and self-editing) needed testing given the scarcity of research on the issue with child participants. We thus had to verify their suitability in terms of children's cognition. Finally, stimulated recall has been scarcely used in empirical research with children, especially with regards to tapping into writing processes where this population is still underrepresented. In this regard, most of the studies on writing processes have resorted to concurrent verbal protocols such as TAP (e.g. Manchón & Roca de Larios, 2007; Roca de Larios et al., 2006, 2008). Recent research trends have shown that retrospective verbal protocols are useful to gather data on the writing processes, especially after engaging in digital writing (e.g. Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2019; Sasaki, 2000). Thus, our piloting intended to shed light on the usefulness of this technique for the final data collection, and the potential pitfalls.

Among the many outcomes that this piloting yielded, it is important to note that timing conditions for the task were restricted to 30 minutes. The decision of establishing this time limit for the writing process was dually founded on our experience with the piloting and on previous studies which accounted for an average time of 20 minutes in the young EFL students' writing process (e.g. Cánovas, 2017), right in accordance with other studies involving Secondary EFL students, where the time restrictions were set in 20 minutes (García Mayo & Labandibar, 2017). Building on the preceding information, timing conditions were set on 30 minutes for the piloting in an attempt to verify whether our children participants were able to successfully fulfill the task without major time concerns. Piloting this aspect brought to light that young EFL learners were able to attain the objectives set for the task in due time (M= 25.54 min; SD= 3.18).

Another concern was the children's ability at typing. Although they did not evidently show advanced fluency at typing or using computer commands, they certainly overcame any difficulty in very little time, either on their own or by asking me any questions concerning technical computing aspects. The keystrokes produced per minute by the participants in the pilot study (M= 19.38; SD= 3.78) indicate their difficulty at typing and also the cognitive effort of writing the task. Compare it, for instance, with Chukharev-Hudilainen's (2014) study where undergraduate's typing rate was 110 keystrokes per minute (SD= 52).

Furthermore, one of the major concerns was the way in which participants were going to engage in a stimulated recall interview, especially at such a young age. Despite the preference of concurrent verbal protocols in writing processes research, there is only a precedent study by Myhill (2009), which looked into L1 children's patterns of composition, in which stimulated recall with children was used. For the present study, we attempted to verify the usefulness of this verbal protocol technique to tap into L2 children's perceptions of their writing processes. Also, we were interested in checking if children understood our queries in the stimulated recall. In this sense, another objective of the pilot study was to check that our prior holistic analysis of the Inputlog data generated would be sufficiently effective to identify the relevant aspects of the writing process to be dealt with in the stimulated recalls.

To summarize all of the above, piloting served as a basis for the final data gathering procedure. It allowed for testing the validity of the task with children as participants, the suitability of the computer-mediated environment, and the identification of any potential technical issue (e.g. computer issues, typing issues in children) or methodological drawbacks (e.g. the accessibility of Inputlog as a research tool).

4.3.3. Data collection procedure

As already mentioned and graphically depicted in figures 16 and 17, our research entailed a three-stage data collection procedure. In section 4.1, we have provided an overview of each stage as well as a description of the research instruments associated with each stage. In what follows, we elaborate further on the methodological decisions taken for data collection at each stage (see Table 4).

Table 4.
Data collection structure of the study.

	Day 1 - Time 1 ***Stage 1***	***Day 2 - Treatment*** ***Stage 2***	***Day 3 - Time 2*** ***Stage 3***
Feedback Group (EG)	Writing 1	Phase 1: Feedback processing (model) with prompt for unguided noticing Phase 2: Stimulated recall	Writing 2 (rewriting task) without access to any of the previous materials
Non-feedback Group (CG)	Writing 1	Phase 1: Prompt for self-editing (no feedback). Phase 2: Stimulated recall	Writing 2 (rewriting task) without access to any of the previous materials

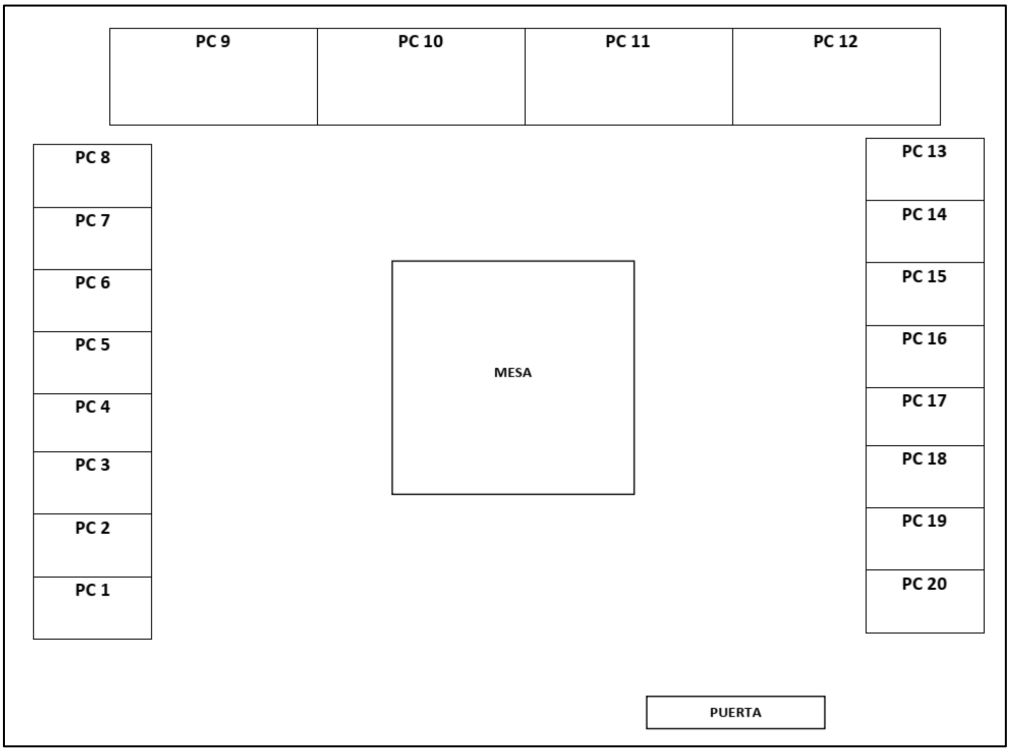
►Stage 1

At this stage, which we shall refer to as *composing stage*, participants were asked to write a story by taking a picture-based prompt as reference (see "Task" section). The prompt was handed in to participants as a black and white DIN A4. In order to ensure full understanding, the instructions for the task were provided orally in Spanish inasmuch as using English might have hindered task completion on the part of the children participants. All the participants belonging to each group wrote their texts at the same time.

Participants were asked to write their texts using the computer via the KSL tool selected, *Inputlog* (Leijten et al., 2012; Van Waes et al., 2011). The screens on the participants' computers already showed the Inputlog window interface, where participants wrote their personal information (namely their full names and age). Afterward, they were requested to press on the "Record" button to start both the Inputlog and writing session. Once they clicked on it, an *MS Word* window opened, and participants were encouraged to start as soon as possible. Participants were allowed to use up to 30 minutes for the completion of the task. On average, they made use of 17-23 minutes out of the whole timing given. They were informed about the time left every 10 minutes, and on a more frequent basis by the end of the session.

Both EG and CG completed the task in the same room, a decision that did not influence the collection of the data on account of the organisation of the computer room (see Figure 25), which avoided any possible visual or oral contact amongst them. There were wall divisions between each computer station. The participants completed their tasks individually and communication between them was not allowed under any circumstances. The supervision of the process was undertaken by the researcher himself.

Figure 25.
Arrangement of the school computer room.



Once participants had concluded their writing, they were requested to inform the researcher about it so that the Inputlog session could be conveniently terminated and the data stored within the computer.

►Stage 2

Stage 2, referred to as *comparison stage* in studies dealing with models as WCF (to name a few, Cánovas, 2017; Coyle, Cánovas & Roca de Larios, 2018; García Mayo & Labandibar, 2017; Kim, 2015), took place the following day after Stage 1. There were two phases in this stage 2.

Phase 1. Once participants were sitting in front of their computers, their initial text written at Stage 1 appeared on the screen. The participants in the feedback group were provided with a printed DIN A4 containing the corrective feedback in the form of a model text, and a prompt for unguided noticing where they were asked to note their mistakes or provide free explanations. This allowed them to improve their texts after establishing a comparison between their own output and the output provided as was done in previous research (Cánovas, 2017; García Labandibar & Mayo, 2017). In the case of the nonfeedback group, they were not provided with any type of feedback. Instead, they had text 1 on screen and a self-editing sheet in which they were prompted to observe their text and note down any improvement they deemed appropriate.

Both groups were provided with an MS Word document with their composition, which they were allowed to modify by highlighting those linguistic units or ideas which they thought could be modified or which they wanted to comment on. They were allowed a maximum of one hour to complete this part of the study, which in most cases was completed in 30 minutes hence confirming the tendency observed in the piloting stage.

Phase 2. Building on recent research stating its usefulness on L2 writing research (Latif, 2018; Révész, Michel & Lee, 2019), both groups participated in an audio-recorded stimulated recall protocol once they had finished their feedback comparison stage. After Phase 1, participants were led to a separate room where they were provided with a replay of their writing at time 1 provided by Inputlog (that is, Stage 1). One by one, the researcher would then ask questions related to their writing processes, and pausological behavior. Once they finished, they would go back to the classroom with their teacher. Those participants who had not finished phase 1 yet were supervised at all moments by a fellow researcher in the computer room. This prevented that participants were told about this stimulated recall procedure by any of the other participants, safeguarding external validity.

Participants were asked about why they took so long to write a word, allowing for differentiation between possible planning or, for instance, lexical retrieval. Thus, participants were shown their text, and the researcher highlighted or signalled the part where, for instance, a long pause was occurring. Furthermore, given the researcher's access to the revision-based analysis, questions were also asked on why some linguistic units were modified (for instance, be it spelling or grammatical issues). Apart from clarifying these aspects, the researcher also noted down the participants' responses. Finally, the stimulated recall did not have any specific timing restriction, but the average duration fluctuated from two to five minutes. Participants did not have contact between them after leaving the room where the stimulated recall was taking place.

Unfortunately, the stimulated recall data for Writing 1 in this stage 2 was eventually not reported in this PhD for two reasons: (1) the data were scarce and uninformative, and (2) the stimulated recall did not largely provide substantive evidence of the writing process. Likewise, this poor data led us not to carry out the stimulated recall procedure for Writing 2. This is obviously a limitation that will be further looked into in Chapter 7 (Conclusion).

►Stage 3

This third stage, or the "*Rewriting stage*", involved the rewriting of the original texts following exactly the same procedure as in Stage 1.

After participants had finished their task, they were immediately provided with two questionnaires: a socio-demographic questionnaire, which both EG and CG completed, and an exit questionnaire with questions about the WCF provided, which only the EG completed.

4.4. Data analyses procedures

4.4.1. Operationalization and measurement of the variables

This section provides a detailed account of the principles that have guided our operational definitions and resulting choice of measures in the temporal analysis of both (a) the writing process and (b) pausological behavior.

Bearing in mind the research questions of our study, and following previous research (especially contributions to Lindgren & Sullivan, 2019) which indicated the multidimensional nature of writing processes and subprocesses, we firstly decided which different operationalizations of each one of our dependent variables (engagement with planning, revision, formulation, on the one hand, and pausological behavior, on the other) would be selected.

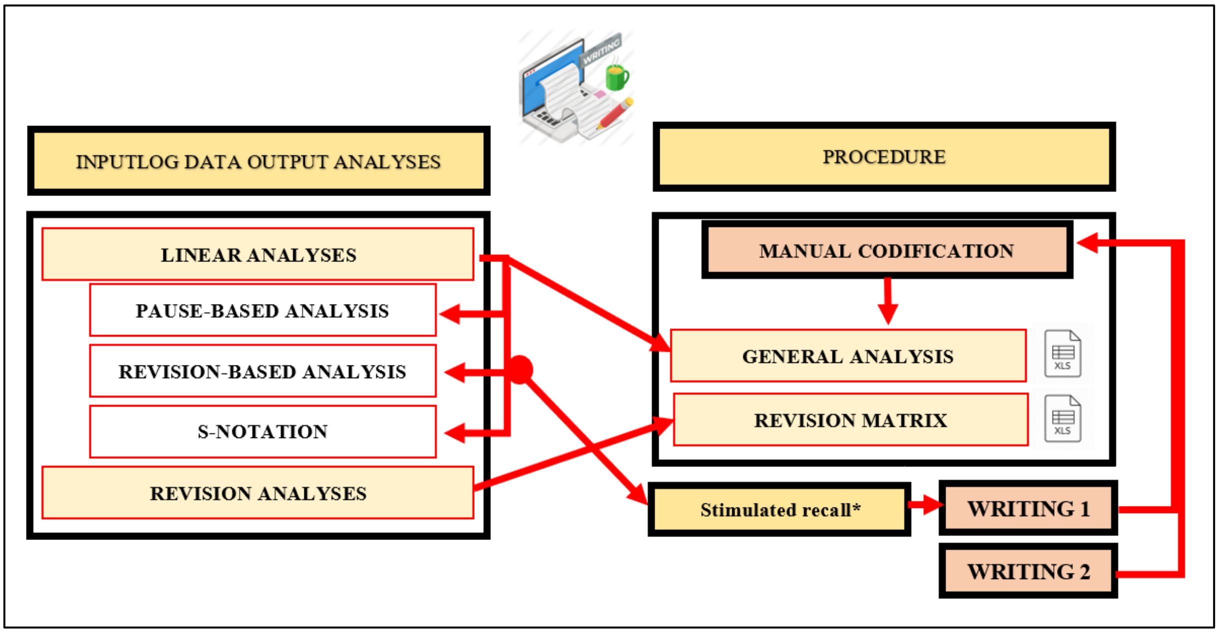
Second, another important consideration is the use of both general and specific measures in both (a) and (b): while general measures are used to refer to global calculations (e.g. total duration), specific measures are used to narrow the focus on a single discrete characteristic (e.g. total duration in interval 1). Finally, the data from Intpulog was supplemented with the use of the the side-by-side comparison between other output from the Inputlog analyses, as will be further explained. In the following sections, a detailed description of the operationalization and measures employed in this dissertation will be provided.

4.4.1.1. Temporal dimension of writing processes: constructs and measures

Table 5 below provides a detailed description of the steps followed in the analysis of the data. The operational definition of each writing process and their measurement will be further described in the following sections. As was mentioned before, pauses constitute the core of some of the writing processes (especially planning) given their fundamental role as visible indicators of these underlying writing processes (Alamargot et al., 2007). Nonetheless, we shall be cautious given that pauses are only "*indirect* evidence of writers' underlying cognitive processes in writing" (Barkaoui, 2019, p. 531. Emphasis added).

In Figure 26, we provide an overview of the different analyses used to process the data, and a schema of the triangulation of the data. After the primary phase in stage 2, the ensuing analyses consisted of a side-by-side comparison of both the data from Writing 1 and Writing 2. These data were further classified and fine-grained in an XLS spreadsheet. As mentioned in the "Instruments" section, the *Linear Analyses* were computer-generated as preliminary data in Stage 2 given the useful information they provided on the writing process *per se*.

Figure 26.
Triangulation of the data to code each of the writing processes.



Note: *As was anticipated 4.3.3., the data from the stimulated recall aiming to tap into the cognitive processes behind Writing 1 were not considered for the present dissertation.

Table 5 below presents an overview of the different steps followed in the analysis of the Inputlog data output, the different methodological decisions taken, the rationale that justified each decision, the Inputlog analyses used, and the relation to the research questions.

Table 5.
Description of the steps in the analysis.

Step in the analysis	**Decisions**	**Rationale**	**Inputlog data output: analyses used**	**Relation to research questions**
1. Identification of pauses with writing processes	Observation and identification of pauses within the text produced in a linear fashion (i.e. including revisions and thus the full process of writing a text). Careful examination of Inputlog data output analyses to match those pauses with writing processes.	Pauses constitute the basis for underlying cognitive processes (Alamargot et al., 2007; Barkaoui, 2019). A preliminary identification of what writing processes pauses might underlie will facilitate a further more in-depth categorization. This preliminary identification leads to: (1) establishing connections between these pauses and the surrounding events during text production, after the global holistic analysis of the Inputlog data reports and (2) considering globally the position and role of these pauses.	■ *Linear Analyses*: pause-based and revision-based analyses. ■ S-Notation. It provides information on the locations in text where modifications were made. *Nota bene*: these analyses were examined carefully side-by-side to identify the most notable pausological behavior, and a global overview of how the writing process developed.	RQ 1
2. Identification and classification of the Inputlog data into spreadhseets.	All the Inputlog output data analyses were converted into a spreadsheet where the information was categorized.	The organization of the data into spreadsheet files allowed for easier handling of the data to be analyzed.	■ *General Analysis*. All the data provided by Inputlog are included in this analysis, although not presented and merged as in the case of the other output analysis (e.g. Revision	RQ1

			Matrix), while the previous analyses process them for specific purposes. The main spreadsheet file where all the data were categorized and measured was General Analysis.	
3. Division of the writing process into intervals.	Temporal division into three intervals, related to initial, medium and final stages of the writing process (as in Kellogg, 1987, 1988; Roca de Larios et al., 2008). Each composition time was divided into three equal intervals.	Dividing the writing process into intervals allows for observing how each writing process differs in terms of duration or frequency in each phase (Van Waes & Leijten, 2015).	■ *Pause Analysis*. This type of analysis was used since, besides information on the nature of pauses, it also provides an automatic temporal division of the writing process into intervals. ■ *General Analysis.* Once we obtained the time-distributed division into intervals from Pause Analysis, the subdivision per interval was manually done in the spreadsheet containing the output of this Inputlog analysis. As said, General Analysis contains all the information, but it is not further organized or merged as in the output of other analyses.	RQ1 RQ2
4. Identification and coding of planning	1■ Setting the pause threshold at 2 seconds as in	Pauses indicate covert underlying processes, particularly above 2	■ *General Analysis*. The planning processes were manually categorized in	RQ1

processes in writing 1 and writing 2.	previous research (Barkaoui, 2016, 2019). 2■ Manual classification of pauses at different textual boundaries to infer the potential underlying type of planning (see section "Analysis of planning" for a description). These pauses are considered as *planning episodes*. 3■ The manual classification is supported by the side-by-side comparison of the data in the analyses in the first step of this data analysis procedure.	seconds (Wengelin, 2009). The position and location of the pause, mainly before sentences and paragraphs, but also words, allow for classifying whether those pauses are planning processes, and subsequently, their nature (planning globally or locally). The position may provide essential information on why this pause is planning, and its scope. The behavior preceding and ensuing this pause is equally revealing.	the spreadsheet file containing these data. The following analyses were used as a support for the side-by-side comparison of the writing process, and for observing the preceding and follow-up behaviors after the pauses: ■ Linear Analysis: pause-based and revision-based analyses. ■ S-Notation.	
5. Identification and coding of formulation processes in writing 1 and 2.	1■ Observing the boundaries between the stretches of text production and revisions through the data in *Revision Matrix* provided by Inputlog. 2■ Further manual categorization in General Analysis with the support of a side-by-side comparison with the Revision Matrix. 3■ Matching the time duration of the categorization in Revision	The formulation process is generally associated with the conversion of thoughts into written language, and the production of stretches of text, with and without having to engage in problem-solving (Roca de Larios et al., 2008). Given the impossibility of discerning whether a certain pause refers to this problem-solving formulation process, those pauses will be assimilated	■ *General analysis*. The formulation processes were manually categorized in the spreadsheet file. Other processes associated with planning (previously classified) were filtered out to avoid mixing the data. The following analyses were used as a support for the side-by-side comparison of the writing process, and for observing the preceding and follow-	RQ1

	Matrix (i.e. "normal production") with the manual categorization in General Analysis (see section "Analysis of planning" for a further description of this process). This entails that the pauses classified as planning are filtered out.	as part of global planning. The ones below 2 seconds included between the stretches of text production will be considered formulation pauses.	up behaviors after the pauses. ■ Revision Analysis: Revision Matrix. ■ Linear Analysis: pause-based and revision-based analyses.	
6. Identification and coding of revision processes in writing 1 and 2.	1■ Manual categorization of type of revisions (microscopic and macroscopic; see section "Measuring microscopic and macroscopic revisions" for a description of these processes) from the data in the *Revision Matrix*. 2■ Isolation of revision episodes in the data in the *Revision Matrix*. 3■ Filtering out the revision episodes in the data in the *General Analysis*. 4■ Categorizing the data in the General Analysis together with the type of revision, merging action time (i.e. the actual time devoted for a keypress)	Revisions entail the modification of already written text at any point of inscription (Barkaoui, 2007). Given the difficulty to distinguish between self-corrections (that is, revisions in formulation) from actual revisions which occur after the process of formulation, these will be included as part of one subtype of revision (i.e. microscopic revisions). As there is no possibility to delve into the students' inner speech while writing, revisions will be considered from the product perspective (see section "Analysis of revision" for a detailed description).	■ *Revision Analysis: Revision Matrix*. Firstly, the revision processes were manually categorized in the spreadsheet file containing these data. Other analyses were used as a support for this first categorization: ■ Linear Analysis: pause-based and revision-based analyses. ■ S-Notation. This file contained a linear representation of the text written, indicating deletions and insertions as well as the order in which these have taken place. ■ *General analysis*. After having categorized the revision episodes in the Revision Matrix, these are	RQ1

	with pauses and other movements (e.g. back keyboard presses).		manually categorized once again in this spreadsheet file following this previous classification.	

These analyses were used to codify the data in terms of the different macro processes of the composing process i.e. planning, formulation and revision. Such data were obtained from the analyses and factual information processed by Inputlog, and subsequently classified accordingly. Previous research (for instance, Beetwelt et al., 1994; Roca de Larios et al., 2008; Tillema et al., 2011) guided both data coding and analyses, especially regarding the operationalization of the temporal distribution of the writing process.

Taken as a whole and building on previous research, the writing process was subdivided into three intervals, which do not refer to the writing subprocesses themselves according to the scientific literature (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes, 2012). These intervals make reference rather to the initial, medium and end stages of the whole writing activity as used in previous research (e.g. Barkaoui, 2019 or Roca de Larios et al., 2008). Inputlog allows for this division into three intervals through one of its analysis (see Figure 27 below). In order to shed light upon the writing processes (i.e. planning, formulation and revision), we will examine pauses in their context (see section 2.7.3. for a detailed review of pausing behavior) in an attempt to verify potential indicators of higher-level cognitive processes (Spelman Miller, 2006; Schilperoord, 1996; Wengelin, 2006).

To identify the writing processes, several KSL analyses were carried out within Inputlog: *General Analysis*, *Revision Analysis*, *Linear Analysis* (pause-based and revision-based analyses) and *Pause Analysis*. As mentioned in section 2.7. (Chapter 2, "An overview of Inputlog"), each of these analyses provided useful information on the global quantitative measures of the composing process, such as total composing time, active writing time, and other more itemized information, for instance, the number of revisions made and their duration or their temporal allocation. Building on these data, micro-processes were subsequently identified manually within the different macro-processes, namely planning and revision.

The analyses provided raw data of the writing process, which needed fine-grained manual analysis, even if the data obtained was automatically quantified by Inputlog in terms of frequency, location, and duration. Thus, each of the previously mentioned Inputlog analyses (i.e. General Analysis, Revision Analysis, Linear Analysis and Pause Analysis) were computer-generated after the completion of stages 1 and 3. Each of the files was in XML format, which was later converted into an XSL (spreadsheet) file, as will be illustrated in the ensuing sections. For instance, *Pauses Analyses* provided direct data of the three intervals in which the writing process was divided (see Figure 27) Despite the vast amount of data provided by each analysis, not everything is to be exploited in our research study.

Figure 27.
Extract of a Pause Summary Logging File.

Summary per Interval	
Interval #1	
Start Time	00:00:00
Number of Pauses	19
Arithmetic Mean of Pauses (s)	10.957
Median Pause Time (s)	8.473
Geometric Mean of Pauses (s)	8.439
95% CI Log-Transformed - Low Boundary (s)	5.971
95% CI Log-Transformed - High Boundary (s)	11.926
Coefficient of Variation	82.078 %
Standard Deviation (s)	9.339
Interval #2	
Start Time	00:04:53
Number of Pauses	8
Arithmetic Mean of Pauses (s)	46.196
Median Pause Time (s)	17.468

In what follows, I provide a detailed analysis of the methodological decisions guiding the measurement of the variables of this study. To do this, I attempt to justify these decisions on the basis of how previous literature has conceptualized each of the writing processes and pausological behavior.

►Analysis of planning

Identification of the process

As indicated before, planning has been studied through the analysis of pauses, which are thought to be indicators of this and other writing processes (Alamargot et al., 2007). Previous research (e.g. Ellis & Yuan, 2004; Manchón & Roca de Larios, 2007; Révész et al., 2017; Roca de Larios et al., 2008; Sasaki, 2000, 2002, 2004, 2018) has operationalized planning as a process used by writers to generate and organize ideas and content, retrieve lexical or morphosyntactic features in a pre-formulation state. For instance, the analysis of pauses at certain points of inscription in the composing process has been the most widespread indicator of planning behavior used in both traditional studies (e.g. Goldman Eisler, 1972; Holmes, 1984, 1988; Spelman Miller, 2006a) and more recent studies involving the use of keystroke logging (e.g. Barkaoui, 2016, 2019; Révész et al., 2017; Spelman Miller, 2000, 2006b). Scholarly attempts have considered both the length and location of the pause as potential indicators of the planning process (e.g. Barkaoui, 2016, 2019), while other studies, e.g. Révész et al. (2019), stated that inter-clause and inter-sentence pauses reflected planning and pauses between words which are more likely to reflect either planning at a local level or formulation. Hence, these tendencies indicated in previous research have helped us set the stage for the manner of characterizing the different pauses as indicators of planning. Studies such as Barkauoi's (2016, 2019) reported that pauses before certain linguistic and textual units might indicate higher-level processes such as planning. Nonetheless, Barkaoui (2016, 2019) did not present his results with a systematized classification of these pauses as belonging to a specific writing process, but only referred to pauses before certain boundaries as indicators of such processes without a manual classification.

As it was explained above, most of the studies to date have resorted to analyzing pauses from a global perspective, resorting to inferences but without classifying these pauses as part of a specific writing process. Thus, our study intends to be a pioneering attempt at identifying and categorizing these pauses – which other studies have considered globally – with more specific processes of planning. Our methodological approach to analyze those pauses consisted of manually identifying and coding pauses indicating planning at both macro and micro levels. As was extensively reviewed in section 2.6., pauses might reflect planning at a macrolevel, i.e. text organization and content, while pausing at lower-level textual units is associated with lower-level writing processes, for instance, planning at a microlevel, i.e. lexical retrieval (Révész et al., 2019). We combined the data from several Inputlog analyses as announced in previous sections. Only the duration, location, frequency of pauses and revisions as well as other characteristics of the process (such as text production) were examined. I was pretty aware that this procedure may entail several drawbacks since the sole use of keystroke logging "does not allow for making inferences about the specific cognitive processes that underlie pausing behaviors" (Révész et al., 2019, p. 609). Despite this obvious limitation, shared by other studies where neither think-aloud or stimulated recall protocols were used (e.g. Barkauoi, 2019), we classified the data obtained from the analyses in Figures 28 to 30.

The proposal for analysis procedure followed is supported by what scientific literature has referred to regarding the duration and location of pauses, as they allow the researcher to infer the potential type of the writing processes, for instance planning, monitoring or otherwise. Building on previous empirical endeavors, we operationally define planning as those pauses preceding textual units at different lengths above 2000 ms, followed by text production of some kind. This, however, should be cautiously considered since pauses are indicators of the presence of planning, and not the planning process itself. The 2000-ms threshold is adequate for the present study since, when looking into high level cognitive processes (such as planning), the pause threshold has been traditionally set at 2 s in the scientific literature with adults (Van Waes et al., 2009). There are, however, no current guidelines to what pause threshold would be more convenient for keystroke-logging studies with children owing to the scarcity of studies with this population. Different measurements for planning have been used in the scholarly literature. In a recent study, Gánem-Gutiérrez & Gilmore (2018) used episode frequency and total duration, both in percentages, to account for the different writing processes including planning, though these authors did not distinguish the planning process per se, but rather "pausing" episodes separately. Likewise, Tiryiakioglu et al. (2019) resorted to both frequency and total duration to account for the different episodes of writing processes from the data obtained in the think aloud protocol and the KLS.

Categorizing the process

For the classification of the data into global and local planning, the Inputlog output from *General Analysis* (see Figure 28 below) gave us all the data required for the calculation of measurements. Thus, *pauseTime* indicated the amount of pausing together with *pauseLocationFull*, which accounted for the specific location of the pause. The *Process* column, manually created, indicated the type of planning involved, as will be further detailed in the ensuing subsections (*Measuring global planning* and *measuring local planning*). Such a classification was performed once paused-based and revision-based analyses were both manually compared. For instance, the pause {38641} is at greater length and might indicate planning at a global level. As can be seen, its location is *before sentences*, and by looking attentively at its place within a linear text (see Figure 29), one can observe that there is very limited text production afterwards. As this is the beginning of the composition process, one may infer that the participant further planning globally what to write next. Some hesitation is present with the subsequent pausing ({2871}), and this is followed by a deletion. Right after that, the {11352} pause indicates that some higher-level thought process was involved, possibly planning globally again. A series of pauses are followed, which could be part of consecutive planning, except for {3672}, which appears to be a formulation pause owing to its presence right in the middle of the verb 'to be'.

Figure 28.
General Analysis spreadsheet.

Interval	Value7	id	type	output	positionFull	doclengthFull	charProduction	RawStart	RawEnd
1	True	5	keyboard	o	0	1	1	2457359	2457546
1	True	9	keyboard	d	4	5	5	2461538	2461714
1	True	13	keyboard	t	8	9	9	2468778	2469010
1	True	39	keyboard	s	12	13	24	2569690	2569850

startTime	startClock	endTime	endClock	actionTime	pauseTime	Suma Act+P	Process	pauseLocationFull
10891	00:00:10.891	11078	00:00:11.078	187	5547	5734	Global	BEFORE SENTENCES
15070	00:00:15.070	15246	00:00:15.246	176	1568	1744	Local	BEFORE WORDS
22310	00:00:22.310	22542	00:00:22.542	232	5304	5536	Local	BEFORE WORDS
123222	00:02:03.222	123382	00:02:03.382	160	49440	49600	Global	BEFORE WORDS

Figure 29.
Revision-based analysis (Linear Analysis).

Interval	Output
0: INSERT	{4000}[Movement][LEFT Click][Movement]{38641}la{2871}ura·

1: DELETE	[BACK][BACK][BACK][BACK][BACK][BACK]
0: INSERT	[BACK]{11352}the·{2032}she·{5712}i{3672}s·happy{2536}·{9408}

Figure 30.
Pause-based analysis (Linear Analysis).

Interval	**Output**
Pause 3 (11352 ms)	ura·[BACK][BACK][BACK][BACK][BACK][BACK][BACK]
Pause 4 (2032 ms)	the·
Pause 5 (5712 ms)	she·
Pause 6 (3672 ms)	i
Pause 7 (2536 ms)	s·happy

Quantification or levels of measurement

Regarding levels of measurements for all the planning processes, we opted for using *frequency* to report the number of planning episodes (i.e. each pause was considered an episode), as well as *total duration* in percentages for the total time devoted to planning out of the total composing time. Some pausological studies have made use of measures such as median pause length (see Michel et al., 2020), but we opted for the total duration in percentages to standardize the data, as we are dealing with writing processes and relative data. Such a level of measurement neutralises inter-subject variability according to previous research (Roca de Larios et al., 2001, 2006, 2008).

The following subsections describe the operationalization of both *local and global planning*, and the measures used. As a prior note, the methodological decision of identifying two different categories within planning has been motivated not only empirically by previous research, but also methodologically. The reason for this decision is based on two premises central to the purposes of our research study: children are considered as more limited capacity processors than adults (Izumi, 2003; Leow, 2015; McLaughlin et al., 2015) and emergent planners (Bereiter & Scardamalia, 1987), along with their consideration as low proficient language learners and unskilled writers. This has led us to merge several sub-categories of planning from previous empirical attempts with adults or teenagers (e.g. Manchón & Roca de Larios, 2007; Roca de Larios et al., 2008) into these broader categories of global and local planning.

Measuring global planning

Sasaki (2002) identifies global planning as rough thematic planning, which is part of two broader categories called pre-task planning (also strategic planning), and online or within-task planning (Foster & Sekehan, 2001; Robinson, 2003), as was already detailed in section 2.6.2. (Chapter 2). *Global planning* has been defined as "the overall planning of the content of the text to be written" (Sasaki, 2018, p. 3) and, in terms of processes, it has been characterized as longer pauses indicating macroplanning processes such as paragraph organization (Spelman Miller, 2006b). Global planning implies paragraph organization, the generation of ideas or procedures. We have analyzed and categorized pauses before long textual units (such as sentences and paragraphs) as indicators of global planning, considering what other studies have suggested regarding pauses at these boundaries (e.g. Barkaoui, 2019; Sasaki, 2002). Three aspects are central to the analysis and measurement of global planning: the duration of a pause, its location (mainly before long textual units), and the surrounding environment since the text produced and the revision behavior around the pause is a significant factor to identify global planning or any other writing process. Key to understanding this methodological decision has been the consideration of young learners as limited capacity processors (Izumi, 2003; Leow, 2015; McLaughlin et al., 1983), as mentioned previously. Young EFL learners are emergent planners, and based on Bereiter & Scardamalia (1987), they write following a knowledge-telling model, without much previous preparation. Likewise, long pauses have been equated with formulation thought processes, such as reflection or pondering over morpho-syntactic features. In our study, as stated in Table 2, it has not been possible to discern between long pauses which reflect global planning and those which are reflective formulation processes. To solve this, global planning as a type of planning process will assimilate those long pauses (i.e. above the two-second pause threshold) that might refer to formulation processes. As our first RQ is mainly centered on gauging the weight of each writing process before and after the provision of WCF, the inclusion of this type of planning process will give us rich data on whether the generation or retrieval of ideas, for instance, was more evident after the treatment.

Measuring local planning

Local planning has been referred to as "planning what to write next without any sense of the overall picture of the composition to be formed" (Sasaki, 2018, p. 3). This leads to planning at the level of grammatical and lexical structuring of the mental activity for a further inclusion in the text (Hayes & Nash, 1996). This type of planning may be more aligned with the aforementioned consideration of young EFL learners as knowledge-telling writers (Bereiter & Scardamalia, 1987), and with the recurrent planning occurring when text generation (i.e. formulation) is taking place without an overall prior global plan. Hence, local planning has been inferred through shorter pauses in length than in global planning and which thus occur at lower-level units (i.e. words or phrases, or even within-words). Accordingly, they might reflect grammatical and lexical choices (Barkaoui, 2019; Spelman Miller, 2006b). This measure will allow us to observe whether young EFL learners resort more to planning locally after the provision of WCF, given the purported effect of model texts to notice lexical features (Coyle & Cánovas, 2019; Coyle et al., 2018; Roca de Larios et al., 2015; Martínez & Roca de Larios, 2010¸).

►Analysis of formulation

Identification and categorization

As stated in section 2.6. (Chapter 2), the *formulation* process, also called *translation* (Hayes & Flower, 1980; Hayes, 2012; Kellogg et al., 2013), has been defined as the process of converting thoughts into language (Roca de Larios et al., 2006) and, as such, it involves the translation of set goals into a textual framework (Roca de Larios et al., 1999). Formulation has been operationalized as continuous bursts of writing (Leijten & Van Waes, 2013; Leijten et al., 2012; Van Horenbeeck et al., 2012), that is, stretches of text production as indicators of formulation behaviour along with certain pauses part of this process (see Barkaoui, 2019). Under this same reasoning, we adopted Barkaoui’s (2019) criteria of evidence of formulation, which regarded this process as those bursts of text production whose pauses are “immediately followed by the keyboarding of additional text at the point of inscription” (p. 539). Inputlog classifies text production that is not interrupted by revisions as "normal production", which is "new text produced at the end of the text produced so far" (Van Waes & Leijten, 2019, p. 88). This "normal production" are episodes, that is, stretches of number of words containing formulation pauses. For our study, we substracted the pauses belonging to planning processes, and subsequently considered as formulation pauses those that were not categorized either as unknown pauses or whose length was below 2 seconds.

The elements of measurement that were categorized included frequency of the formulation episodes (by counting the number of normal production episodes in the *Type* column of Revision Matrix), their total duration (obtained from General Analysis output) and two additional non-temporal measures: the edits and words produced per interval. Van Waes & Leijten (2019) define *edits* as "the number of actions that the writer needed to complete the action. This also includes each keystroke, shift key, cursor movements, mouse clicks etc." (p. 89). We obtained these data from Inputlog, thus no manual coding was necessary except for the classification of individual processes. The analyses where this information was taken from were primarily *General Analysis*, and *Revision Matrix*. Additionally, Inputlog provides the timestamps for each keystroke and cursor or keyboard movement. These data are included in the General Analysis and are imported into an XLS spreadsheet. In doing so, keystrokes, pauses and other operations may be classified as processes, but also pauses below the two-second threshold which are not classified as part of the planning process. Each row contains either a keystroke, cursor movement or any other operation captured. Figure 18 below presents a close-up of the output of this analysis. The column *Process* is manually completed by indicating the writing process involved, just as it was done for planning.

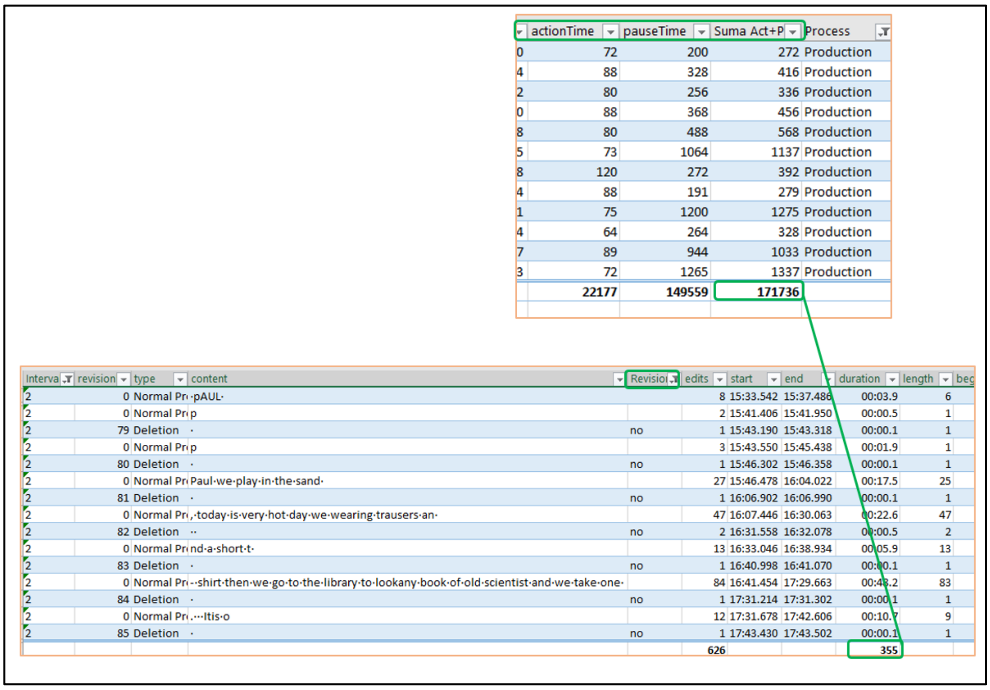
Quantification and levels of measurement

Before categorizing and quantifying the formulation process, we filtered out the planning process from the *Process* column using the filtering option from Excel in an attempt to avoid mixing the time allocated to this process with the formulation one.

Once this was done, a side-by-side comparison was performed with the Revision Matrix output analysis (in Figure 31, the lower-centered chart). The Revision Matrix includes information about the revisions (insertions and deletions), and also about Normal production. The side-by-side comparison between the information in the General Analysis and Revision Matrix was made to obtain an overview of the amount of text produced. Afterwards, we were able to check whether the time devoted to it was similar to the one shown in General Analysis.

Figure 31.

Comparison of the total duration of the formulation process between General Analysis (in the upper-right corner) and Revision Matrix (the lower-centered chart).



	actionTime	pauseTime	Suma Act+P	Process
0	72	200	272	Production
4	88	328	416	Production
2	80	256	336	Production
0	88	368	456	Production
8	80	488	568	Production
5	73	1064	1137	Production
8	120	272	392	Production
4	88	191	279	Production
1	75	1200	1275	Production
4	64	264	328	Production
7	89	944	1033	Production
3	72	1265	1337	Production
	22177	**149559**	**171736**	

Interva	revision	type	content	Revisio	edits	start	end	duration	length	beg
2	0	Normal Pr	·pAUL·		8	15:33.542	15:37.486	00:03.9	6	
2	0	Normal Pr	p		2	15:41.406	15:41.950	00:00.5	1	
2	79	Deletion	·	no	1	15:43.190	15:43.318	00:00.1	1	
2	0	Normal Pr	p		3	15:43.550	15:45.438	00:01.9	1	
2	80	Deletion	·	no	1	15:46.302	15:46.358	00:00.1	1	
2	0	Normal Pr	Paul·we·play·in·the·sand·		27	15:46.478	16:04.022	00:17.5	25	
2	81	Deletion	·	no	1	16:06.902	16:06.990	00:00.1	1	
2	0	Normal Pr	,·today·is·very·hot·day·we·wearing·trausers·an·		47	16:07.446	16:30.063	00:22.6	47	
2	82	Deletion	··	no	2	16:31.558	16:32.078	00:00.5	2	
2	0	Normal Pr	nd·a·short·t·		13	16:33.046	16:38.934	00:05.9	13	
2	83	Deletion	·	no	1	16:40.998	16:41.070	00:00.1	1	
2	0	Normal Pr	··shirt·then·we·go·to·the·library·to·lookany·book·of·old·scientist·and·we·take·one·		84	16:41.454	17:29.663	00:48.2	83	
2	84	Deletion	·	no	1	17:31.214	17:31.302	00:00.1	1	
2	0	Normal Pr	.···Itis·o		12	17:31.678	17:42.606	00:10.7	9	
2	85	Deletion	·	no	1	17:43.430	17:43.502	00:00.1	1	
					626			355		

To illustrate the procedure we followed, let us take the example in Figure 31 above. The total sum of the process of Production in the General Analysis[2] output points to *171736* (milliseconds), which is the actual active writing time. Conversely, the sum of the normal production column in the Revision Matrix indicates *355* (seconds). As seen, both figures are completely different. The main reason lies in the fact that the data from Revision Matrix include all the pausing time of the overall time on task of the writing act, which underlies the different macro-processes (i.e. planning). Hence, the 183.264 seconds that are missing in the General Analysis correspond to the time spent on these processes. This procedure allows us to obtain the *total duration* of the formulation process in a direct and quantifiable way. In fact, formulation does not only involve the production of stretches of text, it is also a process which entails problem-solving and reflective behavior (Manchón & Roca de Larios, 2007). This is well aligned with Roca de Larios et al.'s (2008) coding of formulation as "episodes which indicate that the writer is

[2] As detailed in section 2.7.1. in Chapter 2, General Analysis contains the individual information of each keypress, and does not conflate the stretches of text in terms of *normal production*. Thus, the information on these episodes only appears in the Revision Matrix as such, and has to be collated with the General Analysis output and ideally, with the linear-based analysis S-Notation.

trying to convert thoughts and ideas into language, with or without having to engage in problem solving" (p. 36). In light of this, such episodes may be accompanied by pauses – reflecting this thought process – which supports the inclusion of formulation pauses within the global writer's time allocation to the formulation process. However, as pointed out in the section on the planning process, the lack of information to discern whether or not a long pause refers to a formulation pause is arduous without the support of other data gathering techniques. On this basis, these formulation pauses above 2 seconds were assimilated as global planning processes.

Regarding the *frequency* of formulation episodes, these were counted with the normal production episodes in the Revision Matrix (see Figure 32 below), and the edits were obtained by counting them as per each formulation episode (see column *edits* in the figure below). This way, accessing the writer’s data on formulation seems more unobtrusive and quantifiable. Furthermore, the indication of edits allows us for gaining deeper understanding of “the number of actions needed to complete the action” (Leijten & Van Waes, 2015-2019, p. 89). Finally, the number of words produced (see *words* in Figure 32) were counted per formulation episode, which are also relevant to identify whether the formulation process is fluent or not (Van Waes & Leijten, 2015).

Figure 32.
Revision Matrix in an XSL spreadsheet for the analysis.

Interval	revisio	type	content	edits	start	end	duration	length	beginPos	endPos	chars	charWithoutSpace	words
1	0	Normal Production	one·day·the·scir	21	00:10.891	00:28.854	00:17,963	16	0	16	13	10	4
1	0	Normal Production	entist·	7	00:32.246	00:36.470	00:04,224	7	15	22	6	5	1
1	0	Normal Production	scientist·is·observe·the·experiment·and·he·drimk·the·experiment	64	02:03.222	04:59.287	02:56,065	63	12	75	54	45	10
2	0	Normal Production	nk·the·experiment·and·he·is·conviert·in·a·cat·but·the·dog·is·angry··	68	05:06.455	07:13.543	02:07,088	68	58	126	52	36	15
2	0	Normal Production	and·the·dog·is·attack·the·scientist·and·end	43	07:16.319	07:52.087	00:35,768	43	125	168	35	27	9
2	0	Normal Production	and·c	5	08:18.751	08:20.799	00:02,048	5	161	166	4	3	2
2	0	Normal Production	scientist·is·crazyy	19	08:25.639	08:35.959	00:10,320	19	165	184	17	15	3
3	0	Normal Production	.	8	18:03.448	18:03.568	00:00,120	1	191	192	0	0	0
		9		235		Total dur	374				181		44
						Mean dur	44						

►Analysis of revision

Identification and categorization

The revision process has been defined as an act of revisioning or reseeing the text (Barkaoui, 2007), sometimes not easily distinguishable from formulations (Roca de Larios et al., 2008). Revision as a process involves the evaluation of text product to modify it for linguistic or content reasons. This leads us to the consideration of revision as any change at any point in the writing process (see Alamargot & Chanquoy, 2001; Fitzgerald, 1987; Hayes, 2004; Van Gelderen & Oostdam, 2004), which coincides with Roca de Larios et al.'s (2008) definition as "those episodes in which the writer changes, adds to, or deletes previously written segments of different length" (p. 37). Hence, any change in the text will be assimilated as a revision without specifying whether the reason is a formulation (self-correction) or revision per se.

In this dissertation, the study of the revision processes was observed through a product-oriented lens, centred on the type of changes made by the writer. We operationally defined revision as any changes or modifications made to the text (either at the point of inscription or in any other part of the text). We measured the revision process by indicating the frequency of revision episodes, as in previous research (see Kim, 2020; Lindgren & Sullivan, 2006; Révész et al., 2019), and their total duration in percentages (Roca de Larios et al., 2008). The measurement of *Total revision*[3] included non-temporal measures, such as in the measurement of the formulation process, since their inclusion would provide great insight into cognitive and mechanical aspects. These non-temporal measures included edits (as they allowed us to observe the number of actions needed to complete the revision episode) and words produced (similar to Xu’s [2018] measure of scope, which included letters deleted and produced as proof of the writer's revised stretches of text). Such non-temporal measures were only used for the general process of revision, and not for the micro-processes of revision (i.e. microscopic or macroscopic revisions).

Previous research has categorized revisions as either product or process-oriented taxonomies. Our decision is based on this previous research, which will be clarified further. Product-oriented revisions have been mainly centred on the type of change made. As Lindgren & Sullivan (2006) point out, these are defined according to the operation performed (e.g. insertion, deletion), the syntactic level (e.g. word or sentence), or according to the object of revision (e.g. content or text organisation). For instance, Faigley & Witte’s (1981) taxonomy identified *surface changes* and *text-based changes* according to the inclusion of new information to the text, that is, surface

[3] The reader is reminded that *Total revision* includes all revisions undertaken by the writer without specifying the type of revision.

changes involved the modification of the same information. The super-category of surface changes was further subdivided into other categories (e.g. formal changes, meaning-preserving changes) and text-base changes was divided into micro-level and macro-level changes. This taxonomy bore great resemblance to the one used in Daiute's (1986) and Kehagia & Cox's (1997) empirical research. Children, as low L2 proficient writers, tend to focus on surface-level changes, and prioritize their concerns about revision throughout the process instead of at the end. This is also related to the computer-mediated writing environment (Van Waes & Schellens, 2003).

Quantification or levels of measurement

In our study, revisions were computed and classified from Inputlog output data obtained from the *Revision Matrix (Revision Analysis)*. As described in Chapter 2 and mentioned throughout this chapter thus far, this type of analysis provided information on the revisions made and their levels of measurement (i.e. the revision count, the type of revision [insertion or deletions], the time spent revising as well as the words produced). These measures for the variable *Total Revision*, one of the macro-processes, were obtained automatically from this *Revision Analysis* in Inputlog (see Figure 33 below), but were further classified manually in regard to the type of revision (i.e. microscopic or macroscopic revisions, as will be further detailed). As seen, information on "insertions" and "deletions" are provided.

Figure 33.
Extract from the Revision Matrix (Revision Analysis).

0	Normal Production	batle	8	13:42.476	13:50.700	00:08.224	127	132	5	5	5	1
22	Insertion	·and·g	11	14:05.388	14:07.364	00:01.976	126	132	6	4	2	2
23	Deletion	g	1	14:07.924	14:08.004	00:00.080	132	131	1	1	1	0
24	Insertion	cat·	4	14:08.500	14:09.476	00:00.976	131	135	4	3	2	1

In the following subsection, I explain the operationalization of these micro-processes: both microscopic and macroscopic revisions, and the measures used. As referred to earlier, our study is an attempt to look into how the writing processes are developed throughout the composition time before and after receiving model texts as WCF. Thus, the need to classify revisions according to a specific taxonomy allows us to understand how this process was potentially affected by the provision of model texts. In fact, previous research on the effect of models as WCF, even if mostly based on handwriting environments, has pointed to a focus on lexicon and ideas in the revised versions of the participants' texts (Coyle & Roca de Larios, 2020), which might be reflected in several writing processes, among them revision. Thus, computing the process of revision requires probing into how different these revisions were, that is, microscopic or macroscopic revisions.

Measuring microscopic and macroscopic revisions

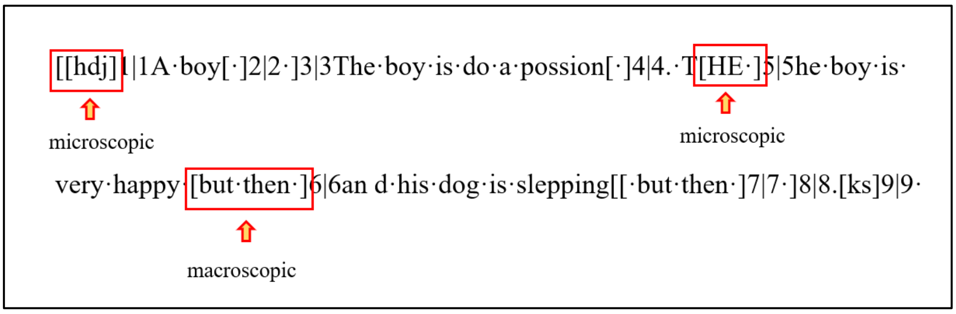
As we have anticipated in the previous section, revision taxonomies are abundant in the scientific literature and we have opted for product-oriented taxonomies focusing on the effect of the revision on the written product. As announced before, numerous studies have used this type of taxonomy (see Bridwell, 1980; Faigley & Witte, 1981; Sullivan et al., 1998; Wengelin, 2002; Chanquoy, 2001, among many others). Bridwell (1980), for instance, operationalized revisions by using a linguistic-based taxonomy focusing on the linguistic unit where the revision was taking place. Given its similarity with Faigley & Witte's (1981) surface-based revisions, we opted for Kim's (2020) taxonomy, which was based on Bridwell's (1980) and precedent taxonomies (Daiute, 1986; Kehagia & Cox, 1997). Our revisions were subsequently classified into *microscopic* and *macroscopic* revisions. As noted by Lindgren & Sullivan (2006), "manual analysis can be a useful complement and is necessary in order to define the content of revisions" (p. 88). Inputlog provides the researcher with very specific data on this type of revisions. Nonetheless, manual analysis is needed for the classification of such revisions, which is what we have done in our study.

Microscopic revisions were studied through the textual level of the revision, and more specifically, whether this revision occurred "at the surface, lexical or phrase level" (Kim, 2020, p. 4). Hence, I classified microscopic revisions as those modifications in the text which were below the phrase level, and resorted to the operationalization provided by Bridwell (1980), and Kim's (2020) recent study. Kim (2020) merged the various categories (surface level, lexical and phrase level) of Bridwell's (1980) revision taxonomy into one single category. Bridwell (1980) devised his categories characterizing "a movement from small to larger linguistic units" (p. 203).

Then, Bridwell (1980) subdivided these linguistic levels further into a series of operations: Surface level contained spelling, abbreviation vs. full form, or morphological conditioning. The lexical level referred to substitution, addition or deletion of words, among others. Finally, the phrase level included substitution, addition, alteration, deletion, but also expansion of word to phrase or reduction of phrase to word (Bridwell, 1980, p. 204). Young EFL learners are reported to concentrate on surface-level changes more than changes at the level of meaning or above phrase level (Luquin & García Mayo, 2021). The attempt to ascertain how children review and modify their texts at these levels makes the inclusion of this type of revision process a necessary condition for our study.

Macroscopic revisions were classified as those text modifications at clause, sentence, and multi-sentence level (Bridwell, 1980). Hence, we operationally define macroscopic revisions as those occurring at those levels, i.e. above phrase level. To illustrate this in a better way, an example of both microscopic and macroscopic revisions is included in Figure 34. As seen in the revised parts of text in brackets, the examples of microscopic revisions are representative of (i) surface level (in the case of THE) and (ii) keyboarding issues (as in the case of the initial revision). Conversely, the two examples on the macroscopic revision "[but·then·]" are understood as above-phrase level. Following Bridwell (1980), sentence level revisions might also refer to a "reduction of sentence to word, phrase or clause (including coordination)" (p. 204). The relevance of macroscopic revisions for the present study is connected with the view that they may reveal whether children are reducing the scope of their revisions or, on the contrary, are editing their texts above the phrase level on a more frequent basis.

Figure 34.
Characterization of the subtypes of revision in an S-Notation output.



The levels of measurement for macroscopic and microscopic revisions will be equal to the ones part of the global revision process (i.e. episode frequency and total duration in percentages), except for the non-temporal measures (i.e. edits and words produced). We decided not to include these last measures since the key concern in these subprocesses is chiefly placed on ascertaining the types of revision and their temporal allocation within the continuum of the writing process. Hence, the non-temporal measures will be globally considered in the macroprocess of revision and not in these subprocesses.

4.4.1.2. Pausological behavior: constructs and measures

The study of writing processes as proposed, operationalized and measured in the previous sections may give us a more direct perspective of how these processes develop throughout the continuum of the composition time. The construct of "pausological behavior" has been primarily operationalized as the frequency, the location and the duration of such pauses given that "depending on where and how long writers pause, pauses are likely to signal differential underlying processes" (Révész et al., 2019, p. 608). Hence, the construct of planning was measured taking into account pauses above the pause threshold of 2000 milliseconds (Wengelin et al., 2007). However, in the case of the rest of the processes (i.e. formulation or revision), the time allocated was used including pauses without bearing in mind the 2000-millisecond pause threshold, which include globally action and pause time. Our study intends to provide a complete analysis of the pausological behavior of children learners, a population scarcely studied in terms of pausing research. The pause threshold in this study has been profusely used for capturing cognitive processes such as planning or revision in previous research (Chukharev-Hudilainen, 2014; Medimorec & Risko, 2017; Tiryakioglu et al., 2019; Wengelin, 2007; Wengelin et al., 2009). Furthermore, the inclusion of pausological behavior provides interesting data to supplement the one classified manually in the writing processes of planning and revision.

The following subsections will respectively deal with the methodological actions undertaken regarding pause location, frequency, and duration.

►***Pause location***

Another measure computed using output from Inputlog was pause location. Pause location refers to the exact linguistic location of the pause (i.e. before words, between words, within words, before sentences, among others). To supplement pause frequency, pause location was computed alongside. Previous empirical evidence (Medimorec & Risko, 2017; Révész et al., 2017; Dean & Zhang, 2015) has suggested that pause location might be revealing as to where pause frequency occurs, possibly unveiling underlying cognitive and writing processes. It has been suggested that pauses before words, which are one of the most difficult to interpret (Conjin et al., 2019), might reflect anticipatory or reflective operations, such as word choice (Van Waes et al., 2016). Thus, the operationalization of pause location refers to where the pause occurred in combination with its frequency, duration, and the ratio of pause location (Barkaoui, 2019). Previous studies (Baaijen et al., 2012; Conjin et al., 2019; Barkaoui, 2019; McArthur & Graham, 2016; Révész et al., 2017; Van Waes et al., 2016; Van Waes et al., 2014), have contributed to understanding how pause location has helped discern patterns in the conceptualization of the writing processes and other cognitive operations. For our current purposes, pause location (see Coding categories below) was highlighted for the EG and CG by indicating the number and mean duration for each location of pauses. Such a general vision allows us to establish comparisons and detect dissimilarities between participants' original texts and the revised text at stage 3. The scales of measurement used for pause location included pause frequency, the mean duration, and the location ratio (Barkaoui, 2019). All of these measures have been successfully used in research with adult and teenagers L2 writing, and none have been empirically tested with children writers in digital writing. Despite the absence of this type of empirical research with children, I will attempt to compare what has been the norm in pausological research with adult and adolescents, and explore what tendencies reveal. The use of these measures is, hence, appropriate for responding to our RQ2.

In what follows, the coding categories are visually presented from the Inputlog data output from the pause-based Linear analysis. As seen, the text is visually represented in a linear fashion, including not only the duration of pauses and text production, but also the keyboard presses or cursor movements performed.

<u>*Coding categories*</u>

(1) {numbers} indicates the pause duration; (2) [cursor movement/key press or release] indicates the specific movement or key used, for instance, BACK refers to deletion, CAPS LOCK to pressing on the capital letter key (Leijten & Van Waes, 2015-2019).

- Within word

Example 1. Within-word pauses in bold.

"{7945}.·[CAPS LOCK]T[CAPS LOCK]he·dog·s**{2008}**aw"

- Before and after word

Example 2. Before and after word pauses in bold.

"{3968}the·pocion**·{4168}**and**·{10385}**he·ch[BACK][BACK]{2152}explosin{6857}"

- Before and after sentences

Example 3. Before and after sentences pauses in bold.

"bea[BACK]cause·the·sciente[BACK]{4120}ist·transform·to·{5008}a·cat**{39797}** | [Movement][LEFT Click][RETURN][Movement]**]{136183}[**CAPS LOCK]T[CAPS LOCK]he ·scient{3192}ist·{6184}start·to·have·{2184}"

- Before and after paragraphs

Example 4. Before and after paragraphs pauses in bold.

"**{2216}**tact **·{12080}[**BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK].[RETURN]{43272}[CAPS LOCK]T[CAPS LOCK]he·dog·[Movement]{2153}atact·to·{2487}s ·c[BACK][BACK]cientist·"

- Other boundaries (i.e. *other pauses*).

Example 5. Other pauses in bold.

[BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK][BACK].[RETURN]**{43272}[**CAPS LOCK]T[CAPS LOCK]he·dog ·[Movement]{2153}atact·to·{2487}s·c[BACK][BACK]cientist·"

►***Pause frequency***

Pause frequency was computed as the number of pauses within a specific textual location. This level of measurement has been used in previous research where KSL software was used (e.g. Barkaoui, 2019; Medimorec & Risko, 2017; Van Weijen et al., 2008; Xu & Qi, 2017). The main rationale behind the use of pause frequency is rooted in the invaluable information that it offers. It gives a global idea of the writers' specific recurrence to pausing, indicating how much a writer has had to interrupt the writing process.

By means of the output data provided in the *Pause Logging File (Pause Analysis)* from Inputlog, the following information about pauses was obtained. Not only pause frequency but also other valuable measures to account for this pausing behavior have been extracted: pauses per keystroke, pauses per minute and interkeystroke interval. The first, pauses per keystroke, was manually computed by dividing the total number of pauses by the total number of keystrokes (Barkaoui, 2019; Wengelin, 2006). The second measure, pauses per minute, was computed by dividing the total number of pauses by the length of the writing session in minutes (Barkaoui, 2019). Van Waes & Leijten (2015) used in their study the total number of pauses and this ratio of pauses per minute as one of the essential fluency measures to be considered in their multidimensional approach to fluency. The third measure, *interkeystroke intervals*, was automatically computed by dividing the total number of keystrokes by the number of pauses (cf. Conijn et al., 2019; Vandermeulen et al., 2020). Van Waes & Leijten (2015) considered this measure as "interkey-transition times or pauses" (p. 86).

In our study, the inclusion of fluency measures responds to the need to examine the amount of cognitive load before and after receiving WCF. Firstly, this cognitive load is reflected in the different writing processes, previously operationalized, and pausological behavior. In fact, "fluent writing processes are characterized by short pausing times, few revisions and a high production rate (MacArthur et al., 2008)" (Van Waes & Leijten, 2015, p. 80), which alludes to the importance of considering pausing times, but also how these production is shaped by these pauses. Fluency measures helped us discern this amount of cognitive load. Nonetheless, fluency, in its multiple definitions, has been characterized differntly depending on what it represents. Transcription fluency, for instance, is considered as execution by Kellogg (1994), and is strongly related to motor movement (either handwriting or typing). Conversely, fluency is related to the access to linguistic knowledge (Van Gelderen & Oostdam, 2002), and is linked to the formulation process, but also, to the cognitive load. In this sense, more and longer pauses may result in more cognitive load, and hence, less fluency. In sum, fluency measures might give us insight into how children's fluency might develop, hinting at their cognitive effort with the measure of *pauses per minute*, and how they write motorically, with the measure of *interkeystroke interval*.

►***Pause duration***

Pause duration was computed as *mean duration* directly from the Inputlog output data. This was calculated for pause location (i.e. the mean duration of pauses within-words) and for pause frequency (i.e. the mean duration of pauses at a given interval during the writing process). The conceptualization of pause duration is one of paramount relevance for the writing process since it reflects the cognitive effort deployed by the writer not only in quantity but also in terms of *where* this occurred (Hoang, 2019). Building on research in the literature (Alamargot et al., 2007; Barkoui, 2019), I have accepted the assumption that both pause duration and its frequency may shed light on whether a writer is engaging in high or low-level cognitive operations, as was announced in the previous sections on the operationalization of writing processes. For instance, pause duration within word may be indicative of either keyboard text production (Grabowski, 2008), or conversely of spelling issues. Apart from referring to cognitive operations in relation to the writing process, the duration of a pause might mirror a reflection or deliberation (Conijn et al., 2019). In summary, the operationalisation of pause duration will be reflected in the analyses by means of *mean duration* in relation to pause location, and also the interval in which these pauses took place. Such measures allow for gaining more insight to identify the processing load carried by such operations. In other words, whether these pauses might reflect high- or low-level operations.

In addition to this, I also took into account the relevance of *process time* measures such as *total process time*, which entails the length of the writing process. Similarly, *active writing time* was also considered given the information that it provides on how much time the writer devoted, on its whole, to producing stretches of text. Finally, *pausing time* was considered as a measure indicating the total time writers paused throughout the composing process.

4.4.2. Intra-rater reliability

Six months after the initial data analysis, 33% of the data for the measures that required manual calculations were recoded, which included planning (local and total planning), formulation and revision (microscopic and macroscopic revision). As can be seen in Table 6, the obtained Cohen's *kappa* coefficients indicated a satisfactory degree (90%) of intra-rater reliability (M= .900, ASE= .013, range= .888-.925).

Table 6.
Cohen's kappa intra-rater reliability coefficients for planning, formulation and revision.

	Writing 1		Writing 2	
Planning	N= 139	.834	N= 136	.855
Formulation	N= 140	.819	N= 220	.848
Revision	N= 272	.935	N= 114	.859

4.4.3. Summary of the measures used in this study

Tables 7 below summarizes the different variables and subvariables for each writing process. As mentioned, each variable corresponding to the planning and revision processes was multidimensional. In the case of formulation, no subvariables were present. Several measures were used for each writing process.

Table 7.
Summary of the writing processes constructs and measures used in the study.

Planning		**Formulation**	**Revision**		
Local	***Global***			***Macroscopic***	***Microscopic***
■Frequency of episodes (raw numbers) ■Total duration (percentages)		■Frequency of episodes (raw numbers) ■Total duration (percentages) ■Words produced (numerical) ■Edits (raw numbers)	■Frequency of revision episodes (raw numbers) ■Total duration (percentages) ■Words produced (raw numbers) ■Edits (raw numbers)	■Frequency of revision episodes (raw numbers) ■Total duration (percentages)	

Table 8 summarizes the construct of pausological behavior into "pauses" and "pausing behavior". In the case of the former, a series of measures are provided to account for the number of pauses, the time-on-task, as well as two variables for fluency (one referring to mechanical aspects, and the other to temporal ones). Regarding "pausing behavior", pauses were classified according to the ratio, their frequency, where they appeared and their duration. These measures chosen for the construct of pausological behavior are objective, allowing for a quantification and direct interpretation. Similarly, these objective measures might ease the comparison with other studies on pausological behavior in digital environments.

Table 8.
Summary of the pausological behavior measures used in the study.

Pauses	**Pausing behavior**			
■Number of pauses (numerical) ■Mean duration (numerical) ■Interkeystroke interval (miliseconds). ■Pauses per minute (nº of pauses/length of writing in minutes). ■Total process time (seconds). ■Total pausing time (seconds). ■Total active writing time (seconds).	**■Pause ratio**	**■Pause Frequency**	**■Pause Location**	**■Pause duration**
	Nº of pauses/PauseLocation_type	Nº of pauses at a given location	The location of the pause (e.g. within words, before sentences).	Mean duration of the pauses at a given location

*The general quantitative measures of pauses "Interkeystroke interval (milliseconds)" and "Pauses per minute (nº of pauses/length of writing in minutes) belong to the construct of fluency (Van Waes & Leijten, 2015).

4.4.4. Statistical analyses

Our research questions answered through a number of descriptive and inferential statistical analysis on the different variables associated with this piece of research. A summary of the different research questions, constructs, measures and statistical analyses is displayed in Table 9 below.

Different statistical analyses are carried out with distinct values from the pausing and revision behavior obtained out of Inputlog data. In light of the small sample size (n= 18), using statistical analyses to compare pre and post-tests as well as to establish between-subjects comparisons was not appropriate as they would not yield a stable result. In addition to this, inferential statistics are largely dependent upon sample size, especially p values (Larson-Hall & Plonsky, 2015). When an extrapolation or generalization of the results is not possible, and thus the objective of the study is exploratory and centered on the results for the sample, significance testing may not be used. This is why effect sizes are regarded, in these cases, as sufficient and suitable (Neill, 2008). Thus, after careful consideration and consultation of experts in the field, we opted for descriptive statistics to address our research questions. Such descriptives were based on the measure of central tendency as the *mean*, the measure of dispersion as *standard deviation*, and the effect size as expressed with *Hedges' g* and *Hedges'* $g_{adjusted}$.

The main reason behind using Hedges's *g* is rooted in the fact that this effect size is "multiplied by a correction factor for small sample sizes" (Turner & Bernard, 2006, p. 45). To interpret the magnitude of the effect, we will follow the benchmarks proposed by Plonsky & Oswald (2014) for the field of L2 research: between-groups (small: .40; medium: .70; large: 1.0) and within-groups (small: .60; medium: 1.0; large: 1.4). Similarly, the benchmarks indicated for correlations (rs) by the same authors will be taken into consideration: small (.25), medium (.40), and large (.60).

Table 9.
Statistical analyses.

Analysis	RQs addressed	Between-subject factor(s)	Construct > Dependent variables	Statistical tests
1. The temporal dimension of writing processes	RQ 1. To what extent are there differences in young EFL learners' temporal distribution of writing processes (i.e. planning, formulation, and revision) when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?	Written Corrective Feedback: GE: provision of WCF GC: absence of WCF	━Planning > Total planning, Global and Local planning. ━Formulation > Formulation ━Revision > Total revision, miscroscopic and macroscopic revision.	Descriptive statistics + effect size (Hedges's *G and G adjusted*)
2. Pausing behavior in the temporal dimension of Writing Processes	RQ 2. To what extent are there differences in young EFL learners' pausological behaviour when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?	Written Corrective Feedback: GE: provision of WCF GC: absence of WCF	━Pauses & Fluency > Nº of pauses, pauses per minute, interkeystroke interval, total process time, total pausing time and total active writing time. ━Pausing behavior > Pause ratio, pause frequency, pause location and pause duration.	Descriptive statistics + effect size (Hedges's *G and G adjusted*)

CHAPTER 5. RESULTS

In an attempt to answer the research questions guiding our research, the present chapter is organized into two different kinds of analyses: quantitative and qualitative. The quantitative analyses account for the data regarding the (1) temporal dimension of writing processes, that is, the comparison of writing processes between Writing 1 and 2 as mediated by models as WCF (RQ 1), and (2) the data on the pausing behavior related to pause location and duration as well as other quantitative pausological measures as mediated by models as WCF (RQ 2).

5.1. The influence of models as WCF on the temporal dimension of writing processes

►*RQ 1. To what extent are there differences in young EFL learners' temporal distribution of writing processes (i.e. planning, formulation, and revision) when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?*

5.1.1. Overview of the global writing processes

Table 10.
Descriptive statistics for global writing processes by group and writing.

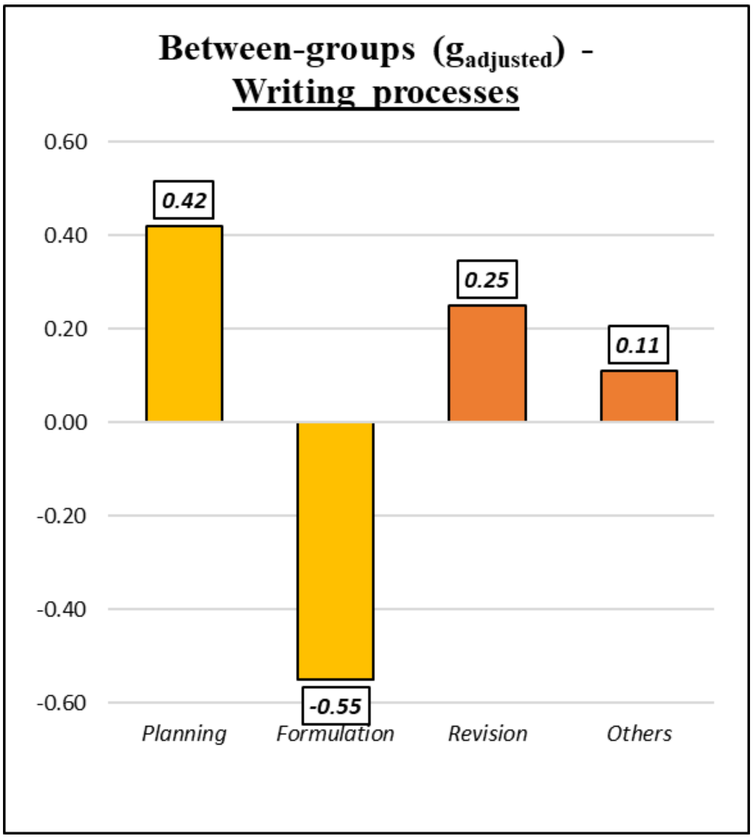
		EG					CG				
		Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
		M	SD	M	SD	*g*	M	SD	M	SD	*g*
Planning	Ratio	.18	.05	.20	.12	**–0.21**	.27	.20	.22	.04	**0.33**
Formulation	Ratio	.47	.09	.53	.09	**0.64**	.50	.13	.60	.07	**0.91**
Revision	Ratio	.09	.12	.07	.05	**0.21**	.08	.04	.05	.06	**0.56**

Note: ES stands for Hedges' G effect size.

We will first account for the temporal distribution of macro writing processes, i.e. planning, formulatiion and revision, across writings and groups.

Table 10 shows that that the EG increased, albeit slightly, the planning time while the CG decreased it although the effect sizes are trivial. In the case of the formulation process, both groups similarly increased the amount of formulation, although such an increase was more considerable in the CG. The differences are relevant for both groups, with a small effect for the EG (g= 0.64) and small-to-medium for the CG (g= 0.91). Regarding revision, both groups decreased the time devoted to that process, although the effect size is trivial.

Figure 35.
Between-group comparison effect size (Hedges' g) for the ratio of macro writing processes.



Regarding the differences between groups, Figure 35 above shows Hedges' $g_{adjusted}$ for the ratio of the three writing processes. As can be seen, the processes which have been more globally and potentially affected by the independent variable are planning and formulation. Nonetheless, the effect size is trivial (g= 0.42 and –0.55, respectively) but at a small magnitude according to the benchmarks proposed in the literature (see Loewen & Plonsky, 2015).

In what follows, the data will be compared within-groups and between-groups through the means and SDs. To determine the magnitude of the effect of the independent variable, the data will be compared with the effect size (cf. Loewen & Plonsky, 2015; Plonsky & Oswald, 2014). This is intended to help us discern the potential impact of the provision (or lack of) of WCF on the different writing processes and subprocesses. The same procedure was then followed for a between-groups comparison to evaluate potential differences between EG and CG as a function of the WCF provided. In this case, we use $g_{adjusted}$ to compensate for the initial dissimilarities in the data in Writing 1.

5.1.2. Planning

Table 11, 12 and 13 report the descriptives for measures (frequency and total duration) of total planning episodes, global planning, and local planning, respectively.

Table 11.
Descriptive statistics and effect sizes for total planning measures by group and writing.

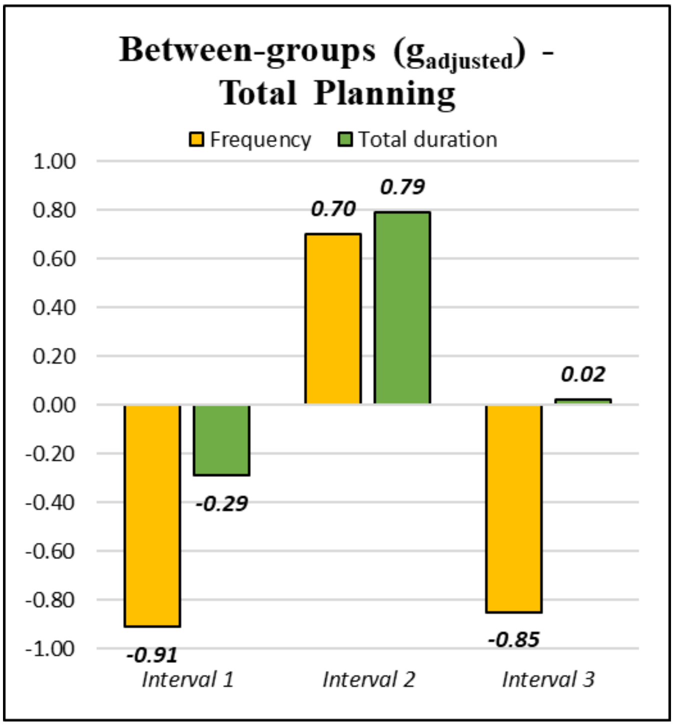
			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Total Planning	Frequency	Interval 1	14.30	6.96	5.70	8.50	**–1.06**	7.62	3.42	4.25	2.25	**–1.10**
		Interval 2	9.80	5.34	8.00	4.05	**–0.36**	6.75	3.91	3.62	1.40	**–1.01**
		Interval 3	9.10	8.55	4.30	3.16	**–0.71**	2.25	2.49	3.87	2.69	**0.59**
	Total duration	Interval 1	.086	.040	.076	.029	**–0.27**	.14	.16	.10	.03	**–0.33**
		Interval 2	.055	.025	.068	.046	**0.33**	.09	.05	.069	.037	**–0.42**
		Interval 3	.048	.041	.066	.077	**0.28**	.043	.053	.058	.039	**0.30**

As can be seen in Table 11, descriptives for the frequency of *Total Planning* reveal a major tendency towards decreasing the amount of planning for both EG and CG. In the first interval, the EG decreased the frequency of planning, as evidenced in the medium-to-large effect size (g= –1.06), though variations in Writing 2 are to be noted owing to the large SD (=8.50). Likewise, the CG decreased the amount of planning with an equally medium-to-large effect size (g= 1.10). In terms of total duration of planning episodes, a slight decrease of time spent on planning was observed in the data of both EG and CG, but the effect size was trivial.

In the second interval, the CG decreased the frequency of planning, as can be attested by the large effect size (g= –1.01). The EG, however, decreased very slightly the frequency of planning episodes, although this difference from Time 1 to Time 2 is trivial given the very small effect size (g= –0.36). In terms of total duration, the EG increased the total time spent on planning although this difference was trivial (g= 0.33), while the CG decreased the time spent on planning, although also in a trivial manner (g= –0.42).

Finally, the third interval reveals opposing tendencies for both groups: While the EG experienced a substantial decrease in frequency supported by the observed small-to-medium effect size (g= –0.71), the CG increased the amount of planning slightly, although the SDs (Writing 1: 2.49 and Writing 2: 2.69) point to a large variation among writers, which is trivial (g= 0.59). Regarding total duration in the planning process, both groups increased the time spent in equal terms, but with trivial effect sizes (EG: g= 0.28, and CG: g= 0.30).

Figure 36.
Between-group comparison with the effect size for frequency and total duration of Total planning.



Finally, to discern the purported effect of WCF, we observed the effect size ($g_{adjusted}$) of the frequency and total duration of total planning in Writing 2 between EG and CG. As can be seen in Figure 36, results point to non-trivial effect sizes in the first interval (g= –0.91) where the EG maintained a higher frequency of planning episodes than the CG (EG: 5.70; CG: 4.25). In terms of total duration, the differences between both groups are trivial ($g_{adjusted}$ = –0.29). Regarding the second interval, the medium effect size ($g_{adjusted}$ = 0.70) indicates that the EG planned more frequently than the CG (EG: 8.00; CG: 3.62). As for total duration, the medium effect size ($g_{adjusted}$ = 0.79) points to greater differences between both EG and CG. Although the values are similar (.068 and .069, respectively), the within-group differences indicate that there was a very sharp decrease in the CG (Time 1= .090, Time 2= .069). In the third interval, there is a medium effect size ($g_{adjusted}$ = –0.85) in the difference between both groups since, even if the CG increased slightly the frequency of planning (Time 1= 2.25, Time 2= 3.87), and the EG decreased it (Time 1= .043, Time 2= .058), the values of the latter were higher than those of the former. In the case of total duration, the descriptives show that the EG devoted more time to planning than the CG, which also increased planning at a lower extent. Despite these data, the effect size was trivial ($g_{adjusted}$ = 0.02).

We next report the data on the subprocesses of planning, that is, global planning and local planning.

5.1.2.1. Global planning

Table 12.
Descriptive statistics and effect size for global planning by group and writing.

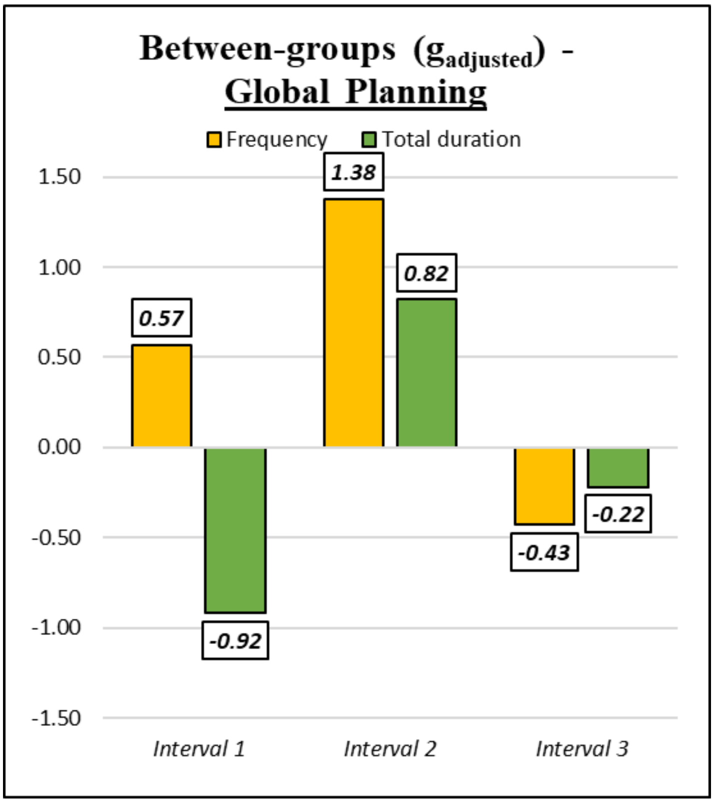
			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Global Planning	Frequency	Interval 1	7.30	2.62	4.10	1.66	**1.40**	5.00	1.92	1.87	.99	**1.94**
		Interval 2	5.00	3.09	3.40	1.17	**0.66**	3.75	2.25	1.37	.92	**1.31**
		Interval 3	3.20	3.76	1.60	1.26	**0.55**	1.00	1.19	1.25	1.03	**0.21**
	Total duration	Interval 1	.064	.03	.046	.02	**0.57**	.12	.17	.080	.03	**0.30**
		Interval 2	.039	.02	.039	.03	**0.01**	.060	.05	.038	.02	**0.75**
		Interval 3	.024	.03	.039	.05	**–0.35**	.016	.02	.035	.03	**–0.65**

Table 12 above displays the descriptives and effect sizes for the subprocess of global planning. Regarding frequency, both groups display a downward trend from the beginning to the end of the composition process. In the first interval, the EG resorted less frequently to planning (Time 1= 7.30, Time 2: 4.10), as evidenced in the large effect size observed (g= 1.40). Similarly, the CG decreased substantially the frequency of global planning episodes (Time 1= 5.00, Time 2= 1.87), with a very large effect size (g= 1.94). Regarding the total duration of global planning, the EG decreased the time spent on planning globally (Time 1: .024, Time 2: .039) with a small effect size (g= –0.57). The CG decreased the total duration of global planning substantially as can be seen in the descriptives in Table 12, although the effect size is small (g= 0.30), probably as a result of the large variation among participants observed at Writing 1 (SD= .17).

In the second interval, the EG resorted slightly less frequently to global planning episodes (Time 1: 5.00, Time 2: 3.40) with a small effect size (g= –0.66) after the use of WCF, while the CG reduced planning episodes to a considerable extent (Time 1: 3.75, Time 2: 1.75), maintaining the large effect size (g= –1.31), as in the previous interval. As regards the total duration of planning time, the EG maintained the time spent on global planning (Time 1: .039, Time 2: .039) while the CG decreased it by 50% (Time 1: .060, Time 2: .038), with a small-to-medium effect size (g= –0.75).

Finally, in the third interval, trivial differences were observed in any of the two groups in terms of global planning episodes. Nonetheless, the descriptive statistics show that the EG decreased the frequency of global planning episodes (Time 1: 3.20, Time 2: 1.60), while the CG maintained the same values (Time 1: 1.00, Time 2: 1.25). In terms of total duration, the EG increased substantially the time spent on global planning (Time 1: .024, Time 2: .039). The effect size, however, was very small (g= 0.35). The CG increased the amount of time allocated to global planning with a small effect size (g= 0.65).

Figure 37.
Between-group comparison with the effect size for frequency and total duration of global planning.



To discern the purported effect of the WCF, the means of Writing 2 were compared for both EG and CG (i.e. the between-groups comparison). As evidenced by Hedges' $g_{adjusted}$ (see Figure 37 above), a small effect size ($g_{adjusted}$ = 0.57) was observed for the frequency of global planning in the first interval, where the EG kept higher values (EG: 4.10, CG: 1.87). In the second interval, the EG planned more frequently than the CG (EG: 3.40, CG: 1.37) with a medium effect size ($g_{adjusted}$ = –0.82). With regards to interval 3, a very small effect size was found between both EG and CG ($g_{adjusted}$ = –0.43), as they kept similar values (EG: 1.60, CG: 1.25). Regarding the total duration of global planning, the greatest differences were observed in the first interval with the CG devoting more time to planning globally than the EG (EG: .046, CG: .080). In the second interval, the effect size was large ($g_{adjusted}$ = 1.38) although both the EG and CG had similar means (EG= .039, CG= .038). The main reason behind this large effect size is that, when looking at the different within-groups, the CG reduced the time devoted to global planning in the second interval (Time 1= .060, Time 2= .038). In the third interval, the effect size was trivial ($g_{adjusted}$ = –0.22), and both the EG and CG devoted similar amounts of time to global planning (EG= .039, CG= .035).

5.1.2.2. Local planning

Table 13.
Descriptive statistics and effect sizes for local planning by group and writing.

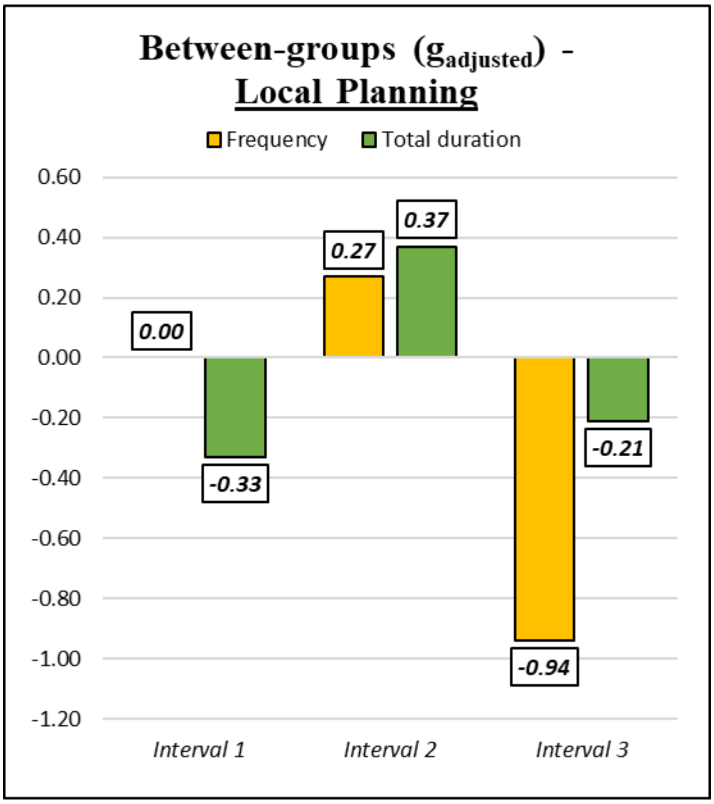
			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Local Planning	Frequency	Interval 1	7.00	4.64	5.30	2.94	**0.42**	2.62	1.92	2.37	1.68	**0.13**
		Interval 2	4.80	2.82	4.60	2.95	**0.07**	3.00	1.85	2.25	1.03	**0.47**
		Interval 3	5.90	5.82	2.70	2.31	**–0.69**	1.25	1.83	2.62	1.99	**0.68**
	Total duration	Interval 1	.023	.01	.028	.01	**–0.37**	.017	.01	.026	.01	**–0.94**
		Interval 2	.017	.01	.029	.02	**–0.68**	.020	.01	.027	.02	**–0.42**
		Interval 3	.023	.03	.027	.05	**–0.09**	.027	.05	.039	.02	**–0.29**

Table 13 above displays the descriptive statistics for the subprocess of local planning. In the first interval, the effect sizes in both the EG (g= 0.42) and CG (g= 0.13) for the frequency of local planning were trivial. In the second interval, the effect sizes were trivial as well for either the EG (g= 0.07) or the CG (g= 0.47). In the third interval, however, the EG decreased the frequency of local planning (Time 1= 5.90, Time 2= 2.70) with a small effect size (g= –0.69). There was great variation among the participants as observed in the SD (Time 1= 5.82, Time 2=

2.31). In the case of the CG, a slight increase was observed (Time 1= 1.25, Time 2= 2.62) with a small effect size (g= 0.68).

Regarding the total duration of local planning, in the first interval, the EG increased the time spent on this process (Time 1= .023, Time 2= .028) but the effect size was trivial (g= –0.37). The CG increased the time spent on this process (Time 1= .017, Time 2= .026) with a medium effect size (g= –0.94). In the second interval, however, the EG increased the time devoted to local planning (Time 1= .017, Time 2= .029) with a small effect size (g= –0.68). The CG increased it as well (Time 1= .020, Time 2= .027) but the effect size was trivial (g= –0.42). In the third interval, no differences were observed either for the EG (Time 1= .023, Time 2= .027) with a trivial effect size (g= –0.09) or the CG (Time 1= .027, Time 2= .039) with a trivial effect size too (g= –0.29).

Figure 38.
Between-group comparison of frequency and total duration of local planning.



To discern the effect of the WCF, the frequency and total duration of local planning in Writing 2 were compared and the effect size (Hedges' $g_{adjusted}$) was computed (see Figure 38 above). In the first interval, the effect sizes were trivial either for the frequency or total duration of local planning. In the second interval, there were trivial effect sizes for either of the two measures. In the third interval, the effect size for the frequency of local planning was nearly large ($g_{adjusted}$ = –0.94), where the EG planned less frequently than the CG (EG= .027, CG= .039). In the case of the total duration of local planning in this interval, the effect size was trivial ($g_{adjusted}$ = –0.21). Both groups, however, increased the total time devoted to local planning (EG: Time 1= .023, Time 2= .027; CG: Time 1= .027, Time 2= .039).

5.1.3. Formulation

Our first research question asked whether there was any difference on the temporal distribution of writing processes (i.e. planning, **formulation**, and revision) with and without the provision of WCF. Thus, Table 14 presents data on the frequency and total duration of the **formulation** process in terms of means, standard deviation (SD), and Hedges' *g* effect size (ES).

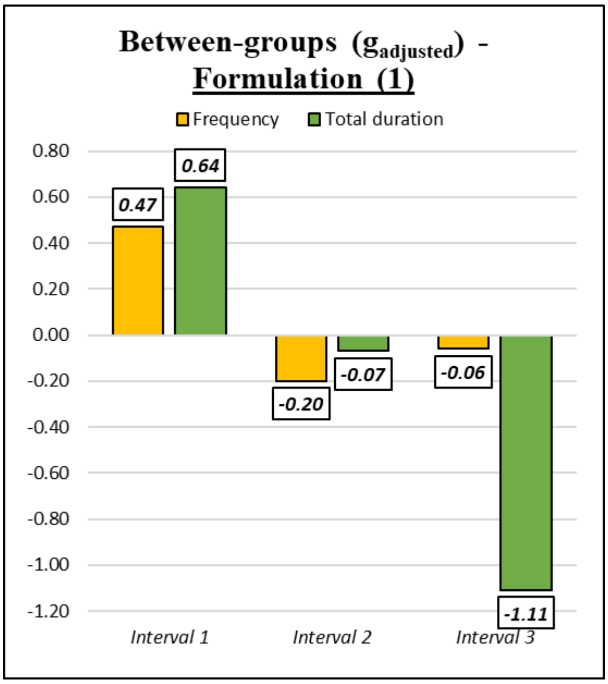
Table 14.
Descriptive statistics and effect sizes for frequency and total duration of formulation by group and writing.

			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Formulation	Frequency	Interval 1	20.00	12.49	17.80	13.56	**0.16**	11.37	4.95	3.87	1.35	**1.95**
		Interval 2	18.30	11.59	16.00	13.52	**0.17**	7.62	3.70	6.00	2.50	**0.48**
		Interval 3	13.70	9.73	10.50	8.73	**0.33**	7.12	6.35	5.25	4.83	**0.31**
	Total duration	Interval 1	.18	.06	.23	.04	**–0.84**	.234	.06	.236	.07	**–0.04**
		Interval 2	.16	.06	.19	.05	**–0.37**	.20	.10	.22	.10	**–0.23**
		Interval 3	.11	.07	.10	.07	**0.14**	.06	.07	.13	.06	**–0.98**

In the first interval, both groups decreased the frequency of their formulation processes. The effect size for the EG was very small (g= 0.16). The CG decreased the frequency of formulation processes (Time 1= 11.37, Time 2= 3.87) with a very large effect size (g= 1.95). In the second and third intervals, there were no effects for the frequency of formulation processes were yielded as observed in the effect sizes in both groups (Interval 2: EG: g= 0.17, CG: g= 0.48; Interval 3: EG: g= 0.33, CG: g= 0.31), which were trivial. It must be noted that the EG displayed great variability across all intervals for the frequency of formulation as observed in the SDs (see

Table 14 above). Regarding the total duration of the formulation process, in the first interval, the EG increased it (Time 1= .18, Time 2= .23) with a medium effect size (g= –0.84). The CG devoted a similar amount of time (Time 1= .234, Time 2= .236) and the effect size was trivial (g= –0.04). In the second interval, no difference were found for either the EG (Time 1= .16, Time 2= .19) or the CG (Time 1= .20, Time 2= .22), with both displaying trivial effect sizes (EG: g= –0.37, CG: g= –0.23). In the third interval, the EG devoted a similar amount of time to formulation (Time 1= .11, Time 2= .10) with a trivial effect size (g= 0.14), while the CG increased it (Time 1= .06, Time 2= .13) with a medium effect size (g= –0.98).

Figure 39.
Between-group comparison with the effect size for the frequency and total duration of formulation.



Regarding the between-group comparison between EG and CG in Writing 2 (see Figure 39 above), in the first interval, we found that the frequency of formulation in the EG was higher than in the CG (EG= 17.80, CG= 3.87) with a small effect size (g= 0.47). In the second and third intervals, the effect sizes were trivial (Interval 2: g= –0.20, Interval 3: g= –0.06), although the descriptives show that the EG resorted to the formulation process more frequently than the CG. As for the total duration of the formulation process, we found that both groups devoted similar amounts of time to this process (EG= .23, CG= .236) with a nearly medium effect size (g= 0.64).

In the second interval, there were no differences between both groups. In the third interval, we found that the CG devoted more time to formulation than the EG (EG= .10, CG= .13) with a large effect size (g= –1.11).

Table 15.
Descriptive statistics and effect sizes for edits and words produced during formulation by group and writing.

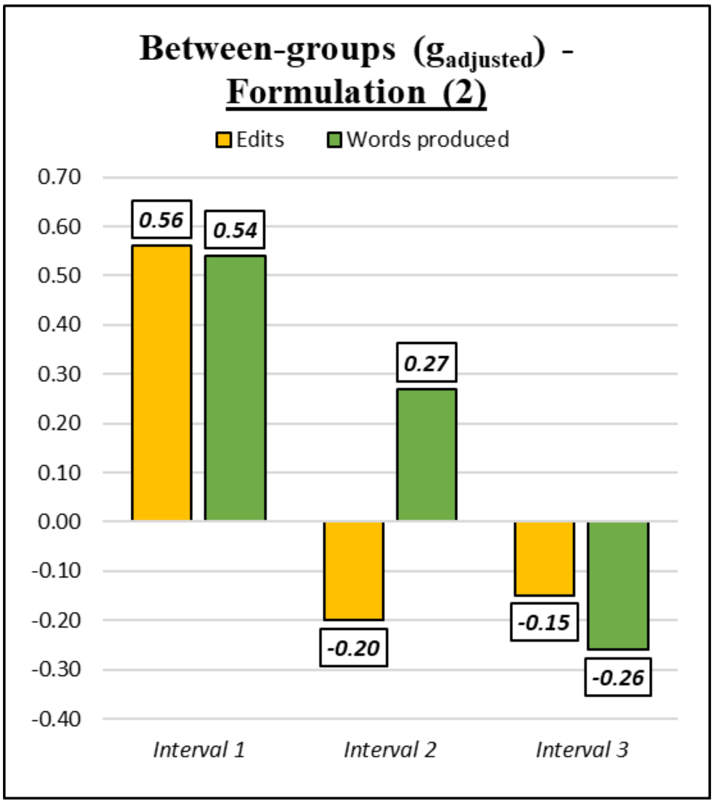
			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	g	M	SD	M	SD	g
Formulation	Edits	Interval 1	307.40	185.52	320.80	214.67	**–0.06**	185.00	73.19	89.62	32.79	**1.59**
		Interval 2	292.20	186.78	280.90	220.25	**0.05**	101.10	73.22	98.87	48.12	**0.03**
		Interval 3	229.50	172.81	191.10	160.65	**0.22**	92.87	81.72	88.25	52.71	**0.06**
	Words produced	Interval 1	62.40	48.61	64.60	48.61	**–0.04**	34.37	14.53	16.50	7.80	**1.45**
		Interval 2	71.10	70.32	55.30	51.98	**0.24**	37.00	29.18	20.00	8.84	**0.75**
		Interval 3	69.70	83.93	50.90	67.32	**0.24**	19.75	21.99	23.87	28.45	**–0.15**

Table 15 above displays the descriptive statistics (i.e. means and SDs) and effect sizes (Hedges' *g*) for the non-temporal measures, that is, edits and words produced. Both groups are characterized by a decrease in edits with different magnitudes in the first interval. We found that the CG decreased the number of edits (Time 1= 185.00, Time 2= 89.62) with a very large effect size (g= 1.59). In the second interval, no differences were observed for either the EG (Time 1= 292.20, Time 2= 280.90) or the CG (Time 1= 101.10, Time 2= 98.87) with trivial effect sizes (EG: g= 0.05, CG: g= 0.03). In the third interval, the effect sizes were trivial (EG: g= 0.22, CG: g= 0.06). It is worth mentioning, however, the great variability existing in the EG (Time 2: SD= 160.65).

In terms of words produced, in the first interval, we found that the EG maintained a similar number of words during formulation with a trivial effect size (g= –0.04). Conversely, the CG decreased the number of words during this process (Time 1= 34.37, Time 2= 16.50) with a very large effect size (g= 1.45). In the second interval, we found that the CG decreased the number of words produced during the formulation process (Time 1= 37.00, Time 2= 20.00) with a small-to-medium effect size (g= 0.75). In the third interval, the effect sizes were trivial for both EG (g= 0.24) and the CG (g= –0.15). It should be mentioned that there was great variability in both the

EG (Time 1: SD= 83.93, Time 2: SD= 67.32) and the CG (Time 1: SD= 21.99, Time 2: SD= 28.45) for words produced.

Figure 40.
Between-group comparison with the effect size for edits and words produced of formulation.



We next report on the between-groups comparison for the edits and words produced in the formulation process. As regards the edits, in the first interval, we found that the EG had higher edits than the CG (EG= 320.80, CG= 89.62) with a small effect size ($g_{adjusted}$ = 0.56). In the second and third intervals, the effect sizes were trivial (Interval 2: $g_{adjusted}$= –0.20, Interval 3: $g_{adjusted}$ = –0.15). Despite the triviality of the between-groups comparison in these intervals, as can be seen in Table 15, the descriptive statistics shows that the EG performed more edits than the CG.

5.1.4. Revision

Our first research question asked whether there was any difference on the temporal distribution of writing processes (i.e. planning, formulation, and **revision**) with and without the provision of WCF. Thus, this section will present the data on the process of revision, and the types of revision (microscopic and macroscopic revision). Table 16 below contains the quantitative data for frequency and total time spent on revision for both EG and CG.

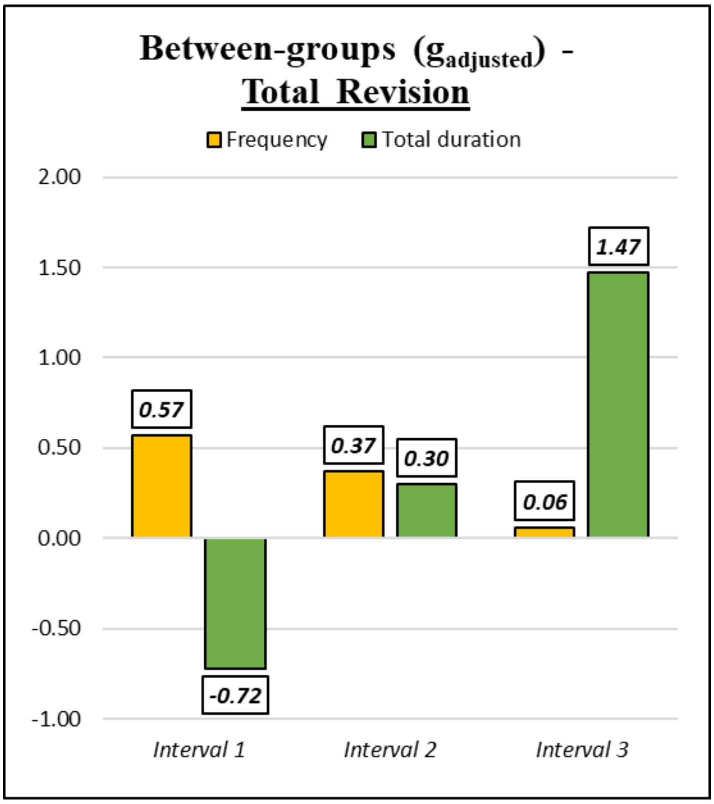
Table 16.
Descriptive statistics and effect size for total revision by group and writing.

			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Total revision	Frequency	Interval 1	18.10	14.47	10.40	6.34	**0.66**	7.37	3.62	2.62	2.19	**1.50**
		Interval 2	13.90	8.79	12.50	11.72	**0.13**	9.12	6.53	3.62	2.13	**1.07**
		Interval 3	26.60	36.48	24.20	43.66	**0.03**	9.87	6.17	3.37	3.96	**1.19**
	Total duration	Interval 1	.021	.01	.023	.04	**–0.06**	.014	.005	.024	.06	**–0.20**
		Interval 2	.057	.13	.019	.05	**0.36**	.031	.04	.022	.10	**0.12**
		Interval 3	.014	.01	.011	.07	**0.06**	.035	.02	.014	.06	**0.39**

As can be appreciated from Table 16 above, in the first interval, both EG and CG decreased the frequency of revision episodes (EG: Time 1= 18.10, Time 2= 10.40; CG: Time 1= 7.37, Time 2= 2.62), with a nearly medium (EG: g= 0.66) and large effect size (CG: g= 1.50). In the second interval, we found that the CG decreased the frequency of revisions (Time 1= 9.12, Time 2= 3.62) with a large effect (g= 1.07). In the case of the EG, the effect size was trivial (g= 0.13). In the third interval, the CG decreased the frequency of revisions (Time 1= 9.87, Time 2= 3.37) with a large effect size (g= 1.19). The EG, however, maintained a similar frequency of revision (Time 1= 26.60, Time 2= 24.20), with a trivial effect size (g= 0.03).

Concerning the total duration of Total revision, we found that no effects were observed, given the trivial effect sizes, (EG: g= –0.06, CG: g= –0.20) in the first interval for either the EG (Time 1= .021, Time 2= .023) or the CG (Time 1= .014, Time 2= .024). In the second and third intervals, the effect sizes were very small. It is, however, worth noting the great variability intra-subjects existing across intervals in the total duration of total revision for both groups.

Figure 41.
Between-group comparison with the effect size for the frequency and total duration of Total revision.



Moving on to the between-groups comparison, Figure 41 above shows the effect sizes for the frequency and total duration of the process of revision. As for the frequency of revision, we found that in the first interval the EG resorted more frequently to revision than the CG (EG= 10.40, CG= 2.62), with a small effect size ($g_{adjusted}$ = 0.57). In the second and third intervals, no effects were observed according to the effect sizes, which were trivial, (Interval 2: $g_{adjusted}$ = 0.37, Interval 3: $g_{adjusted}$ = 0.06), although the descriptive statistics display that the EG resorted more frequently to revisions than the CG.

Regarding the total duration of the revision process, we found that the EG and CG maintained similar values in Time 2 (EG= .023, CG= .024), with a medium effect size ($g_{adjusted}$ = –0.72). In the second interval, the effect size was trivial ($g_{adjusted}$ = 0.30). In the third interval, we found that the CG devoted more time to revisions than the EG (EG= .011, CG= .014), with a large effect size ($g_{adjusted}$ = 1.47).

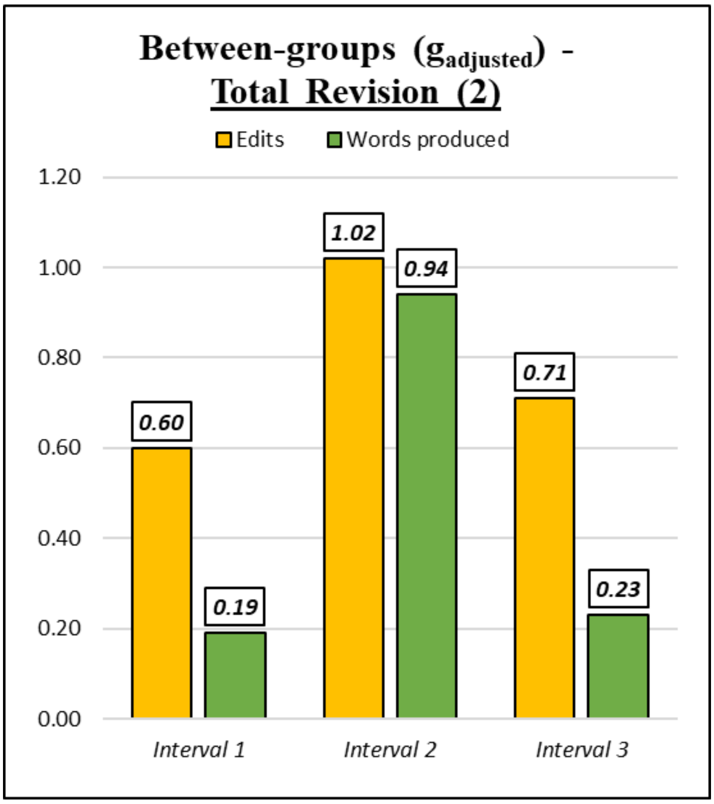
Table 17.
Descriptive statistics and effect size for edits and words produced for total revision by group and writing.

			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Total revision	Edits	Interval 1	58.39	67.10	71.00	68.93	**–0.18**	36.62	13.19	14.75	19.68	**1.23**
		Interval 2	70.50	57.16	71.40	51.09	**–0.02**	53.87	38.74	16.50	12.44	**1.23**
		Interval 3	125.90	168.40	160.70	162.31	**–0.20**	95.75	24.55	37.25	51.49	**1.04**
	Words produced	Interval 1	19.60	19.13	11.10	8.58	**0.55**	6.75	2.86	3.62	3.70	**0.89**
		Interval 2	26.30	39.52	17.30	16.77	**0.28**	22.25	30.67	3.37	2.44	**0.82**
		Interval 3	67.70	168.40	89.70	144.16	**0.13**	18.50	17.54	20.87	31.27	**0.09**

Our first research question enquired about the temporal distribution of writing processes in Time 1 and Time 2 after the provision or absence of model texts. Revision is a process which, as formulation, is characterized by modifications in the text, and also, text production. Table 17 above displays the descriptive statistics (i.e. means and SDs) and effect sizes (Hedges' *g*) for non-temporal measures: edits and words produced. In the first interval, we found that the CG performed fewer edits (Time 1= 36.62, Time 2= 14.75) than the EG (Time 1= 58.39, Time 2= 71.00), with a large effect size (g= 1.23). In the second interval, the EG did not vary the number of edits (Time 1= 70.50, Time 2= 71.40), but the CG decreased the number (Time 1= 53.87, Time 2= 16.50), with a large effect size (g= 1.23). In the third interval, we found that the CG maintained this tendency and decreased the number of edits (Time 1= 95.75, Time 2= 37.25), with great variability in Time 2 (SD= 51.49), and a large effect size (g= 1.04).

In terms of words produced during revision, there was a decrease in the EG (Time 1= 19.60, Time 2= 11.10) but it was trivial (g= 0.55). The CG decreased the number of words (Time 1= 6.75, Time 2= 3.62), with a medium effect size (g= 0.89). In the second interval, the effect size was trivial for the EG (g= 0.28), but the CG decreased the number of words produced (Time 1= 22.25, Time 2= 3.37) with a medium effect size (g= 0.82). In the third interval, the effect sizes were trivial for the EG (g= 0.13) and the CG (g= 0.09).

Figure 42.
Between-group comparison with the effect size of edits and words produced of Total revision.



As regards the between-groups comparison between the EG and CG (see Figure 42 above), the descriptives point to visible dissimilarities, further backed by the adjusted effect size. In terms of edits, we found that, in the first interval, edits were higher in the EG than in the CG (EG= 71.00, CG= 14.75), with a small effect size ($g_{adjusted}$ = 0.60). In the second interval, we found that values for edits in the EG (Time 2= 71.40) were higher than in the CG (Time 2= 16.50), with a large effect size ($g_{adjusted}$ = 1.02). In the third interval, the EG (Time 2= 160.70) kept higher edits than the CG (Time 2= 37.25) with a medium effect size ($g_{adjusted}$ = 0.71).

Regarding the words produced during revision, there were no meaningful differences according to the effect sizes in the first interval ($g_{adjusted}$ = 0.19) and third interval ($g_{adjusted}$ = 0.23), which were trivial. The descriptive statistics show, however, that the EG produced or modified more words than the CG in these intervals. In the second interval, the effect size was large (g = 0.94), since the EG produced more words than the CG (EG= 17.30, CG= 3.37).

5.1.3.1. Microscopic revision

We next report on the frequency and total duration of microscopic revisions. As a reminder for the reader, microscopic revisions are those modifications in the text which are below the phrase level (Bridwell, 1980; Kim, 2020). At this point, edits might be confounded with microscopic revision, but the truth is that, in essence, they are different as they are not modifications to the text. As detailed in Chapter 4, edits include the number of actions needed to complete the action, including each keystroke, shift key, cursor movements, mouse clicks, among other actions.

Table 18 below contains the quantitative data for the frequency and total time spent on microscopic revisions for both EG and CG.

Table 18.
Descriptive statistics and effect size for microscopic revisions by group and writing version.

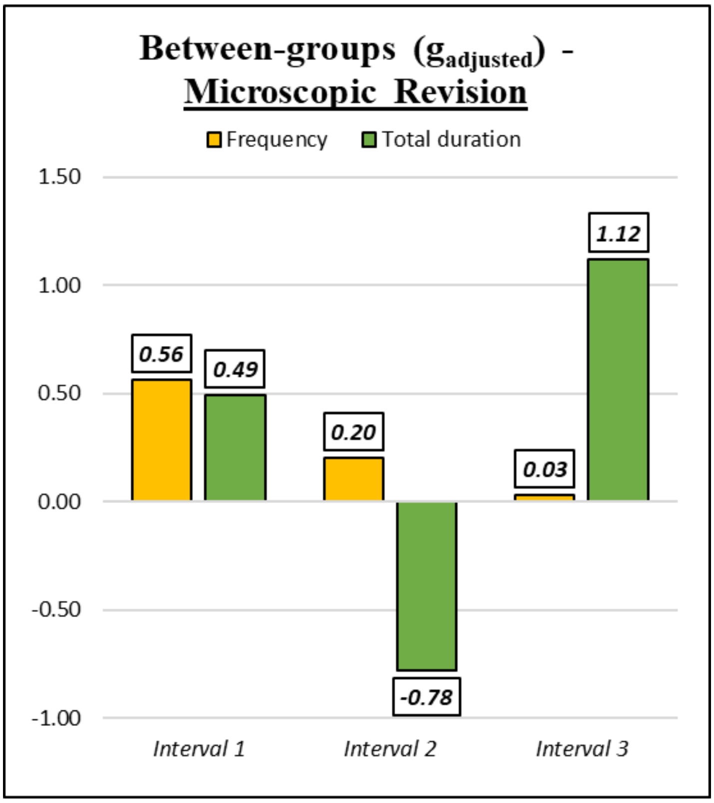
			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
	Measure		M	SD	M	SD	*g*	M	SD	M	SD	*g*
Microscopic Revision	Frequency	Interval 1	15.50	14.47	8.80	6.08	**0.58**	7.00	3.62	2.37	1.99	**1.50**
		Interval 2	12.60	8.99	10.40	11.67	**0.20**	8.25	5.94	3.62	2.13	**0.98**
		Interval 3	22.50	34.43	18.40	38.39	**0.11**	8.75	5.82	2.37	2.50	**1.35**
	Total duration	Interval 1	.008	.008	.011	.00	**–0.32**	.012	.007	.011	.011	**0.15**
		Interval 2	.055	.130	.007	.004	**0.49**	.020	.010	.009	.006	**0.69**
		Interval 3	.011	.010	.015	.016	**–0.29**	.023	.019	.009	.012	**0.77**

In the first interval, we found that the CG decreased the frequency of microscopic revisions (Time 1= 7.00, Time 2= 2.37) with a large effect size (g= 1.50). In the second interval, the CG resorted less frequently to microscopic revisions as well (Time 1= 8.25, Time 2= 3.62) with a medium effect size (g= 0.98). In the third interval, the EG decreased the frequency of microscopic revisions (Time 1= 8.75, Time 2= 2.37) and a large effect size (g= 1.35). Across the three intervals, the EG slightly decreased the frequency of microscopic revisions, but the effect sizes were trivial (Interval 1: g= 0.58, Interval 2: g= 0.20, Interval 3: g= 0.11).

Regarding the total duration of microscopic revisions, no meaningful effects were observed in the EG from Time 1 to Time 2 across intervals as the effect sizes were trivial (Interval

1: g= –0.32, Interval 2: g= 0.49, Interval 3: g= –0.29). We found that the CG devoted less time to microscopic revisions in the second interval (Time 1= .020, Time 2= .009) and third interval (Time 1= .023, Time 2=.009), with small-to-medium effect sizes (Interval 2: g= 0.69, Interval 3: g= 0.77).

Figure 43.
Between-group comparison with the effect size for frequency and total duration of microscopic revision.



Moving on to the between-groups comparison (see Figure 43), in the first interval, we found that the EG revised more frequently than the CG after receiving WCF, with a small effect size ($g_{adjusted}$ = 0.56). In terms of the second and third intervals, the differences between EG and CG are trivial (Interval 2: $g_{adjusted}$ = 0.20, Interval 3: $g_{adjusted}$ = 0.03). However, the descriptive statistics show that the frequency of microscopic revisions was higher in the EG than in the CG across these two intervals.

As for the total duration of microscopic revisions, the effect size was small in the first interval ($g_{adjusted}$ = 0.49), since both the EG and CG devoted similar amounts of time (EG= .011, CG= .011) although with different variability (EG: SD= .00, CG: SD= .011). In the second interval, we found that the EG decreased the time spent on microscopic revisions (= .007) while the value was higher for the CG (= .009), with a medium effect size ($g_{adjusted}$ = –0.78). In the third interval, the EG spent more time on microscopic revisions than the CG (EG= .015, CG= .009), as observed in the large effect size ($g_{adjusted}$ = 1.12).

5.1.5.2. Macroscopic revision

We next report on the frequency and total duration of macroscopic revisions. As a reminder for the reader, macroscopic revisions are those modifications in the text which are above the phrase level (Bridwell, 1980; Kim, 2020).

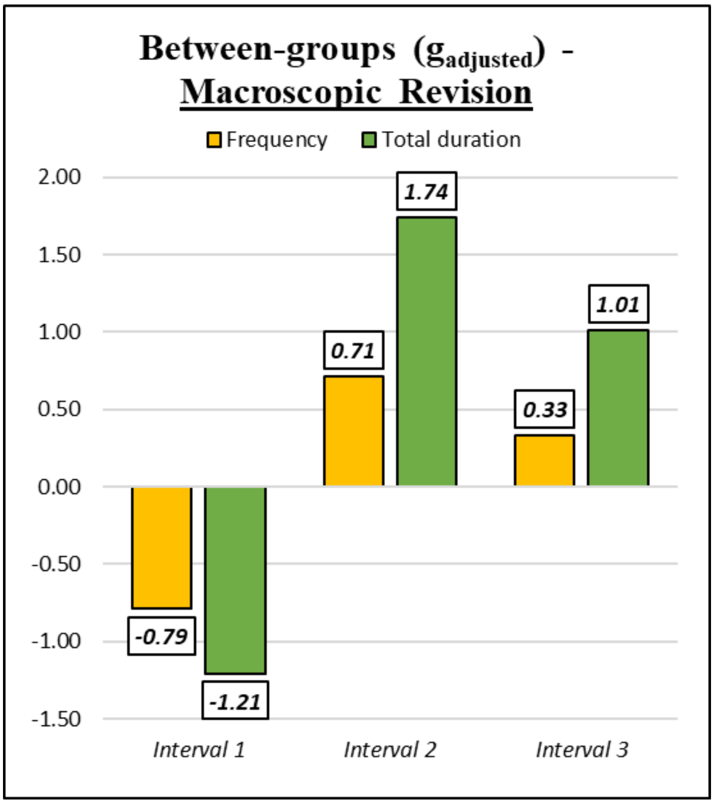
Table 19.
Descriptive statistics and effect size for macroscopic revisions by group and writing.

			EG					CG				
			Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
			M	SD	M	SD	*g*	M	SD	M	SD	*g*
Macroscopic Revision	Frequency	Interval 1	2.60	1.83	1.70	2.54	**0.39**	.37	.51	.25	.46	**–0.23**
		Interval 2	1.30	1.56	2.10	2.64	**0.35**	.87	1.12	.00	.00	**–1.04**
		Interval 3	3.10	4.88	5.80	7.16	**0.42**	1.12	.99	1.00	1.51	**–0.09**
	Total duration	Interval 1	.011	.012	.007	.009	**–0.34**	.002	.002	.018	.04	**0.45**
		Interval 2	.002	.004	.008	.008	**0.87**	.012	.025	.00	.00	**–0.62**
		Interval 3	.004	.005	.020	.027	**0.76**	.013	.018	.011	.02	**–0.10**

Table 19 above displays the descriptive statistics of frequency and total duration for macroscopic revision along with the effect size. In the first interval, the EG decreased it (Time 1= 2.60, Time 2= 1.70) as well as the CG did (Time 1= .37, Time 2= .25), but with trivial effect sizes (EG: g= 0.39, CG: g= –0.23). Similarly, there is variability among participants given the high SDs in Time 1 (EG: SD= 1.83, CG: SD= .51) and Time 2 (EG: SD= .51, CG: SD= .46). In the second interval, we found that the CG did not resort to macroscopic revisions (Time 1= .87, Time 2= .00), with a very large effect size (g= –1.04). The EG increased the number of macroscopic revisions (Time 1= 1.30, Time 2= 2.10) with a trivial effect size (g= 0.35). In the third interval, the effect sizes were trivial for both the EG (g= 0.42) and the CG (g= –0.09).

Regarding the total duration of macroscopic revisions, the results of the first interval showed trivial effect sizes for both groups (EG: g= –0.34, CG: g= 0.45), although the EG decreased the time spent on macroscopic revisions (Time 1= .011, Time 2= .007) and the CG increased it (Time 1= .002, Time 2= .018). In the second interval, we found that the EG devoted more time to macroscopic revisions (Time 1= .002, Time 2= .008) with a medium effect size (g= 0.87). The CG reduced the time spent on microscopic revisions (Time 1= .012, Time 2= .00) with a small effect size (g= –0.62). In the third interval, the EG increased the time spent on this type of revisions (Time 1= .004, Time 2= .020) with a small-to-medium (g= 0.76). The CG slightly decreased the time devoted to macroscopic revisions (Time 1= .013, Time 2= .011) with a trivial effect size (g= –0.10).

Figure 44.
Between-group comparison with the effect size for frequency and total duration of macroscopic revision.



The between-groups comparison indicated that, in the first interval, the EG resorted more frequently to macroscopic revisions than the CG (EG= 1.70, CG= .25) with a medium effect size ($g_{adjusted}$= –0.79). In the second interval, we found that the frequency of macroscopic revisions was higher in the EG than in the CG with a medium effect size ($g_{adjusted}$= 0.71). In the third interval, the EG resorted more to macroscopic revisions than the CG, although the effect size was trivial ($g_{adjusted}$= 0.33).

In terms of total duration of macroscopic revisions, the between-groups comparison were relevant across the three intervals. In the first interval, we found that the EG spent less time than the CG (EG= .007, CG= .018), and the effect size was large ($g_{adjusted}$= –1.21). In the second and third intervals, the EG devoted more time to macroscopic revisions than the CG with very large effect sizes (Interval 2: $g_{adjusted}$= 1.74, Interval 3: $g_{adjusted}$= 1.01).

► ***Summary of results***

Our first research question asked about the effect of models as WCF on the planning, formulation, and revision processes in young learners' writing. Our results show that planning was manifestly different for the group receiving WCF, which planned more frequently than the CG. In terms of total duration, the EG increased the time spent on planning more than the CG, which only increased it toward the end of the composition process. Model texts also affected the type of planning insomuch as the EG decreased the number of global planning episodes, especially at the beginning. However, the CG decreased it across all intervals at a much larger extent than the EG. Additionally, our results show that the EG planned more frequently at a local level (i.e. local planning) than the CG.

The formulation process was more frequently resorted to at the beginning of writing in the EG, with more time spent on task than the CG. Also, the EG decreased the amount of time devoted to formulation toward the end of the process, with the CG spending more time on text production. The EG produced more edits (i.e. the number of actions needed to complete text production) throughout the writing process, but this difference was relevant at the beginning of the composition. Both groups reduced the number of words produced after the provision (or absence of) WCF, but the EG maintained a higher number than the CG.

In terms of revision, our results show that the EG revised more frequently than the CG, especially at the beginning of the writing process. Both the EG and CG reduced the amount of time devoted to revising, especially at the beginning and end of the composition process. Regarding the type of revision, the EG produced fewer microscopic revisions after WCF, but these were higher than in the CG. Likewise, the EG devoted more time to this type of revision especially at the beginning and end of the composition process. In terms of macroscopic revisions, the EG carried out more macroscopic revisions than the CG across the three intervals. This was also reflected in the total duration, where the EG devoted more time to macroscopic revisions in the middle and end of the composition process than the CG.

5.2. Effects of models on pausing behavior throughout composing

►*RQ 2. To what extent are there differences in young EFL learners' pausological behaviour when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?*

Table 20 (see next page) presents a summary of what the reader will find in the section devoted to RQ2. As can be seen, there are two dimensions: (i) global pausing and process measures, and (ii) pause location. In the first dimension, measures related to the pausological behavior (total pausing time, or number of pauses across intervals as well as the mean duration) or fluency-related aspects (e.g. interkeystroke interval, or pauses per minute) were examined. In the second dimension, the measures regarding the location of the pause in terms of number, mean duration and the location ratio were examined.

Table 20.
Summary of analyses performed on the different measures along with the section where they are described.

RQ2: To what extent are there differences in young EFL learners' pausological behaviour when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?		
Dimension	Analyses conducted	Section in the text
1: Effect of WCF on global pausing and process measures.	Analysis of process time and pausing measures during writing and during revising. Analysis of total frequency of pauses and fluency-related measures (interkeystroke interval and pauses per minute).	▪ Writing process: pausing and process measures.
	Analysis of the distribution of pauses across intervals in terms of frequency and duration during writing and during revising.	▪ Pause distribution and duration.
2. Effects of WCF on pause location.	Analysis of pauses per location (i.e. within-word, before-word...) in terms of number, mean duration and location ratio.	▪ Pause location.

5.2.1. Writing process: pausing and process measures

In this section of results, data on pausing behavior in the global process of writing will be presented. Thus, important dependent variables such as the time spent on writing actively or pausing, the number of pauses as well as the distribution across intervals will be examined and compared within- and between subjects. In this section, we next report on the results regarding the following set of data: (1) process time measures, which will give us an overview of the amount of time devoted to the whole writing process, the time spent on writing actively, and the amount of pausing time; and (2) total number of pauses throughout the writing process, the interkeystroke interval, and finally, pauses per minute. These measures will be presented from a within-groups (i.e. the effect from Writing 1 to Writing 2) and between-groups perspective (i.e. to observe the effects of WCF between the EG and CG)

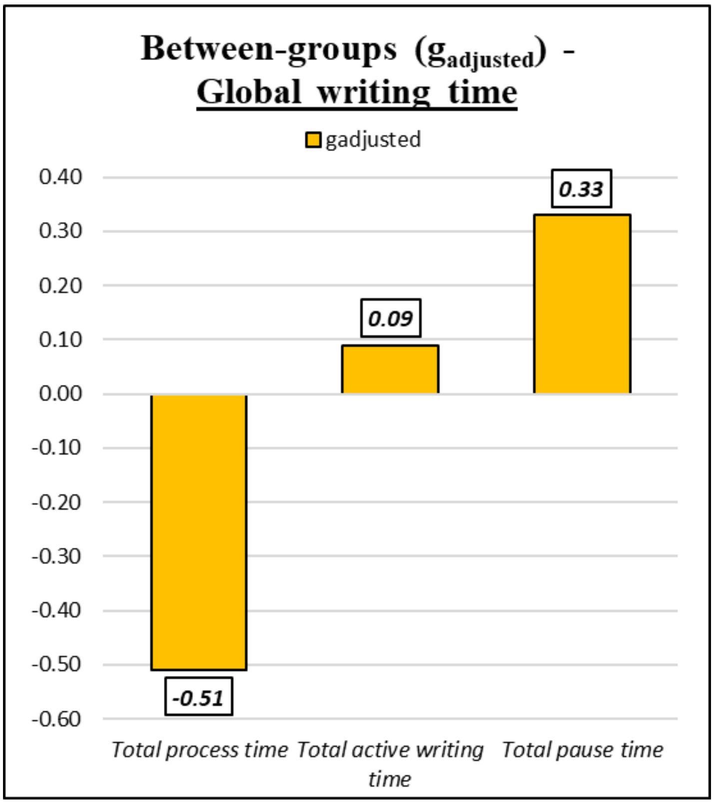
Table 21.
Descriptive statistics for pausing and time measures in the writing process by group and writing version.

	EG					CG				
	Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
	M	SD	M	SD	*g*	M	SD	M	SD	*g*
Total process time	1428	296.26	1042	417.38	**1.02**	852.80	56.88	377.90	155.66	**3.83**
Total active writing time	596.30	307.47	543.30	370.08	**0.15**	312.20	109.74	181.70	94.55	**1.20**
Total pause time	832.10	293.05	498.30	237.73	**1.20**	540.50	115.42	196.60	90.51	**3.13**

Table 21 above shows the descriptive statistics and effect size (Hedges' *g*) for the dependent variables of total process time, total active writing time, and total pause time. There was a decrease between time 1 and time 2, with wide variability across participants and variables in both groups. In the case of the EG, there is a decrease in the total writing process time with a large effect size (g= 1.02), which was even more marked in the case of the CG (g= 3.83).

Following with the time actively spent on writing, the EG slightly devoted less time after WCF (Time 1= 596.30, Time 2= 543.30) although the effect size was trivial (g= 0.15). The CG, however, devoted less time (Time 1= 312.20, Time 2= 181.70) with a large effect size (g= 1.20). In the case of total pause time, the EG decreased by half the time devoted to pausing (Time 1= 832.10, Time 2= 498.30) with a large effect size (g= 1.20). Likewise, the CG decreased the time spent on pausing (Time 1= 540.50, Time 2= 196.60) with a very large effect size (g= 3.13).

Figure 45.
Between-group comparison with the effect size of global writing time measures.



To understand better the potential effect of WCF provision (see Figure 45 above), the between-groups comparison of Writing 2 revealed that the EG devoted more time to the whole writing process than the CG with a small effect size ($g_{adjusted}$= -0.51). Similarly, the EG devoted more time than the CG to both active writing (EG= 543.30, CG= 181.70) and total pause time (EG= 498.30, CG= 196.60), although the effect sizes were trivial ($g_{adjusted}$= 0.09 and $g_{adjusted}$= 0.33, respectively).

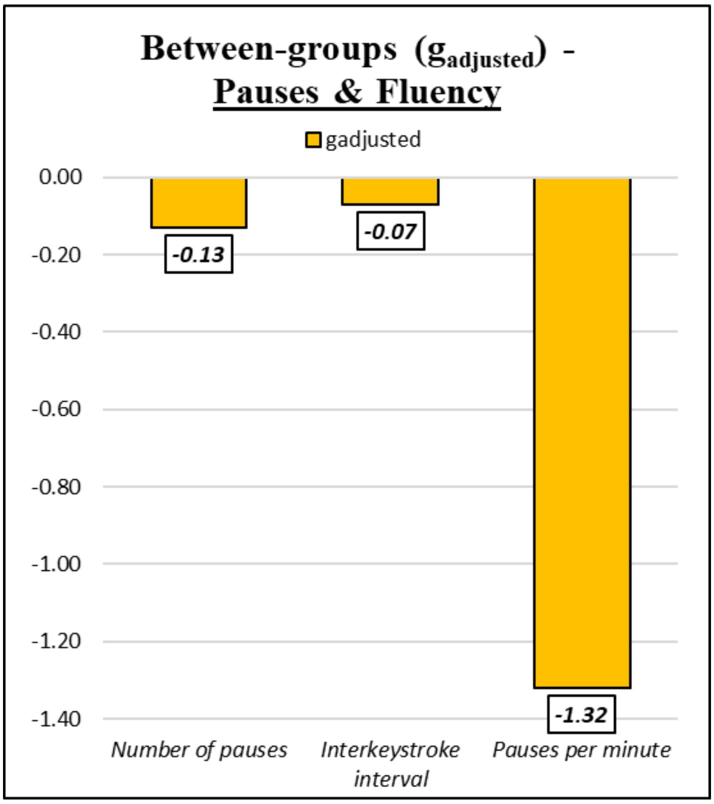
Table 22 below displays the descriptive statistics along with the effect sizes for the number of pauses across the composition process, the interkeystroke interval, and the number of pauses per minute.

Table 22.
Descriptive statistics for the number of pauses, transcription fluency (milliseconds) and fluency by group and writing version.

	EG					CG				
	Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
	M	SD	M	SD	*g*	M	SD	M	SD	*g*
Number of pauses	105.50	35.51	72.71	31.43	**0.94**	63.25	13.36	36.75	17.49	**1.61**
Interkeystroke Interval	210.67	40.69	202.20	46.18	**0.19**	192.27	39.15	186.65	30.71	**0.15**
Pauses per minute	4.40	1.10	4.29	.80	**0.11**	4.55	.93	6.15	1.60	**1.16**

Beginning with the number of pauses performed by each group, there is a tendency to decrease in both the EG and CG in their Writing 2. In terms of number of pauses, both the EG (Time 1= 105.50, Time 2= 72.71) and the CG (Time 1= 63.25, Time 2= 36.75) decreased the number of pauses with large effect sizes (EG: g= 0.94, CG: g= 1.61). Regarding interkeystroke interval, no effect was found within-groups in neither the EG (g= 0.19) nor the CG (g= 0.15), with trivial effect sizes. Finally, we found that the EG did not modify the pauses per minute after WCF (Time 1= 4.40, Time 2: 4.29), but the CG increased it (Time 1= 4.55, Time 2= 6.15) with a medium-to-large effect (g= 1.16).

Figure 46.
Between-group comparison with the effect size of pauses and fluency measures.



As for the between-groups comparison (see Figure 46 above), there were no differences between the EG and the CG in either the number of pauses ($g_{adjusted}$ = –0.13) or interkeystroke interval ($g_{adjusted}$ = –0.07), given the trivial effect sizes. However, there was a difference between the EG (= 4.29) and the CG (= 6.15) according to the large effect size ($g_{adjusted}$ = –1.32).

5.2.2. Pause distribution and duration

These analyses provide useful information on how pauses were distributed and the mean duration of each pause across the different intervals. This constitutes valuable information on the pausological behavior of both the EG and CG, and the effect of WCF. Table 23 displays the descriptive statistics and effect sizes for the frequency of pauses and mean duration across intervals of the composition process within and across groups.

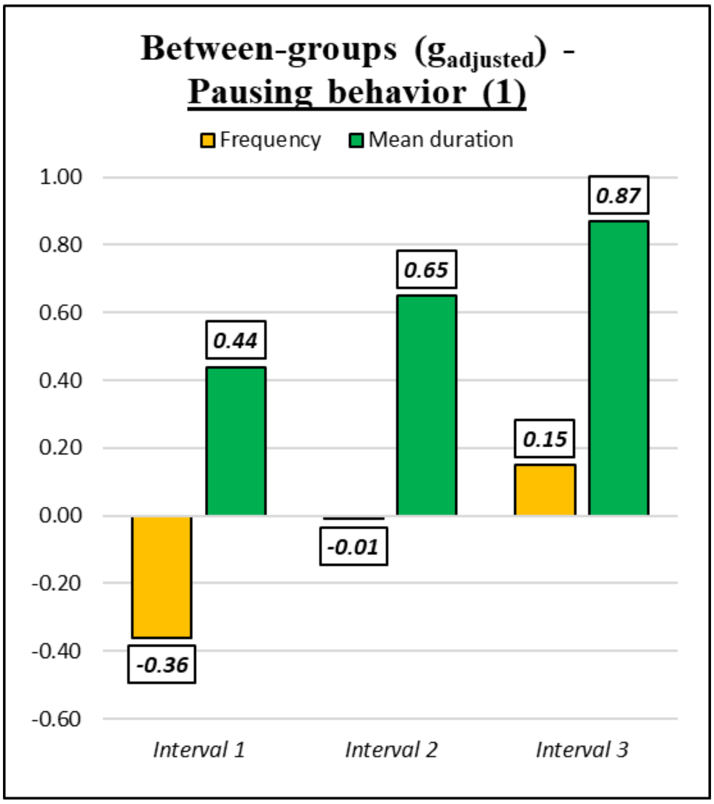
Table 23.
Descriptive statistics for pause frequency and mean duration (in seconds) per interval by group and writing version.

		EG					CG				
		Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
		M	SD	M	SD	*g*	M	SD	M	SD	*g*
Pause frequency	Interval 1	39.40	10.04	27.30	11.06	**1.10**	25.13	3.76	13.88	6.73	**1.95**
	Interval 2	37.00	15.37	24.70	10.33	**0.90**	19.75	6.09	13.00	4.66	**1.18**
	Interval 3	29.10	14.39	20.70	11.67	**0.61**	18.38	5.50	9.87	6.71	**1.31**
Mean duration	Interval 1	7.82	3.30	6.53	2.24	**0.44**	7.88	2.11	5.71	1.25	**1.19**
	Interval 2	13.39	14.39	8.42	5.63	**0.44**	12.73	13.55	5.32	.88	**0.73**
	Interval 3	7.83	4.46	9.77	9.37	**–0.25**	8.81	3.34	5.03	1.43	**1.39**

Globally considered, in the first interval, both groups decreased the number of pauses with a practically significant results (EG: g= 1.10, CG: g= 1.95). The number of pauses decreased as well in the EG, with small-to-medium effect sizes (Interval 2: g= 0.90, Interval 3: g= 0.61) and the CG, with larger effect sizes (Interval 2: g= 1.18, Interval 3: g= 1.31). As observed, the decrease was marked in the case of the CG (Interval 2: Time 1= 19.75, Time 2= 13.00, Interval 3: Time 1= 18.38, Time 2= 9.87).

Regarding the mean duration of pauses, no meaningful differences were observed for the EG across any interval since all the effect sizes yielded a trivial effect (Interval 1: g= 0.44, Interval 2: g= 0.44, Interval 3: g= –0.25). It is worth mentioning the great variability existing in the second interval in both the EG (Time 1: SD= 14.39) and CG (Time 1: SD= 13.55), whose values are larger than the means themselves. This indicate great dispersion of the data. Likewise, the variability in the third interval in Time 2 is noticeable in the EG (Time 2: SD= 9.37). Conversely, we found that the CG decreased the mean duration of pauses across intervals with medium and large effect sizes across all intervals (Interval 1: g= 1.19, Interval 2: g= 0.73, Interval 3: g= 1.39).

Figure 47.
Between-group comparison with the effect size for pause frequency and duration per interval..



Moving on to the between-groups comparison of these pausing measures (see Figure 47), the effect sizes were trivial for the frequency of pauses across intervals (Interval 1: $g_{adjusted}$= –0.36, Interval 2: $g_{adjusted}$ = –0.01, Interval 3: $g_{adjusted}$ = 0.15). Nonetheless, despite the non-significance of the between-groups comparison, the descriptives reveal that the EG paused more frequently across intervals than the CG. Regarding the mean duration of pauses, the effect size was small in the first interval ($g_{adjusted}$ = 0.44), in which the EG paused slightly longer than the CG (EG, Time 2= 8.42; CG, Time 2= 5.71). In the second interval, we found that the EG pause longer than the CG (EG, Time 2= 8.42; CG, Time 2= 5.32) with a nearly medium effect size ($g_{adjusted}$ = 0.65). In the third interval, this tendency is maintained (EG, Time 2= 9.77; CG, Time 2= 5.03) with a medium-to-large effect size ($g_{adjusted}$ = 0.87).

5.2.3. Pause location

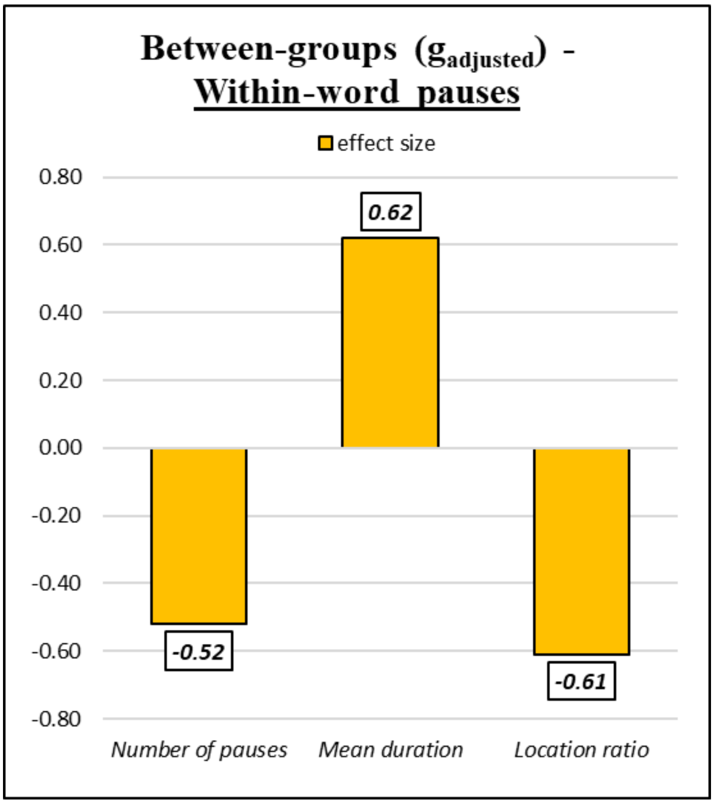
Table 24 below presents the descriptive statistics and effect size for the dependent variables of pause location and the associated measures (i.e. number, mean duration, and location ratio). These linguistic location of pauses are measures of pause patterns part of what constitutes pausological behavior (Barkaoui, 2016), hence the relevance for responding to our RQ2. In what follows, each dependent variable will be described from a within-groups perspective to discern the potential effect of the WCF, and a between-groups perspective will allow for a comparison of the posttest (i.e. Writing 2) itself.

Table 24.
Descriptive statistics and effect size for pause location by group and writing version (Writing 1 vs Writing 2).

		EG					CG				
		Writing 1		Writing 2		ES	Writing 1		Writing 2		ES
		M	SD	M	SD	*g*	M	SD	M	SD	*g*
Within word	Number	15.80	12.18	9.00	6.51	**0.67**	9.12	8.89	8.37	11.21	**0.07**
	Mean duration	5.06	2.31	4.38	2.26	**0.29**	5.25	2.84	3.22	1.62	**0.83**
	Location ratio	.17	.09	.13	.06	**0.44**	.15	.12	.19	.16	**–0.25**
Before words	Number	41.30	17.22	27.20	10.36	**0.95**	22.63	6.76	13.00	4.69	**1.56**
	Mean duration	8.11	4.00	7.70	4.66	**0.09**	11.62	9.50	5.74	1.58	**0.82**
	Location ratio	.46	.13	.46	.11	**0.06**	.44	.10	.42	.11	**0.12**
Before sentences	Number	1.90	2.88	1.10	2.08	**0.30**	.62	.52	.62	.74	**0.00**
	Mean duration	.86	1.96	1.11	2.42	**–0.11**	.00	.00	1.91	5.42	**–0.47**
	Location ratio	.02	.02	.01	.03	**0.23**	.01	.01	.02	.03	**–0.67**
Before paragraphs	Number	1.20	3.16	1.80	3.08	**–0.18**	.12	.35	.00	.00	**0.47**
	Mean duration	.91	2.15	4.09	6.96	**–0.59**	.00	.00	.00	.00	–
	Location ratio	.01	.03	.03	.05	**–0.48**	.001	.00	.00	.00	**0.49**
After words	Number	5.70	4.19	6.10	4.53	**–0.09**	5.50	3.38	4.00	3.51	**0.41**
	Mean duration	10.68	17.34	4.49	2.44	**0.48**	8.35	2.93	3.16	2.97	**1.66**
	Location ratio	.06	.05	.10	.06	**–0.57**	.106	.06	.101	.06	**0.08**
After sentences	Number	.30	.48	.50	.97	**–0.25**	.50	.76	.25	.46	**0.38**
	Mean duration	.00	.00	.65	2.07	**–0.43**	.28	.80	.00	.00	**0.47**
	Location ratio	.005	.01	.006	.01	**–0.09**	.01	.02	.008	.01	**0.16**
After paragraphs	Number	1.20	2.10	.60	1.26	**0.33**	.12	.35	.25	.71	**–0.21**
	Mean duration	1.48	3.28	.41	1.32	**0.41**	.00	.00	.28	.81	**–0.47**
	Location ratio	.010	.02	.009	.02	**0.17**	.003	.01	.008	.02	**–0.30**
Other	Number	23.20	11.76	15.60	9.25	**0.69**	14.00	6.72	6.12	1.25	**1.54**
	Mean duration	6.62	2.13	6.02	2.84	**0.23**	6.20	5.55	4.18	.94	**0.48**
	Location ratio	.243	.06	.246	.06	**–0.05**	.27	.09	.23	.15	**0.25**

Pauses at within-word boundaries experienced a decrease in the EG (Time 1= 15.80, Time 2= 9.00) with a small effect size (g= 0.67). We also found out that the CG decreased the mean duration of within-word pauses (Time 1= 5.25, Time 2= 3.22) with a medium effect size (g= 0.83).

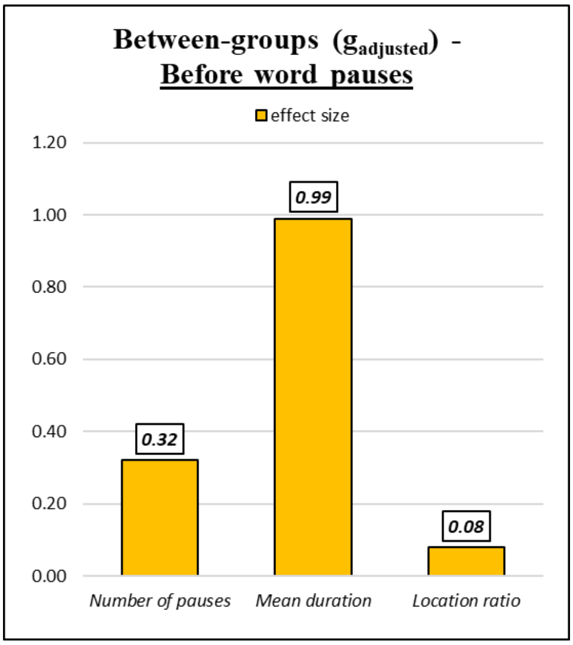
Figure 48.
Between-group comparison with the effect size for within-word pauses.



Turning our attention to the between-groups effect sizes (see Figure 48 above), the comparison of the number of within-word pauses was trivial as evidenced by the effect size ($g_{adjusted}$= -0.52). In the case of mean duration, the EG devoted more time to each of these within-word pauses with a small-to-medium effect size ($g_{adjusted}$= 0.62). The location ratio for within-word pauses was much lower in the EG than in the CG (EG= .13, CG= .19), a result that goes in line with the effect size ($g_{adjusted}$= -0.61), which was small-to-medium.

Moving on to before-word pauses, the EG decreased the frequency of these pauses (Time 1= 41.30, Time 2= 27.20) with a medium effect size (g= 0.95). Similarly, the CG decreased the number of before-word pauses (Time 1= 22.63, Time 2= 13.00) with a large effect size (g= 1.56). We also found that the CG decreased the mean duration of before-word pauses (Time 1= 11.62, Time 2= 5.74) with a medium effect size (g= 0.82).

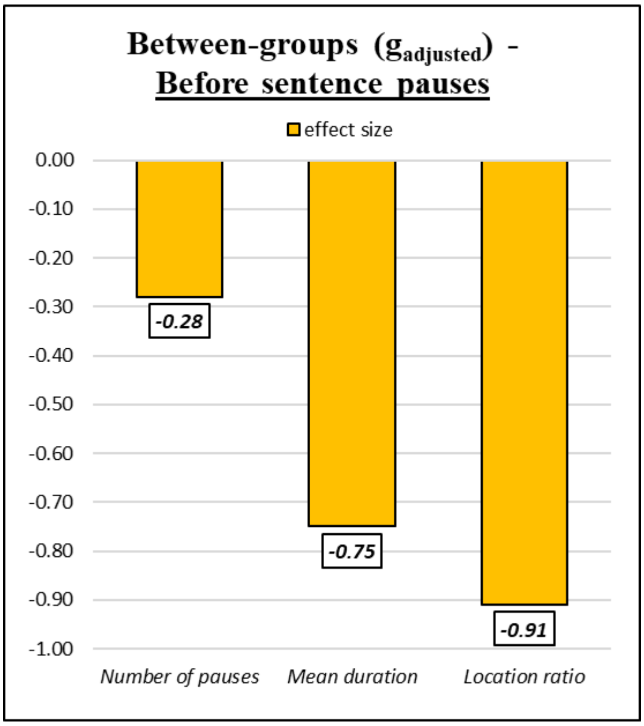
Figure 49.
Between-group comparison with the effect size for before-word pauses.



As regards the between-groups comparison in Writing 2 (see Figure 49 above), the effect sizes were trivial for the number of pauses ($g_{adjusted}$ = 0.32) or the location ratio ($g_{adjusted}$ = 0.08). In contrast, we found that the EG paused for longer before words than the CG (EG= 27.20, CG= 13.00) with a nearly large effect size ($g_{adjusted}$ = 0.99).

Moving on to before-sentence pauses, we found that the location ratio in the CG increased (Time 1= .01, Time 2= .02) with a small effect size (g= –0.67). It is worth mentioning that in both Time 1 and Time 2 there was wide variability among participants (SDs: Time 1= .01, Time 2= .03). As regards the EG, the effect sizes were trivial for all the measures (number: g= 0.30, mean duration: g= –0.11, location ratio: g= 0.23). Likewise, the changes of the CG observed in the number of before-sentence pauses (Time 1= .62, Time 2= .62) and the mean duration (Time 1= .00, Time 2= 1.91) were trivial (g= 0.00 and g= –0.47, respectively).

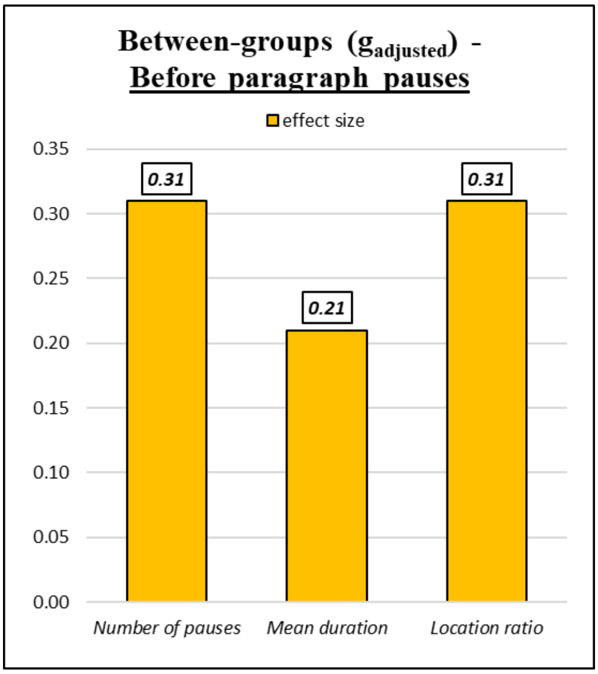
Figure 50.
Between-groups comparison with the effect size for before-sentence pauses.



As seen in Figure 50 above, the between-groups comparison in Writing 2 for the frequency of before-sentence pauses was trivial according to the effect size ($g_{adjusted}$= -0.28). Regarding mean duration, the effect size was medium ($g_{adjusted}$= -0.75), and in this case, the CG paused for longer before sentences than the EG (EG= 1.11, CG= 1.91). Finally, the effect size for the location ratio was medium-to-large ($g_{adjusted}$= -0.91) since there was a reduction for the EG (Time 1= .02, Time 2= .01) in contrast to the increase in the CG (Time 1= .01, Time 2= .02).

Pauses at before-paragraph boundaries are shown to be much more reduced, particularly in the CG. Writers in the WCF group slightly increased the frequency of pauses before paragraphs (Time 1= 1.20, Time 2= 1.80) with a trivial effect size (g= –0.18). In the case of the CG, there is a drastic reduction, but the effect size was also trivial (g= 0.47), possibly as a result of the wide SD in Writing 1 (Time 1= .35). Regarding mean duration of before-paragraph pauses, the EG increased it moderately after receiving WCF, but with a small effect size (g= –0.59). As the CG did not devote any time to before-paragraph pauses, the effect size was null (g= 0.00). With regards to the location ratio, writers in the EG resorted more to pauses at before-paragraph boundaries (Time 1= .01, Time 2= .03), although the effect size was trivial (g= –0.48). Similarly, the CG decreased it, but also in a trivial manner (g= 0.49).

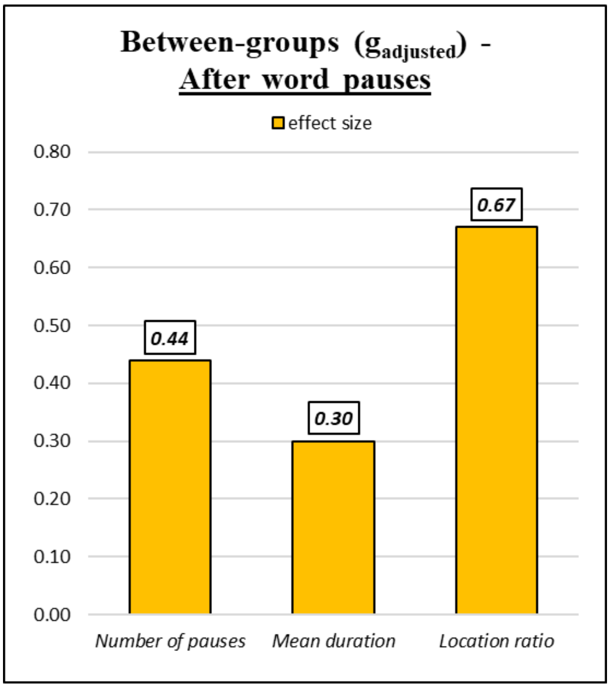
Figure 51.
Between-group comparison with the effect size for before-paragraph pauses.



The between-groups comparison did not yield any practically significant results (see Figure 51 above) since the effect sizes for the before-paragraph pauses were trivial: frequency ($g_{adjusted}$= 0.31), mean duration ($g_{adjusted}$= 0.21) and location ratio ($g_{adjusted}$= 0.31). Such a similarity across effect sizes lies in the values below zero in the CG.

Both groups decreased the number of pauses at after-word boundaries, although the effect sizes were trivial for both groups (EG: g= –0.09; CG: g= 0.41). Similarly, both the EG and the CG decreased the mean duration of pauses after words, although the effect size was trivial (g= 0.48) for the EG (Time 1= 10.68, Time 2= 4.49). The effect size, however, was large (g= 1.66) in the case of the CG (Time 1= 8.35, Time 2= 3.16). In the case of the location ratio, the EG increased it (Time 1= .06, Time 2= .10) with a trivial effect size (g= –0.57). Similarly, the CG maintained it (Time 1= .106, Time 2= .101), although the effect size was also trivial (g= 0.08).

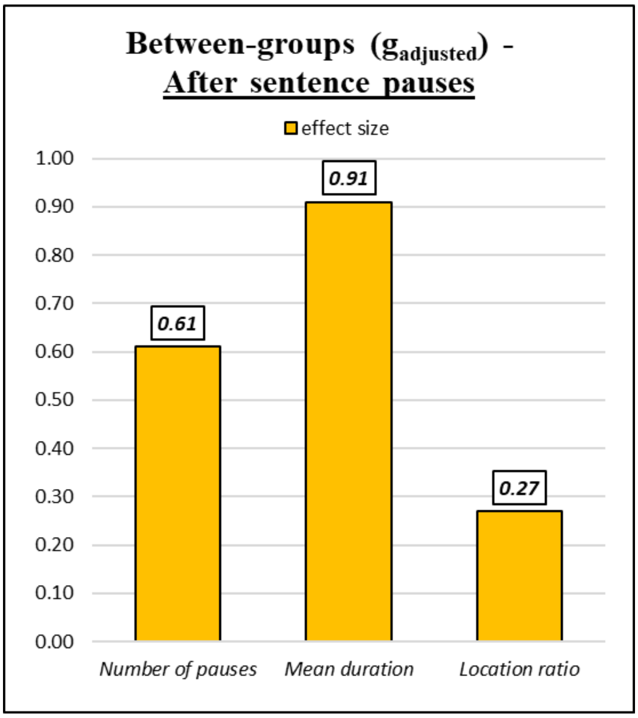
Figure 52.
Between-group comparison with the effect size for after-word pauses.



From a between-groups perspective (as can be seen in Figure 52), we observed that the EG paused more than the CG after the provision of WCF (EG= 6.10, CG= 4.00) with a small effect size ($g_{adjusted}$ = 0.44). In a similar way, the effect size was nearly medium for the location ration ($g_{adjusted}$ = 0.67), which the EG increased (Time 1= .06, Time 2= .10) and the CG maintained (Time 1= .106, Time 2= .101). Regarding mean duration, the effect size was trivial ($g_{adjusted}$ = 0.30).

Moving on to after-sentence pauses, the EG increased the number of these pauses (Time 1= .30, Time 2= .50) but the effect size was trivial (g= –0.25). The CG decreased the number of pauses at this boundary (Time 1= .50, Time 2= .25), but also with a trivial effect size (g= 0.38). In terms of mean duration, the EG increased the duration of after-sentence pauses (Time 1= .00, Time 2= .65) while the CG decreased it (Time 1= .28, Time 2= .00), although both effect sizes were trivial (g= –0.43, and g= 0.47, respectively). Regarding the location ratio of after-sentence pauses, the effect sizes were trivial either for the EG (g= –0.09) or the CG (g= 0.16).

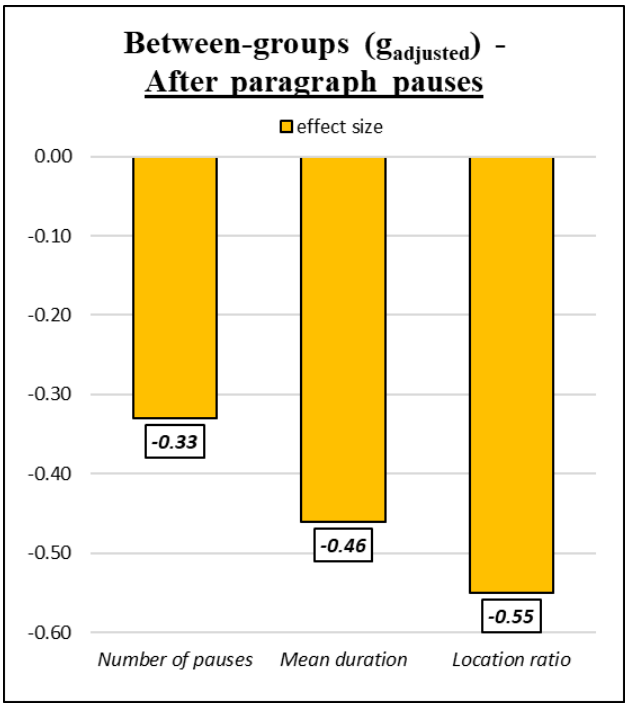
Figure 53.
Between-group comparison with the effect size for after-sentence pauses.



The between-groups comparison (see Figure 53 above) for the number of after-sentence pauses displayed a small-to-medium effect size ($g_{adjusted}$ = 0.61). The EG paused more frequently at this boundary than the CG (EG= .50, CG= .25). Also, we found that the EG paused for longer after sentences than the CG (EG= .65, CG= .00) with a nearly large effect size ($g_{adjusted}$ = 0.91). For the location ratio, the EG kept a lower value than the CG (EG= .006, CG= .008), but the effect size was trivial (g= 0.27).

In terms of pauses after paragraphs, the effect sizes were trivial across all the measures. We next report on the tendencies of these measures despite the trivial effects. Regarding the number of pauses at this boundary, the EG decreased them (Time 1= 1.20, Time 2= .60) while the CG increased them (Time 1= .12, Time 2= .25) with these trivial effect sizes (g= 0.33 and g= –0.21, respectively). The same occurs with mean duration, where the EG decreases the mean duration while the CG increases it. In both cases, as announced before, the effect sizes were trivial (g= 0.41 and g= –0.47, respectively).

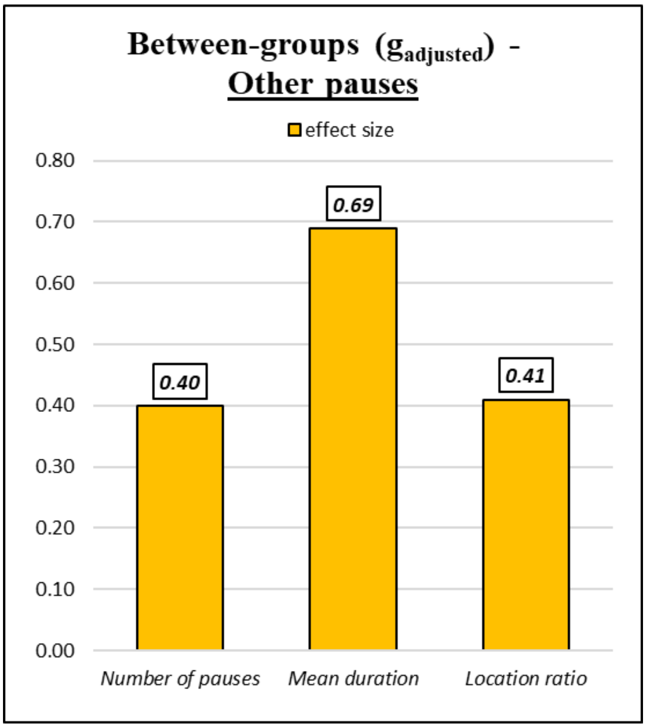
Figure 54.
Between-group comparison with the effect size for after-paragraph pauses.



From a between-groups perspective (see Figure 54 above), the EG devoted more time to pauses after paragraphs than the CG (EG= .41, CG= .28) after WCF, and the EG's location ratio was higher than the CG (EG= .009, CG= .008), both with very small effect sizes (g= –0.46 and g= –0.55). In the case of the number of pauses, the EG paused more than the EG (EG= .60, CG= .25), but with a trivial effect size (g= –0.33).

Regarding other pauses (i.e. those classified as *unknown* and not belonging to any of the previous categories), we found that the EG decreased the number of these pauses (Time 1= 23.20, Time 2= 15.60) with a small effect size (g= 0.69), and the CG decreased them too (Time 1= 14.00, Time 2= 6.12) with a large effect size (g= 1.54). In terms of mean duration, there were no visible differences for the EG (Time 1= 6.62, Time 2= 6.02) while the CG paused for shorter (Time 1= 6.20, Time 2= 4.18), although both with trivial effect sizes (g= 0.23 and g= 0.48, respectively). No effects were observed in the case of the location ratio for either the EG (g= –0.05) or the CG (g= 0.25).

Figure 55.
Between-group comparison with the effect size for other pauses.



The comparison between subjects yielded practically significant results (see Figure 55 above). In terms of frequency, the EG paused more than the CG in this type of pauses (EG= 15.60, CG= 6.12), with a small effect size ($g_{adjusted}$= 0.40). As regards the mean duration, we found that the EG devoted more time to these pauses than the CG (EG= 6.02, CG= 4.18), with a medium effect size ($g_{adjusted}$= 0.69). Finally, the EG had a slightly higher location ratio than the CG (EG= .246, CG= .23) after WCF, with a barely small effect size ($g_{adjusted}$= 0.41).

► ***Summary of results***

Our second RQ asked how young learners' pausological behavior in its multiple dimensions was affected by model texts as WCF. Overall in the writing process, we found that the EG used more time to the whole process than the CG. We also observed that learners not receiving WCF paused much more than the EG. The effect of WCF was also observed in the mean duration of pauses, which was higher for the EG.

Regarding pause location, the effect of the WCF can be observed in the number of within-word and before-word pauses in the EG. Similarly, WCF may have affected the mean duration of before-paragraph pauses, and the location ratio of after-word pauses, which the EG increased. The overall impression of the non-provision of WCF is reflected in the pausological behavior of the CG, which displayed a tendency to decrease the number of pauses and, most importantly, the mean duration, especially at within-word, before-word, and after-word pauses.

CHAPTER 6. DISCUSSION

The aim of the current chapter is to discuss the results obtained in terms of the connection and contribution to previous research in this area. To do so, we will interpret the answer to the research questions guiding the study, namely:

RQ 1. To what extent are there any differences in young EFL learners' temporal distribution of writing processes when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?

RQ 2. To what extent are there differences in young EFL learners' pausological behavior when writing and when revising the same text on the computer, with and without the provision of WCF in the form of models?

Results will be discussed as follows: (1) *findings* will be presented and referred to when significant results are found according to the effect sizes. In case there is a trivial effect size, we will then refer to them as *tendencies*; (2) after the presentation of these findings or tendencies, we will interpret and discuss them accordingly..

6.1. The temporal dimension of writing processes

Our first research question explored what the temporal distribution of writing processes is in young EFL learners' writing, and the extent to which such a distribution changes depending on whether they are provided or not with model texts as WCF. As stated in Chapter 4 (Method), we took into consideration three writing processes: planning, formulation, and revision, which were operationalized with the following measures: frequency, total duration, and in some of the variables, edits, and words produced. There is an empirical motivation to examine these measures: firstly, the need of adding new empirical evidence as regards children's writing processes in an L2 in a digital environment, and second, to what extent the different writing processes are affected by the use of model texts as WCF or the absence thereof.

We found that there were significant differences among the writing processes, in which writers that received WCF in the shape of models modified some of their writing processes. In this regard, the EG seemed to have planned for longer or to have maintained similar amounts of time throughout some intervals of the composition process in contrast to the tendency showcased by the CG group. Bearing in mind the interdependency between the writing processes (see Hayes & Flower, 1981; Kellogg et al., 2013; Roca de Larios et al., 2008), the provision or absence of WCF had a consequent impact on the distribution of some of these processes. The following sections include the interpretation of the results in terms of the planning, formulation, and revision processes in an attempt to reveal to what extent the provision of WCF (or lack thereof) was meaningful.

6.1.1. Planning

We measured planning in terms of frequency of planning episodes and total duration of planning time on the basis of a thorough analysis of previous studies (see Bridwell, 1980; Barkaoui, 2016, 2019; Gánem-Gutiérrez & Gilmore, 2018; Manchón & Roca de Larios, 2007; Sasaki, 2002, 2018). For the purpose of this dissertation, planning was subdivided into two individual subprocesses (i.e. global and local planning), as detailed in Chapter 4 (Method). As previously stated, to our knowledge, this is the first empirical study delving into the intricacies of both types of planning as regards young EFL learners' writing processes, and the purported effect of WCF on planning behavior.

We found that children in the EG reduced the frequency of planning in the first and third intervals of this writing process, while the CG resorted less frequently to planning in the first and second intervals. These differences are attestable in the comparison of the revised version, in which the effect of the provision or absence of WCF was visible across the three intervals. In this case, the EG maintained a higher frequency of planning than the CG. The reduction of planning at the beginning of writing in both groups may be a consequence of the practice effect, given their acquaintance with the writing task (Hayes, 2012). Although both groups reduced the frequency of planning, the EG resorted more to planning than the CG. A potential explanation could be rooted in the assumption that ideas may not have been generated in the CG as much as in the EG in the revised version. Even despite the reduction of planning across intervals in the EG, the higher values may indicate that young EFL writers were pondering over ideas or lexical chunks, precisely what models have been reported to activate in previous research with both children (see Coyle & Cánovas, 2019; García Mayo & Luquin, 2021) or adults (Hanaoka, 2007).

Another point worth mentioning is that although both groups reduced planning at the beginning of the actual composition process, they did engage in more planning at the beginning of the composition process than in the rest of the stages. However, after receiving WCF, the EG resorted more frequently to planning in the middle of the composition process while the CG maintained the initial tendency of concentrating on planning at the beginning. This finding runs counter to Tillema's (2012) study with high school students since it lends support to the claim that low L2 proficient writers do not engage in more planning activities at the beginning. Indeed, the provision of model texts seems to have contributed to allocate more attentional resources in the middle of the composition process for the EG while the absence of WCF seems not to have modified the tendency, i.e. more planning at the beginning.

In terms of the total duration of the planning process, we found that even if trivial effect sizes were observed for the EG, their tendencies show an increase of planning in the middle and end of the composition. However, in the CG, there was a moderate decrease in those stages of the composition. The comparison of the revised version for both groups showed that a major effect was observed in the middle of the writing process. In this case, the EG increased the time spent on planning while the CG decreased it. That writers in the EG increased their planning time could be explained by the purported cycle of text production reflecting planning-execution (Schilperoord, 1996). Thus, after the EG is provided with the model, the writers may have increased their effort in retrieving ideas as suggested in previous studies with adults (see Kormos, 2011; Révész et al., 2019). This would, in some sense, suggest that young writers may follow more or less the same directional patterns as adults while composing their texts, especially after being provided with WCF.

►The effect of WCF on global planning

Moving on to the first subprocess of planning, i.e. global planning, as we operationalized it, that is, in the measures of frequency and total duration (Sasaki, 2001, 2018), both the EG and CG on their own reduced the frequency of global planning in the first and second intervals with medium and large effect sizes. The comparison between groups in the revised version showed that the major effect was found in the middle of the composition process, in which the EG had more global planning episodes than the CG. Regarding the total duration, the EG decreased the time spent on global planning in the first interval with a small effect size. The CG decreased the time spent on global planning in the second interval, and increased it in the third interval, with a small effect size. The comparison between groups in the revised version showed that the major effect was found in the first interval with large effect size, in which the CG had a higher total duration on global planning than the EG. Also, another between-group effect was found in the second interval with a medium effect size. In this case, the EG had a slightly higher total duration than the CG, but the difference is statistically relevant since the EG maintained the same value in this second interval while the CG decreased it.

As other empirical studies have suggested (Breetvelt et al., 1994; Van den Bergh & Rijlaarsdam, 2001), the generation of ideas in the planning process mainly occurs in the initial stages with a strong effect on text quality – which was not studied in this doctoral thesis. The aforementioned decrease of attentional resources to the frequency of global planning at the beginning and middle of the composition process in both the EG and CG individually considered suggests that model texts did not exert any influence over the frequency in the first and third intervals, while it was relevant in the second. We interpret this decrease to be related to, as was expressed by Fayol (1991), the conjugation of setting goals and generating ideas with the constraints of syntax and lexis. More specifically, the EG's decrease of the total time devoted to global planning than the CG might refer to a more comprehensive understanding of the task environment. This goes in line with Faigley & Witte's (1981) study with adults, which indicated that situational variables – for instance, the writer's familiarity with the writing task, the length of the task, or the projected text – contributed to the reduction of the cognitive load experienced by writers as a result of conjugating goal setting and ideas generation, as the writer is already acquainted with the task itself and the text implied by the task. As stated previously, the findings for the CG are indicative of a reduction in frequency in the first and second interval, and a gradual decrease of total time on global planning except for the third interval, where it moderately increases. One might interpret this finding and the overall low degree of frequency of global planning as a decrease in the conjugation of goals and ideas, but the total duration of global planning, which was higher than in the EG, indicates that writers in the CG might have struggled to obtain ideas at the beginning of the composition process, possibly as a consequence of not being provided with the model text. Thus, this points that the cognitive load might have been higher, most likely as a result of the inexistence or addition of new ideas. Previous research within a young learner population (López et al., 2019) has indicated that content planning (i.e. global planning) accounts for nearly 16.34% of children's writing processes.

We might associate the sharp decrease of the total duration of global planning in the first interval in the CG (12% to 8.4%) against the EG (6.4% to 4.6%) with a drastic reduction of ideas. However, the EG devoted a similar total duration to global planning in the middle of the composition process (3.9%) while the CG decreased it considerably (6.0% to 3.8%). In light of this, one might ascertain that writers within the EG did not devote less time to generating ideas after having received WCF, especially in the middle of the writing process. As will be covered in section 6.2., research with adults with digital writing (e.g. Barkaoui, 2019) associates pausing time with global planning, which in turn is related to the generation of ideas and more thought processes.

►The effect of WCF on local planning

In terms of frequency of episodes of local planning, the most important findings for the subprocess of local planning are, firstly, that the EG decreased the frequency of local planning in the third interval while the CG increased it with a small effect size. The comparison between groups of the children's compositional behavior in the revised version showed major differences in this third interval. Additionally, in terms of the total duration of local planning, we also found that the EG increased the time spent on local planning in the second interval whilst the CG increased the time devoted to that process in the first interval. No effects were observed in the comparison between groups in the revised version for the total duration.

The increase of frequency in the CG might be due to a wider focus on grammatical or lexical choices at the end of the composition process, which is associated with this type of planning as attested in adult research on writing processes (Barkaoui, 2019). Conversely, the findings related to the EG cannot be understood without collating them with the ones in global planning, where there was also a decrease in frequency in the third interval. Writers in the EG might have allocated their attentional resources to other processes at the expense of freeing up the central executive (Kellogg et al., 2013) from the planning process. The reduction of local planning episodes in the third interval could be attributed to a focus on macroscopic revisions, as will be further detailed in the corresponding section. This reduction might have freed up resources to pay more comprehensive attention to local aspects of the language, although with different purposes for the EG and the CG. Model texts as WCF have been reported to prompt learners to expand their lexical repertoire as has been claimed in recent studies (e.g. Kim, 2015; Luquín & García Mayo, 2021; Roca de Larios et al., 2015). Nonetheless, our study suggests a rather contradictory view since local planning is associated with lexical and grammatical choices, and hence a reduction in the frequency of this process for the EG may be related to a lack of influence of models in this respect.

As stated above, the findings related to total duration in local planning pointed to a small effect in the second interval for the EG, and a medium effect in the first interval for the CG. As mentioned previously, these within-group findings fade away when the revised version of the text is compared between groups, with no effects observed. Models did not seem to have an influence on the total duration between groups, but there seems to be some important implications in the results of this measure within groups. The EG may have increased the time spent on local planning in the second interval as a result of the participants' nature as emergent planners (Bereiter & Scardamalia, 1987). These has been reported to interweave planning and text production (Cumming, 1989). This is supported by the scarce changes observed in the frequency of local planning episodes in the second interval for the EG, which was scarcely altered. In regards to the CG, the noticeable increase of the total duration of local planning in the first interval may be explained given that L2 writers tend to consume more reflective time (as suggested in previous research with adults, Barkaoui, 2016; Hall, 1990; Silva, 1993). Thus, such an increase could mean that young writers in the CG are striving for focusing on form or lexical retrieval, perhaps due to the absence of the model text. The CG decreased the amount of time devoted to global planning substantially in the first interval, which lends support to a potential search for lexical or grammatical choices. Studies with adults (e.g. Whalen & Ménard, 1995) have suggested that L2 writers tend to focus their attention on morphosyntactic and lexical considerations in contrast to rhetorical and textual dimensions. These interpretations in our study are tentative given the difficulty to know whether the decrease in the number of goals – and hence in global planning – could be associated with the perceivable increase of local planning in this interval. L2 composing per se involves a decrease in the number of goals if compared with the L1 (Skibniewski, 1998). Our tentative explanation that young writers in the CG might have increased local planning at the expense of global planning as a consequence of this focus on grammatical and lexical items should be, however, taken cautiously.

6.1.2. Formulation

We measured the variable of formulation by employing the measures of frequency, total duration, and two other measures, edits, and number of words produced (Van Waes & Leijten, 2003, 2015). The last two are non-temporal measures, but they certainly provided rich information to gauge the depth of what the temporal ones intended to showcase.

►The effect of WCF on the frequency and total duration of formulation

In general, we found that the EG did not change the frequency of formulation distinguishably after the WCF, and thus no effect was observed. Conversely, the CG reduced the frequency of formulation in the first interval with a large effect. The tendency in the rest of the intervals, although not with a significant effect, show that the CG slightly reduced the frequency of formulation. The comparison in the revised version did not show any effect of the WCF on the frequency of formulation between groups.

A possible interpretation of these findings within-groups might be related to the potential effect of models which, in the case of the EG individually, may have pushed writers to maintain similar attentional resources to formulation episodes at the beginning of the writing process. These findings could be explained by several factors: (1) young learners are not usually fluent writers given the constraints of the L2. Additionally, the important pressure on the central executive and the executor (Kellogg et al., 2013) owing to motorically-related issues might be related to the maintenance of the frequency of formulation but also associated with the maintenance of the frequency of local planning. In this sense, shorter pauses at this stage – linked to microplanning processes (see Barkaoui, 2019) – might be reflective of lexical or grammatical choices. (2) L2 adult writers have been reported to produce stretches of within-clause text as a result of sufficient planning (Chukharev-Hudilainen et al., 2019). In our view, findings from this study with an adult population may reflect a similar but not identical pattern of the children's writing in the EG. Our assumption is based on the idea that this production of within-clause text in the planning stage, and the increase in local planning in the second interval, may have subsequently left the frequency of formulation untouchable in the first interval. As will be further detailed, the increase of time spent on the formulation in the first interval in the EG could signal that young writers were concerned with planning in within-clause locations.

Let us move on now to the findings related to the total duration of the formulation. Globally, the EG increased the ratio of total duration in the formulation process (Writing 1= 47%; Writing 2= 53%) with a small effect size. The CG also increased the ratio of total duration in the formulation process (Writing 1= 50%; Writing 2= 60%) with a medium effect size. The comparison between groups of the global time of formulation showed a small effect size, in which the percentage of formulation time was higher in the CG. In terms of the total duration across the three intervals, the EG increased time spent on this process in the first interval with a nearly medium effect size. Although with no significant effect, the tendencies in the EG show that formulation was predominant at the beginning and middle of the process in comparison with planning and revision. The CG increased the time devoted to formulation in the third interval with a medium effect size. Additionally, the tendencies showed that the CG moderately increased the time spent on formulation mainly in the second intervals, while in the first interval no changes were observed. Subsequently, the comparison of the revised version across groups showed a large effect in the third interval, in which the EG maintained a similar amount of time on formulation while the CG increased it.

The predominance of formulation throughout the writing process has been manifested in previous research, especially with adolescent and adult populations (e.g. Roca de Larios et al., 2008) or with children in traditional pen-and-paper environments, such as López et al.'s (2019) study which reported that children, when writing collaboratively, devoted as much as 52% to formulation process. In our study, young writers in the EG devoted around the same figure to formulation, as stated above. However, the presence or absence of WCF did not seem to have had a differential effect over the augmentation of this process, since both the EG (Writing 2 = 53%) and the CG (Writing 2 = 60%) increased it almost in similar terms. The only implications that the previously presented findings might have are related to the fact that the EG significantly increased the time spent on formulation in the first interval. Such an increase might point to compensatory concerns in an attempt to produce meaningful messages (Roca de Larios et al., 2006). In light of this, the increase of time in the first interval runs counter to the belief that the acquaintance with the task environment could be facilitative and aid to reduce the cognitive load imposed by the reflective and thought-related processes regarding formulation. More time spent on formulation in this first interval may indicate that more compensatory concerns were present during text production. This finding is opposed to the fact that the CG maintained the same time spent on formulation in the first interval, and hence leading us to believe that the amount of reflection was not present at this stage of the writing process. Nevertheless, the CG augmented the time spent on formulation in the third interval, which could be related to issues in the executor system (Kellogg et al., 2013). To illustrate this, learners may be writing more slowly, but not necessarily making more pauses. This might be a consequence of typing-related issues beyond the purported constraints of the L2. Research with adolescents (Roca de Larios et al., 2006) has suggested that formulation problems in the L2 are not only attributable to concerns with linguistic issues. This tentatively suggests that young EFL writers may share some common issues with older writers when writing in an L2.

►The effect of WCF on the edits and words produced during formulation

Concerning the edits – which are the number of actions required to produce a stretch of text – no significant effect was observed in the EG. However, one tendency was that the EG decreased the number of edits in the third interval more substantially than in the CG. Nonetheless, this was a trivial effect. We found that the CG decreased the number of edits in the first interval with a very large effect. The tendency in the rest of the intervals also showed a moderate decrease, although there was no effect in terms of statistical analyses. The comparison between groups of the revised text, after the provision of the WCF, showed that there was a small effect in the first interval, in which the EG performed more edits than the CG.

Regarding the explanation of our findings in the EG, we would like to suggest that the maintenance of similar values across the first and second intervals could be inherently related to the degree of transcription fluency (Kellogg, 1994). In their study with adults, Chukharev-Hudilainen et al. (2019) argued that each keypress entails several cognitive processes: content generation followed by "syntactic, lexical, and orthographic processing; [ending] with the motor planning of finger movements" (p. 586). In our study with young EFL learners, a great number of edits may be associated with larger stretches of text since the more keypresses there are, the more edits there will be. Additionally, the number of edits may also account for non-typing issues, especially those related to cursor movements. Hence, Chukharev-Hudilainen et al. (2019) claim that a delay in the upstream processes (i.e. syntax, spelling, content generation, or, equally relevant, keypresses) will propagate down the cascade of cognitive processes and trigger longer pauses. Hence, the maintenance of edits in the EG could be associated with potential disfluencies – which implied decreased transcription rate (Medimorec & Risko, 2016) – in writing as a result of the cognitive effort in remembering the content of the WCF. Regarding the CG, they decreased substantially the number of edits in the first interval as opposed to the EG, which maintained it. We would suggest that the decrease in the CG could be related to the fact that there was no cognitive effort involved in remembering the content of the WCF, and hence the absence of effortful upstream processes (for instance, content generation or lexical retrieval) may have contributed to writing less and reducing the disfluency.

Additionally, the high number of edits maintained in the EG across intervals, and the tendency to decrease in the CG in the second and third intervals, could be attributable to typing issues that young EFL learners generally face during the execution stage (Michel et al., 2019). Both of these aspects are expected to be reflected in the interkeystroke interval measure – which will be further discussed in section 6.2.1.

Concerning the number of words produced, we found that the EG did not produce more, or fewer words given the trivial effect sizes. The tendency shows that there were scarcely any changes in the first interval, while there was a moderate decrease in the second and third intervals, although these were trivial effects. Additionally, we found that the CG decreased the number of words produced in the first and second intervals. The comparison between groups of the revised version of the text showed there was an effect of the WCF in the first interval.

To interpret these findings, some studies have reported that model texts foster attentional and noticing mechanisms being allocated to lexical aspects as well as to ideas and expression. For instance, Coyle & Roca de Larios (2021) found that, while writing collaboratively, children participants noticed lexis after being provided with model texts. Similarly, Luquin & García Mayo (2021), in another study with children writing collaboratively, found that young writers noticed lexical aspects when provided with WCF, but only incorporated discursive, mechanical, and formal features. We could thus tentatively suggest that children in the EG might have maintained a similar number of words in the first interval in contrast to the sharp decrease in the CG as a result of the provision of WCF. Similarly, we believe that models might have slightly sparked the generation of ideas and the incorporation of linguistic structures used to convey the expression of such new ideas – these might be the stylistic features fostered by model texts in the feedback group of Luquin & García Mayo's (2021) study.

As was previously discussed in section 6.1.1., writers in the EG tended to plan more frequently at a local level than globally after the provision of WCF. However, the total duration of global planning was higher than local planning, especially in the first interval. In our view, it could be tentatively suggested that there exists a relation between the higher time spent on global planning and the maintenance of a higher number of words at the beginning of the writing process, which could be well intertwined. This potential association could be interpreted as suggesting that the influence of the WCF induced writers in the EG to ponder over the ideas and lexical items included in the model text and attempt to recall and include them into the revised text. Also, our assumption that they may have created more ideas at the beginning of the writing process could be supported by studies such as Tiryakioglu et al. (2019), in which low-proficient high school L2 writers encountered difficulties in expressing their ideas while searching for the most appropriate vocabulary and grammar to materialize them. Furthermore, in the second and third intervals, the

EG devoted the same amount of time to global planning while fewer words were produced. Hence, young EFL learners in the EG might have strived for the appropriate vocabulary and grammar that would suit the ideas they had in mind (Chenoweth & Hayes, 2001).

6.1.3. Revision

As stated in Chapter 5, we found that the EG decreased the frequency of revisions globally in the first interval, although there was a small effect. The tendency in the EG points to the maintenance of the number of revisions across the second and third intervals. In contrast, the CG decreased substantially the frequency of revisions globally across all intervals with very large effects. The comparison between a group of the revised version showed differences in the first interval, in which the frequency of revisions was higher in the EG than in the CG.

These results indicate that the potential influence of the WCF might have been in the first interval given the reduction of revisions at this stage. In this vein, writers in the EG did not modify the number of revisions considerably in the rest of the intervals, while the CG did. Importantly, previous research (Roca de Larios et al., 2008) with adolescents and adults has stated that writers are expected to plan more if their revisions are lower. Our study contributes to previous research by adding new empirical evidence with an under-represented population, young EFL learners. Our EG, as discussed previously, did not devote less time to local or global planning, but rather maintained it. Nonetheless, the CG did not resort to more revisions even if there was a substantial decrease in the global planning process. Previous research with adult writers in digital environments has suggested that unskilled writers avoided engaging in final revisions given the "proactive investment of efforts" (Xu, 2018, p. 22) or even their reluctance to revise the produced text (Barkaoui, 2007). Our study shows that young EFL learners are more likely to devote equal time to revisions towards the end of the composition process after having received model texts as WCF. Young EFL writers in the EG may have been in a position to make fuller use of the revision process after having completed the whole text. These revisions are likely to fall into the immediate revisions category, which are the commonest type among L2 writers regardless of their proficiency level (Barkaoui, 2016; Stevenson et al., 2006; Xu, 2018).

Regarding the results related to the total duration of the revision process, there were no within-group effects for the EG. However, the tendencies revealed that there was a decrease in time spent on revision in the second and third intervals. Similarly, there were no within-group effects for the CG, whose tendency showcased a decrease in the time spent on revision in the second and third intervals. The major differences, however, were observed in the between-group comparison of the revised version of the text. Here, a moderate effect was observed in the first interval, with the EG increasing slightly the total duration of revision while the CG increased it substantially. The larger between-group differences were found in the third interval, where the EG devoted less time to revision than the CG.

The implications of these findings point to an overall concern of the CG to liberate their capacity of working memory, allowing for the simultaneous allocation of the attentional resources from revising in the third interval to other processes such as local and global planning (Xu, 2018). In the case of the EG, the slight decrease of time spent on revision runs counter to the maintenance of frequency levels of this process. This may suggest that the EG was more selective at the time of performing editions to their texts, and thus more fluent at the time of revision (i.e. more revisions than the CG in less time, especially in the second and third intervals). Recalling Roca de Larios et al.'s (2008) claim that planning is expected to increase if revisions are lower, a potential effect of the WCF seems to have had a bearing on the selective choice of writing processes in the EG. While there was an increase in the time spent on planning in Writing 2 (from 18% to 20%) in the EG, there was a decrease in the revision process (from 9% to 7%). Thus, part of the reduction of the weight of time allocated to revision might be related to an increase in the formulation process (from 47% to 53%) and the planning process associated therewith. This modification of the time spent on different processes after receiving WCF might have entailed an augmentation of writers' concerns about lexical retrieval or grammar checking as part of self-corrections or problem-solving concerns (Roca de Larios et al., 2008; Akyel & Kamisli, 1997), possibly taking place as part of internal revisions (Lindgren & Sullivan, 2006a, 2006b), but which were not part of the research focus in this study. These potential concerns go in parallel with the ostensible profile of L2 writers, mainly centered on lexicon and accuracy (Whalen & Ménard, 1995).

As stated in Chapter 5, within-group findings related to the non-temporal measures such as the number of edits and words produced were the following. Although there were trivial effects in the EG across intervals for the first measure, the tendencies revealed that the EG increased the number of edits after WCF. Conversely, the CG substantially decreased the number of edits with very large effects. Also, there were important effects in the comparison between groups of the revised version of the text. Small effects were found in the first interval, while there were larger effects in the second interval and moderate effects in the third interval.

Concerning the words produced, the EG decreased the number of words in the first interval with a small effect. No effects were observed in the remaining intervals, but tendencies revealed that the EG decreased the number of words produced during the revision process in the second interval and increased it in the third interval. The CG, however, decreased the number of words in the first and second intervals with a moderate effect. The comparison between groups of the revised version of the text showed major differences in the second interval, in which the EG wrote or edited more words.

In line with the previous results, there are important implications, which have to be interpreted cautiously. Despite the decrease of the words produced in the revision process in the EG and CG, it is of note to mention that writers in the EG did not significantly reduce the number of words produced during this process. Indeed, the number of words produced revealed that the EG wrote or edited more words than the CG in less time. Despite the short length of texts produced by young learners in an L2, such a tendency might point to a purported effect of the use of model texts as WCF, in line with the inclusion of new ideas or lexical chunks, as opposed to the claims that shorter L2 texts are indicative of a reduced conceptualization (Stevenson et al., 2006). Nonetheless, this claim is tentative given the number of external and learner-internal factors that come into play in L2 writing in addition to situational variables as attested in previous research with adults (Faigley & Witte, 1981). On the one hand, the higher number of edits in the EG is well aligned with the number of words that they produced, since young EFL learners might have been more concerned with solving language-related issues as a result of the cognitive effort in recalling the different language items contained in the model text provided in the comparison stage. Additionally, model texts seem to have prompted the EG to make use of the available resources in an attempt to reduce the frequency of revision, which allows for focusing on specific chunks of texts for a longer time. On the other hand, another potential interpretation of the high number of edits in the EG, and the differences with the CG, could lie in the less-skilled writers' failure to regulate the writing process effectively, and thus it might be continually interrupted by editing problems that bear little relation to substantive meaning-related changes (Barkaoui, 2007). Likewise, the number of edits might point to motoric problems related to keyboard presses, cursor movements, or situational variables such as rereading the prompt. Such a degree of text modification and production in the EG contrasts with the overall decrease of edits and words produced in the CG, especially in the second and third intervals, which can be attributed to the absence of the model text for the CG.

►The effect of WCF on microscopic revisions

Within-group findings related to the frequency of microscopic revisions indicated that the EG decreased these type of revisions in the first interval, although with a small effect size. The CG decreased the number of microscopic revisions across all intervals with a very large effect size. The comparison between groups of the revised version showed moderate differences in the first interval. These findings point to an effect of the WCF on the frequency of revisions, which will be further discussed.

Regarding the effect of WCF on the total duration of microscopic revisions, the effect sizes were trivial across intervals in the EG. However, tendencies revealed that the EG increased the time spent on microscopic revisions in the first and third intervals, while there was a decrease in the second interval. Much to the contrary, the CG decreased the time spent on microscopic revisions in the second and third intervals, with a small effect size. The comparison between groups of the revised version of the text indicated moderate differences in the second interval, in which the CG devoted more time to microscopic revisions than the EG. Larger differences were observed in the third interval, where the EG spent more time on microscopic revisions than the CG.

The pattern shown in the results indicates that, on the one hand, young EFL writers might have allocated more time to microscopic revisions as a result of the WCF. Previous research with adults (Stevenson et al., 2006) has suggested that writers devoted more attention to microscopic revisions in an L2. Our study contributes to add new empirical evidence by showing that young EFL writers are concerned with this type of revision as much as adults are. Equally relevant, after receiving WCF, the EG increased the frequency of microscopic revisions at the beginning, but also, at the end of the composition process, contrary to the CG. In our study, the number of words produced and edited during the revision process was higher in the EG than in the CG, which points to the pivotal role of the so-called editing revisions (Fitzgerald, 1987). Children in the EG slightly maintained the frequency of microscopic revisions in the third interval, and increased this type of revisions in this interval as opposed to the CG. These findings are well aligned with what previous research has stated. Stevenson et al. (2006) found that for adults "error-triggered language revisions involving spelling, grammar and vocabulary were more frequent in FL" (p. 223). Hence, we could tentatively suggest that these revisions in the EG are focused on solving spelling and grammar issues, which might be a direct contribution of using model texts as WCF. In another study with adults, Barkaoui (2007) found that less-skilled writers tended to restrict their attention to micro-level revisions (i.e. single words, within and between sentences), and correct only local or surface mechanical, grammatical, and lexical problems. This attention to microscopic revisions is made evident in our study with children, especially toward the end of the writing process as a potential effect of WCF.

Another potential explanation as to why the frequency of revisions was maintained in the EG may be well rooted in the children's use of computer writing. Research has cautioned that the effect of computer writing on the scope of revision might be attributed to several variables, for instance, time constraint, L2 proficiency, writing expertise, or the task (Barkaoui, 2007). Our study is well aligned with what early research has attested regarding the influence of the computer setting on revisions. For instance, Bean (1983) found that L1 adult writers were affected by the frequency of revisions, but not by the type or level (i.e. local or global). Van Waes & Schellens (2003) added more evidence to the role of revisions in the writing process by shedding light on how L1 writers distributed the revision process, revealing a focus in the first interval, and not in the third interval. Contrary to these findings in previous research with L1 adults, young EFL writers in the EG resorted more to microscopic revisions in the third interval, and after the provision of WCF, more time was devoted to this type of revision in the third interval. Hence, we could argue that, despite the overall increase of the time spent on microscopic revisions in both groups, the EG maintained higher frequency and total time on this type of revision as a potential consequence of the cognitive effort upon recalling how some lexical items appeared in the model text. From another perspective, certain factors influencing the redistribution of microscopic revisions towards the end in the EG might be due to spelling or typographical issues, and the text per se (Bridwell et al., 1985).

The maintenance of the frequency of microscopic revisions is contradictory with the findings in several WCF studies with young learners, e.g. Coyle & Roca de Larios (2014) or Coyle et al. (2018), which indicate that models do not trigger a high number of revisions but do however activate occasional attention at a syntactic and discursive level. Our study points against this direction since young EFL learners have shown a prevalence of microscopic revisions over macroscopic revisions, as we will see below, which goes in line with more recent research on the effect of models on revised versions of the text (e.g. Luquín & García Mayo, 2021) suggesting the value of models to trigger lexis and content-related noticing, but also to other language forms such as grammar or mechanics. Caution, however, is required when establishing parallelisms between our tentative findings in revision and those of WCF: in most of these studies there is an important methodological difference since participants wrote their texts collaboratively, and the research design included delayed post-tests. Nevertheless, our study has revealed that, after the provision of WCF, children in the EG were still in a position to focus their attention on this type of revision.

In Xu (2018), findings showed that low proficient university L2 writers engaged in immediate and distant revisions of small scopes when writing argumentative texts. Although the task and age are different from our participants', we could tentatively suggest that children in the EG might have resorted very similarly to microscopic revisions after WCF as a result of their proficiency level, and their nature as less-skilled writers in an L2. In contrast, the sharp reduction of the frequency of microscopic revisions in the CG might well be related to the practice effect of repeating the task without having received the feedback stimulus. This is congruent with Barkauoi's (2016) study with adults, which found that low proficient writers opted for overall and precontextual revisions, most of which were low-level-based ones. Our finding that children in the EG kept a higher revision frequency at the end of the process runs counter to Van Waes & Schellen's (2003) study with L1 adults, which found that adult writers do not revise systematically at the end of the writing process. This is an important contribution made by our study, which indicates that children receiving WCF concentrate their revisions toward the end of the process.

►The effect of WCF on macroscopic revisions

No statistical effects were reported for the frequency of macroscopic revisions in the EG. However, tendencies revealed that the EG decreased the frequency of this type of revisions in the first interval and increased them in the second and third intervals. The CG decreased the frequency of macroscopic revisions in the second interval. The comparison between groups of the revised version of the text showcased differences in the first and second intervals, with medium effect sizes, in which the EG resorted more frequently to macroscopic revisions than the CG.

Regarding the total duration of macroscopic revisions, the EG increased the time spent on this type of revisions in the second and third intervals with a small effect size. The CG decreased the time spent on macroscopic revisions in the second intervals. The comparison of the revised version revealed large differences across intervals in the total duration of macroscopic revisions, on which the EG spent more time in the second and third intervals. Conversely, the CG allocated more time to macroscopic revisions than the EG in the first interval.

These findings in terms of frequency and total duration are related to the presence of global revisions over low-level ones. Stevenson et al. (2006) found that there was little inhibition of low-level revisions on high-level ones in L2 writing. Similarly, Barkaoui (2016) found that, even if L2 writers' revisions were more focused on form than on content, the strategic use of revisions was largely dependent on L2 proficiency. Although these studies were based upon adult L2 writers, both coincided that less skilled L2 writers were more concerned with pre-contextual revisions (i.e. at the point of inscription). In our study, the EG slightly increased the frequency of macroscopic revisions in the second and third intervals in contrast to the CG. Such a finding is also associated with the overall increase of time spent on macroscopic revisions, which was higher in the EG. In this respect, previous research (e.g. Barkaoui, 2016; Chenoweth & Hayes, 2001) has manifested that adult L2 writers revised more at a local level but that does not impair the existence of a competition of attentional resources to global issues, which might refer to meaning-related changes or large-scope modifications of the text (Xu, 2018). The maintenance of the levels of frequency and the increase of the time spent on macroscopic revisions in the EG may tentatively suggest that models might have raised the EG's awareness of ideational and lexis-related content, exerting influence over this type of large-scope revisions. Additionally, if less-skilled writers are not constrained by the allotted task time, they are bound to attend to macroscopic revisions during text production, as suggested by Xu (2018). Thus, the increase of the EG on the time spent on macroscopic revisions might be interpreted as young writers not being constrained by the task time, and additionally, the provision of model texts might have contributed to reducing potential constraints on time limitations given that they have more available alternatives or sample linguistic units to use in their own revised texts. In the case of the CG, the increase in the time spent on macroscopic revisions at the beginning of the writing process was not significant given the trivial effect size. Nonetheless, such a tendency might suggest that children in the CG might have noticed certain features while self-editing their texts to a larger extent at the beginning in an attempt to write with more accurate grammar or use other vocabulary items. This pattern was also attested in previous research (e.g. Cánovas et al., 2015; Luquín & García Mayo, 2021) where learners in the CG incorporated features noticed in their revised texts. Yet, in our study, the frequency of macroscopic revisions was much more reduced in the CG than in the EG. One may speculate that after the provision of model texts the EG was more concerned with modifying their texts at a larger scope, e.g. whole phrases or sentences, but at the same time making an effort to recall features noticed in the WCF.

6.2. Pausing behavior in the temporal dimension of writing processes

Our second research question asked about the potential effect of WCF on pausing behavior. Several studies have addressed the issue of how pausing behavior moulds the writing processes themselves (e.g. Barkauoi, 2019; Medimorec & Risko, 2017; Révész et al., 2019), but to our knowledge, none have dealt with the potential effect of model texts as WCF on pausing behavior and its relation to the writing processes in any age group. As explained in Chapters 2 and 3, this is a relevant question because our study contributes to adding new empirical evidence to the body of research on pausological behavior in digital writing with an underrepresented population, in an attempt to shed more light upon how children L2 writers pause, and how their pausing behavior is affected by model texts..

6.2.1. The effect of models on global pausing behavior

To measure global pausing behavior and fluency, we took into consideration the variables corresponding to total process time (i.e., time spent on task), total active writing time (i.e. total time spent on producing text), and total pausing time as expressed in seconds. Additionally, we observed the total number of pauses during the process, the interkeystroke interval (as a measure of motoric and transcription fluency) and pauses per minute (as a measure of fluency). Finally, we also considered the distribution of pause frequency and pause duration (i.e., the mean duration of pauses) across the three intervals in which the writing process was subdivided. As mentioned in Chapter 4 (Method), there is an empirical motivation to examine these measures: firstly, the need of adding new empirical evidence as regards the global pausing behavior of children writers in an L2 in a digital environment, and second, exploring time measures related to the duration of the writing process, and thirdly, fluency measures allow us to delve deeper into our participants' degree of amount of cognitive load.

►The effect of WCF on the duration of the writing process

For the measure of total process time, we found that the EG and the CG decreased the total time spent on the task. The comparison of the revised version of the text revealed a small difference between the EG and the CG. In this case, the EG spent more time on task than the CG. Regarding the time spent actively writing, the EG did not decrease the active writing time while the CG decreased it with a very large effect size. The comparison between groups did not yield any significant difference. Finally, concerning pause time, the EG decreased the total pausing time on task by half with a large effect size. Similarly, the CG decreased the total pausing time with a very large effect size. The comparison of the children's compositional behavior in the revised text did not show any differences between the groups.

The interpretation and empirical implications of these results are relevant for explaining the potential effect of triggering noticing (Swain, 1998, 2005) which models are thought to activate, especially given the substantial difference of time between the EG and the CG in writing the revised version of their text. In essence, young EFL writers in the EG might have engaged in writing their revised version of the text, and as a consequence of the provision of model texts, they might have become more aware of the problems in their interlanguage. That explains why children in the EG devoted a similar amount of time to writing actively despite the significant decrease in the total pause time. In essence, such a finding might be interpreted as a counter-effect maintained throughout the writing process in Time 2, which may have conditioned the higher-level cognitive activities in which the writers may have engaged. As opposed to this, we could tentatively suggest that the CG reduced the time spent on the task given the absence of this triggering noticing effect, even if in some of the processes, e.g. macroscopic revisions, children in this group may have potentially noticed certain features through self-editing. Additionally, the reduction of the time spent writing actively and the plummeting pausing time points to a number of factors, which include the lack of ideas to be expressed – and hence less cognitive pressure – as well as a further deactivation of higher-level processes, reported to consume more active writing time and pausing time.

Based on the above, the WCF might have had a slight effect on pausing time. Pauses are traditionally related to high-level cognitive operations involving planning or content in writing processes terms (see Chenu et al., 2014; Olive et al., 2009; Medimorec & Risko, 2017; Barkaoui, 2016, 2019), as was accounted for by models of writing (see Bereiter & Scardamalia, 1987; Flower & Hayes, 1980; Hayes & Flower, 1981). Pauses are considered as traces of "covert cognitive processes" (Barkaoui, 2019, p. 530). In any event, the apparent reduction of pausing time throughout the composition process might indicate that writers in the EG did not experience the same amount of cognitive load as when writing before receiving WCF. In Writing 1, the EG devoted nearly as much as 58% of the whole composition time to pausing, which drastically diminished after having received WCF. This percentage is broadly consistent with empirical evidence in adult research that states that pauses amount to nearly half of the total writing time (Alamargot et al., 2007; Alves et al., 2007; Medimorec & Risko, 2017). Young writers decreased the amount of time devoted to pausing after the provision of WCF to 48%. A purported explanation for this changing situation might be related to the reduction in problem-solving and other related issues.

►The effect of WCF on the number of pauses, transcription fluency, and fluency

Regarding the total number of pauses throughout the writing process, the EG decreased the number of pauses with a medium effect size while the CG decreased it substantially with a very large effect size. Nevertheless, the between-group comparison eliminated any differences. In terms of motoric fluency, there was no effect of the provision or absence of WCF on the measure of interkeystroke interval for neither group. Finally, cognitive fluency seemed to be largely affected since the EG maintained a similar number of pauses per minute while the CG increased it with a very large effect. The between-groups comparison displayed a large difference, in which the CG paused more per minute than the EG.

These results have several implications. Firstly, the reduction of pauses from Writing 1 to Writing 2 occurred in both the EG and the CG. However, the reduction was more marked in the CG. In light of this, while the fact that children in the EG paused more in Writing 1 may be associated with potential hesitant phases while writing, as suggested in studies with adults (see Alamargot et al., 2007; Spelman Miller, 2006a; Van Waes & Leijten, 2015), the decrease of pauses in Writing 2 in the EG may point to benefits at the level of managing attentional resources as a result of the provision of WCF. However, this claim is tentative since the CG also decreased their number of pauses, and such a decrease could also be associated with a better management of attentional resources. Nevertheless, other individual factors such as typing skills are to be taken into account since young EFL writers – either in the EG or CG – have not completely automatized motor output (e.g. for handwriting), and thus keyboarding is likely to impose a high cognitive demand on both the central executive and the executor, and on working memory (Bourdin & Fayol, 1994; Olive & Kellogg, 2002; McCutchen, 1996; Kellogg et al., 2013). These keyboarding skills are not only associated with the number of pauses but also with the measures of interkeystroke interval (i.e. motoric fluency) and pauses per minute (i.e. cognitive fluency). As mentioned in the first paragraph of this section, the milliseconds spent between intervals were not affected either within or between groups, and hence the effect of the provision or absence of WCF is not visible. Such a finding indicates that there were no significant inter-personal variations in terms of typing skills. As regards cognitive fluency, that is, pauses per minute, findings point to a potential effect of the WCF in maintaining a similar amount of pauses per minute – possibly as a result of the cognitive effort involved – while the CG paused significantly more per minute. This finding might be interpreted as a cognitive struggle on the CG's part given the absence of WCF, and thus, more cognitive load might have been placed on the central executive (Kellogg et al., 2013) to perform either anticipatory processes such as planning, or the executor in terms of motoric issues. All of these reasons are related to the role that working memory plays in children's writing performance, which may affect the distribution and resources allocated to the different writing processes. Additionally, working memory may have been affected by the provision or absence of WCF.

►The effect of WCF on the distribution of pauses across intervals

Our findings revealed that the EG decreased the number of pauses across intervals, with the higher changes in the first and second intervals with a medium and large effect size. The tendencies are similar after children are provided with WCF: there is a peak of pauses at the beginning, with a slight decrease in the middle and toward the end. The CG substantially decreased the number of pauses across the three intervals with a larger effect size. The comparison of the revised version of the text between groups did not indicate any differences, and thus, no effect of the WCF was observed statistically.

Regarding the mean duration of pauses across the three intervals, our study revealed that the EG was not affected by the WCF since values were similar in Writing 1 and Writing 2 after the provision of WCF. Conversely, the CG decreased the mean duration of pauses across the three intervals with large effect sizes, especially in the first and third intervals. The between-group comparison of the revised version indicated substantial differences in the second interval, with medium effect size, in which the EG paused longer than the CG. In the third interval, there was a large effect since the EG paused longer as well.

These findings run counter to other tendencies revealed by previous studies with adults (e.g. Barkaoui, 2019), in which participants paused less but longer at the beginning, whilst in the middle and end of the composition process, pauses were more frequent though shorter. In our study, however, young writers in the EG paused more frequently at the beginning of the process both before and after receiving WCF but for shorter periods. Conversely, children in the CG displayed lower mean duration values but maintained the first interval as the one in which they paused more frequently and for longer. The reduction of the mean duration of pauses and the frequency of pauses across intervals in the CG may be related to a certain cognitive struggle, since they paused more per minute, as mentioned in the previous section. While the tendencies exposed showcase very clearly that a series of high-level cognitive operations were occurring both at the beginning and middle of the writing processes in the EG and CG, such as planning the overall content or structure of responses (Wengelin, 2006, 2007), the peak in pause duration in the middle of Writing 1 for both the EG and CG points to potential problems in the formulation of ideas as well as the rereading of the prompt or sections of the text, similarly to L2 adult writers (Barkaoui, 2019). Interestingly, the decrease in the frequency of pauses at the end of the writing process in the EG and the time spent on them suggest that children had to revise their text production at that point. EG writers paused longer but less frequently in Writing 2 than the CG, which might indicate why there was an increase in the time spent on macroscopic revisions as was evidenced in section 6.1.3.

6.2.2. The effect of models on pause location

In this section, we will discuss the results regarding pause location and the different associated measures. Thus, to ensure full comprehension of the effects of the WCF on each of these pause locations, we will proceed by analyzing and discussing each pause location individually, which are: *within-word pauses*, *before-word pauses*, *after-word pauses*, *before-sentence pauses*, *after-sentence pauses*, *before-paragraph pauses*, *after-paragraph pauses*, and *other pauses*. All these variables are going to be discussed from the perspective of different levels of measurement: number (of pauses at this boundary), mean duration, and location ratio.

►The effect of WCF on pauses at word boundaries

To start with, regarding within-word pauses, the EG decreased the number of pauses with small effect size and maintained the same mean duration of pauses at this boundary. Although non-significant, the EG decreased the ratio. Conversely, the CG maintained the same number of within-word pauses in Writing 1 and 2 but decreased the mean duration of pauses at this boundary with a medium effect size. The tendency showed that, for the location ratio, the CG increased it with a trivial effect size. The comparison of the revised version of the text revealed a small effect size across the three measures for pauses within words.

The fact that writers in the EG decreased the number of pauses while the CG maintained a similar number, points to a possible effect of the WCF as a result of being able to solve spelling-related issues more efficiently, similarly to adult L2 writers (Barkaoui, 2019). However, the EG did not significantly reduce the mean duration of these pauses, which might suggest existing difficulties related to keyboarding skills at the time of recalling how a specific word was written, and thus resulting in slow typing speed (Alves et al., 2007; Barkaoui, 2019).

Regarding pauses before words, we found that the EG decreased the number of pauses at this boundary with a medium effect size. No statistical effects were observed for mean duration or location ratio in before-word pauses. The CG decreased the number of pauses at this boundary with a large effect and decreased the mean duration with a medium effect size. No changes were observed in the location ratio, although the tendency revealed a decrease in this measure for before-word pauses. The comparison of the revised version between groups showed a major difference in the mean duration of pauses before words, with a large effect size.

These findings point to important implications concerning the potential effect of model texts as WCF. Firstly, previous research with adults (Barkaoui, 2019; Schilperoord, 1996; Spelman Miller, 2006a; Wengelin et al., 2009) has confirmed that L2 writers tend to pause more frequently at word boundaries. Our study contributes to adding new empirical evidence with children, supporting these findings in studies with an adult population, since the location ratio for pauses before words amounted to 46% of the pauses in the EG, and 42% in the CG. Even though the difference between the number of pauses before words did not have a statistical effect, the tendency showed that the EG paused more before words (Time 1= 41.30; Time 2= 27.20) than the CG (Time 1= 22.63; Time 2= 13.00). However, the decrease in number is evident in the case of the EG, which points to a potential effect of the absence of the WCF. Hence, we could tentatively suggest that, as before-words pauses have been associated with local planning in previous research with adults (see Alves et al., 2007; Barkaoui, 2019; Van Waes & Leijten, 2015), young EFL writers in the EG had to resort less frequently to this type of planning as has been attested in section 6.1.1.

Moving on to the results of pauses after words, no statistical effects were observed for the EG. However, the tendencies revealed that the EG increased the number of pauses, the mean duration, and the location ratio of pauses at this boundary. Conversely, the CG decreased the mean duration of after-word pauses, while no effects were observed for the rest of the measures. The tendencies revealed that the CG maintained a similar number of pauses and location ratio at this boundary. The comparison of the revised version of the text between groups indicated small differences in the number of pauses after words, in which the EG paused more frequently than the CG. Also, medium differences were found in the location ratio since the EG increased it while the CG maintained it.

These results might be interpreted from the perspective of what cognitive processes are associated with these pauses. Previous research with adults (Van Waes & Leijten, 2015) has suggested that after-word pauses are related to rereading, monitoring as well as local planning. In our study, the effect of the WCF might have contributed to increasing slightly the number of pauses after words as a result of a stronger need of managing attentional resources towards the previously produced text – for instance, pauses after words might be associated with the revision process. Nevertheless, the mean duration in our study is much shorter after the provision of WCF, which is aligned with the decrease of time spent on the revision process as stated in section 6.1.3.

►The effect of WCF on pauses at sentence boundaries

Results regarding before-sentence pauses revealed no statistical effects in the EG, although the tendency showed a decrease in all measures (i.e. number, mean duration, and location ratio). The CG increased the location ratio of pauses at this boundary, but the rest of the measures were not affected according to the effect sizes. The tendencies revealed that the CG increased the mean duration of pauses, and no changes were observed in the number of pauses. The comparison of the revised version between groups indicated moderate differences in the mean duration of pauses before sentences, in which the CG paused longer at this location than the EG, and large differences in the location ratio, where the CG had higher values than the EG.

Regarding after-sentence pauses, no statistical effects were observed for the EG. The tendencies, however, showed that the EG increased the number of pauses, the mean duration, and location ratio of pauses at this boundary. This pattern was maintained in the CG, whose effects were trivial. The tendencies revealed that the CG increased the number of pauses, the mean duration, and location ratio of pauses at this boundary. The comparison of the revised version of the text between groups indicated moderate differences in the number of pauses (the EG had more pauses at this boundary than the CG), and large differences in the mean duration, in which the EG paused for longer at this boundary.

Previous research with adults has suggested that pauses at sentence boundaries are likely to reflect higher-order processes. For instance, Révész et al. (2019) found that adult L2 writers were more likely to pause between sentence boundaries owing to anticipatory processes such as global planning or rereading longer stretches of text. Similarly, Chukharev-Hudilainen et al.'s (2019) study with adult L1 and L2 writers revealed that, when pausing at after-sentence boundaries, they were likely to look at the previous sentence. The potential explanation to this phenomenon is, according to the authors, that L2 writers are prone to looking at the previous sentence as a memory refresher which, at some point, might be a distractor from the "idea package" (Chukharev-Hudilainen et al., 2019, p. 602). In our study, the WCF seems to have contributed to more monitoring of the previously written text, given the slight increase of the values in this boundary across measures in the EG. The situation varies in the case of before-sentence pauses since children made little use of such type of pauses, which the WCF contributed to decreasing, as observed in the findings related to the EG. This was observed in the location ratio, in which the CG increased moderately in contrast to the EG. As can be seen, the role of pauses at sentence boundaries is rather exiguous throughout the writing process.

►The effect of WCF on pauses at paragraph boundaries

The effect sizes for before-paragraph pauses did not yield any significant result either within or between groups, as stated in Chapter 5. However, the tendencies showed that the EG increased the number of pauses, the mean duration, and the location ratio of pauses at this boundary. In the case of the CG, the tendencies revealed a decrease across all three measures.

Regarding the effect of WCF on after-paragraph pauses, no effects were observed for the EG. The tendencies revealed that they decreased the number of pauses, the mean duration, and location ratio at this boundary. Likewise, no effects were statistically observed for the CG. The tendencies, however, revealed that the CG increased the number, the mean duration, and the location ratio of after-paragraph pauses. The comparison of the revised version of the text between groups indicated differences with a small effect size in the mean duration (the EG paused for longer at this boundary than the CG), and location ratio (the EG had a slightly higher value than the CG).

Previous research with adult L2 writers (Barkaoui, 2019; Medimorec & Risko, 2017; Spelman Miller, 2006a; Wengelin et al., 2009) has associated this type of pause with global planning. Nonetheless, although pauses at paragraph boundaries might reflect such a writing process, children are more likely to write one-paragraph texts as opposed to adult written output, which is likely to be composed of more than one paragraph because of L1 writing literacy training. Additionally, these studies with adult L2 writers used narrative and argumentative tasks with different time constraints, and naturally, varied L2 proficiency levels. Children writers in our study had to write a text based upon a six-picture frame story which, in essence, gave them the narrative of what they were seeing, reducing cognitive struggle, and relegating it to the grammatical and lexical sphere. Besides shedding light on a different level of L2 proficiency and a different type of task, our study contributes to adding further empirical evidence of how children L2 writers behave concerning pausing at paragraph boundaries. Despite the absence of effect of the WCF, tendencies revealed that the EG increased slightly before-paragraph pauses after receiving model texts. As it was explained in previous sections, young learners, either having received WCF or not, devoted much less time to global planning. Thus, it would seem contradictory to subordinate this augmentation of pauses before paragraphs to such a macroplanning process; however, another possible interpretation might be that learners went backward to other points in the text in an attempt to reorganize ideas or content. Likewise, we could tentatively suggest that young writers in the EG, after performing macroscopic revisions, were induced to plan beforehand what they were going to write after this revision.

►The effect of WCF on pauses at other boundaries

Studies on pausing behavior with an adult population have not included other pauses in specific locations or "not automatically identified" (Barkaoui, 2019, p. 539) by Inputlog (for instance, Conijn et al., 2019; Medimorec & Risko, 2017; Révész et al., 2019). In our study, given the absence of studies exploring these unknown pauses in adult as well as children populations, we opted for their inclusion. As stated in Chapter 5 (Results), the EG decreased the number of other pauses with a small effect size. The tendency in the rest of the measures was the maintenance of mean duration and location ratio. The CG substantially decreased the number of other pauses with a large effect size. The tendency in the rest of the measures was that the CG decreased the mean duration and location ratio of other pauses. The comparison between groups revealed small differences in the number of pauses, in which the EG had a higher number than the CG. Regarding mean duration, there was a medium effect size since the EG had longer pauses at these boundaries than the CG. Finally, there was a small effect size for the location ratio since the differences between groups were very weak.

In our study, a close examination of this type of other pauses revealed that they may be associated with a myriad of processes ranging from pauses between revision operations such as inserting or deleting information to pauses that happen between two cursor movements, just as was suggested in previous research with adults (Barkaoui, 2019). The effect of the WCF is more clearly visible in the mean duration since the EG maintained a similar duration while the CG decreased it. Similarly, the decrease of the location ratio in the CG is linked to the large decrease in the number of pauses. One potential explanation as to why the EG paused for longer in this category might be related to essential operations which may not be associated with the effect of the WCF, such as the existing need of maintaining operations as cursor movements. Another alternative to interpreting the differences between groups might be related to the decrease observed in the frequency and edits in the revision process in the CG. Revision pauses fall under this category of other pauses since these include keyboard presses (such as the [BACK] keypress). A decrease in their number might be related to a subsequent decrease in the total time spent on the revision process.

Pausological studies with adult L2 writers have stressed the importance of other pauses as being the second most frequent pause location. Our study points to a similar tendency as shown in these studies, such as Barkaoui's (2019), in which the participants followed this pattern. This result in the present research adds new empirical evidence regarding these types of pauses in an underrepresented population such as young EFL writers. After the provision of models, a very weak effect on the mean duration and location of other pauses was observed since the EG did not diminish the values as opposed to the CG. A potential explanation might be that due to the purported constraints imposed by computing and keyboarding skills reported in studies with low proficient writers with low-level typing skills (e.g. Bourdin & Fayol, 1994; McCutchen, 1996; Olive & Kellogg, 2002), young EFL writers might have had to resort to considerable cursor movements as well as a higher degree of revision.

Another important implication in our findings relates to the text boundary effect as evidenced in studies with adult L1 and L2 writers (e.g. Alves et al., 2007; Chanquoy et al., 1996; Medimorec & Risko, 2017; Schilperoord, 2001). This text boundary effect implies that pause length increases from smaller to longer discourse units (Wengelin et al., 2009). Our findings of pauses did not successfully reflect the text boundary effect. A potential explanation to support this result might be related to the participants' nature as young EFL learners since, in this case, young writers might not have paused for longer before sentences or paragraphs, as discussed in the previous paragraphs. Hence, potential higher cognitive processes related to global planning – which in adult L2 pausological research has been associated with pauses at sentence and paragraph boundaries – have likely been concentrated on word boundaries as proof of cognitive demands. The age factor, but also L2 proficiency, might explain why such a text boundary effect was not reflected in the data with children. The provision or absence of WCF does not seem to have contributed to this text boundary phenomenon.

6.3. Synthesis of the main findings

What follows is a synthesis on the main findings obtained in this PhD, following the RQs that guided this study: a) the effect of model texts as WCF on the temporal dimension of writing processes, and b) the effect of model texts as WCF on the pausological behavior of young EFL learners.

6.3.1. The effect of model texts as WCF on writing processes

The effect of WCF was mainly observed in the subprocesses of planning and revision. Young EFL learners receiving WCF significantly reduced the frequency of planning at the beginning and end of the writing process, but spent more time on the planning process during writing, hence adding to the alleged contribution of model texts as fostering children's ideas and lexical units. This is clearly reflected in the fact that children in the WCF group performed fewer global planning actions especially in the middle of the writing process although with a higher value than the CG. The percentage of time spent on this process was still high, thus indicating a cognitive struggle to find ideas that suited the writers' needs (Fayol, 1991). In contrast, writers in the WCF group resorted more to local planning than the CG even if they decreased the frequency of episodes. This finding points to the purported benefit of model texts to attention-to-form as well as expanding the lexical repertoire (Eschholz, 1980; García Mayo & Labandibar, 2017; Roca de Larios et al., 2015). Thus, the maintenance of local planning episodes in the WCF group adds further evidence to the initially claimed view that children are emergent planners (Cumming, 1989), as previous studies with low proficient L2 writers have suggested (Manchón & Roca de Larios, 2007; Whalen & Ménard, 1995). As regards the process of formulation, previous research with adult writers had indicated that it amounted to more than half the time spent on writing (Roca de Larios et al., 2008; Gánem-Gutiérrez & Gilmore, 2018, among others). On the basis of this, we have concluded that young learners concentrate most of the formulation process at the beginning and middle of the composition process. In this case, such a tendency was not considered as an effect of the provision or absence of model texts as WCF since both the EG and CG maintained a similar pattern. Accordingly, young EFL writers in both groups are likely to have increased a similar amount of time in the formulation process owing to compensatory concerns (Roca de Larios et al., 2006), and most possibly, keyboarding issues (Kellogg et al., 2013) which delayed their typing process. Building on this, and supported by the following findings, children in the WCF group needed slightly fewer actions to complete their texts. The number of edits was still higher in contrast to the CG, which decreased edits as a result of less effortful upstream processes which did not seem to reflect cognitive struggle (Chukharev-Hudilainen et al., 2019; Kormos, 2019). Children receiving WCF also produced larger stretches of text than the CG, especially at the beginning of the formulation process. In light of this, young EFL learners might have augmented the lexis or ideational content after having received WCF, thus supporting the increase of time spent on formulation at the beginning as a possible consequence of recalling or searching for new vocabulary or expressions (Chenoweth & Hayes, 2001).

As stated above, revision (together with planning) is the writing process where the effect of the WCF was more noticeable. The process of revision has led us to conclude that young writers receiving WCF engaged in revising their text at the end of the composition process more than in the rest of the intervals, as in Chenoweth & Hayes (2001). Similarly, children in the EG congruently produced more words at the end of the revision process, thus pointing to benefits at the lexical level (Roca de Larios et al., 2008). Young EFL learners in the EG considerably decreased the time spent on revision in the middle of the writing process, possibly related to the increase of formulation in this interval. Along this line, young EFL writers in the EG spent more time on microscopic revisions at the middle and end of the writing process, thus pointing to substitution of grammatical items and lexis, supporting Stevenson et al.'s (2006) claims about adult L2 writers, and Barkaoui's (2007) predictions that less skilled adult writers concentrate their attention on surface-level changes. In addition to these reasons, young learners in the EG could have devoted these microscopic revisions to the third interval of the writing process owing to the nature of the medium of writing, which Van Waes & Schellens (2003) posited as triggering revisions in that interval. In relation to this, the large number of edits suggests that young learners in the EG failed to regulate their composing and revision processes, supporting Barkaoui's (2007) claims. As for macroscopic revisions, young learners in the EG resorted to this type of revisions at a more frequent rate towards the end of the process, although differences were more marked at the beginning and middle of the writing process in comparison with the CG. Similarly, writers receiving WCF increased the time spent on macroscopic revisions in comparison with the CG, suggesting that models might have contributed to raising EG's awareness of ideational content, and hence they may have engaged in more large-scope revisions.

6.3.2. The effect of model texts as WCF on pausological behavior

Overall, children in the EG reduced pausing time as well as the CG, although the decrease was sharper in the latter. This decrease of pausing time indicates that the cognitive load when writing the revised version in both the EG and CG might have been reduced, but reflective operations were still required possibly as a result of higher-level cognitive operations (Barkaoui, 2016, 2019; Chenu et al., 2014; Medimorec & Risko, 2017). Additionally, children in the EG decreased the number of pauses, which implies a more beneficial allocation of attentional resources (Alamargot et al., 2007; Spelman Miller, 2006a; Van Waes & Leijten, 2015). Even if both groups decreased the number of pauses, the number was higher in the EG than in the CG, which might also be related to keyboarding issues imposing a cognitive load on the executor (Kellogg et al., 2013).

Findings related to the distribution of pauses indicated that the effect of the WCF was observed in that the EG paused more frequently across the three intervals, and for shorter periods of time, than the CG. Such a tendency seems to underlie potential overall planning (Wengelin, 2007) at the beginning, as well as issues when translating ideas into text (Barkaoui, 2019) in the middle of the composition process. Similarly, the increase in mean duration in the EG towards the end suggests their monitoring of the text produced thus far as well as more revision.

With regard to pause location, young EFL learners in the EG paused less frequently before words, which is suggestive of a decrease in the frequency of global planning (Alves et al., 2007;

Barkaoui, 2019). Typing issues or spelling-related aspects seem to be a concern after receiving model texts as WCF since the EG resorted more frequently to within-word pauses than the CG. Similarly, children resorted more to after-word pauses after the provision of WCF, indicating a possible increase in rereading the previous text, monitoring (Van Waes & Leijten, 2015), or planning locally. At the sentence level, young learners devoted more time to after-sentence pauses, suggesting a potential effect of WCF on monitoring (Révész et al., 2017, 2019) as well as other higher-level cognitive operations such as global planning. At paragraph boundaries, the EG increased pauses before paragraphs, thus inducing us to consider that they planned their content prior to revising macroscopically, in contrast to global planning, as other studies with adults suggested (e.g. Medimorec & Risko, 2017; Wengelin et al., 2009). Children in the EG resorted less frequently but for the same amount of time to other pauses related to revision operations such as deleting or inserting, cursor movements or similar to these ones (Barkaoui, 2019). This may point to an equal need of maintaining operations related to these other pauses. Finally, neither the EG nor the CG were subject to the text boundary effect as in other studies (e.g. Wengelin et al., 2009; Chanquoy et al., 1996; Schilperoord, 2001). Pause length was particularly erratic across all locations, which might imply a potential influence of the age factor given the participants' nature as young EFL learners. Another possible explanation points to L2 proficiency as a variable influencing the length of pauses, which is higher in adult L2 writers (Barkaoui, 2019).

CHAPTER 7. CONCLUSION

The present chapter is structured as follows. Firstly, a summary of the rationale of our study will be presented. Second, we will provide some concluding remarks on what contribution our study has attempted to make, followed by the pedagogical implications that emerged out of the results obtained. Finally, a series of limitations will be acknowledged, and future research avenues will be suggested for the area of our study.

7.1. Summary of the rationale of the study

The ultimate aim of the present PhD dissertation was to advance and contribute to existing research in the following main areas: a) the study and characterization of children's L2 writing processes and pausological behavior, and b) the manner in which these variables are affected by WCF in the shape of models in a digital environment. Accordingly, our study has explored methodological alternatives such as logging tools in a fully digital environment to observe these phenomena unobtrusively.

The main contribution that this PhD dissertation has aimed to make is to shed light on young EFL learners' implementation of writing processes and their pausological behavior as a function of the provision (or absence) of model texts as WCF. Our study has attempted to provide new empirical evidence on these two areas in digital writing by young EFL learners, an underrepresented population in contrast to previous studies on writing processes, where participants were high school students (Roca de Larios et al., 2008; Tiryakioglu et al., 2019) or adults (Barkaoui, 2016, 2019).

In terms of the first area of research of this doctoral thesis, the temporal dimension and the interplay of writing processes have been widely researched in adult and high school L2 writers (see Barkaoui, 2007; Gánem-Gutiérrez & Gilmore, 2018; Manchón & Roca de Larios, 2007; Roca de Larios et al., 2002, 2008; Zimmerman, 2000). However, no studies have probed into characterizing young EFL learners' writing processes and pausological behavior, and how they manage them in a completely digital environment. Hence, our study has attempted to fill this gap by providing an overview of how children strategically allocate their writing processes and pausological behavior to the time-distributed nature of the composition process.

Regarding the second area of research of this dissertation, model texts have demonstrated to be an ecologically valid technique to enhance L2 acquisition (e.g. Cánovas et al., 2014; Coyle et al., 2018; Coyle & Cánovas, 2019; Luquín & García Mayo, 2021) prompting learners to establish cognitive comparisons (Ellis, 1995) and directing their attention to form, lexical chunks or linguistic units. Our study has contributed to bringing to light how model texts as WCF mediate how young learners make use of their attentional resources in each writing process when models are used as a WCF technique from the perspective of individual writing and in digital writing. Concerning this, our study has attempted to fill the existing gap given the existence of studies in paper-based environments with collaborative writing by young EFL children, but the absence of studies in digital environments with individual writing in this population.

7.2. Contribution of our research

From the theoretical and empirical standpoints, the contribution of this thesis is related to the disciplinary debates on (1) the temporal dimension of writing processes and pausological behavior in young EFL learners (as detailed in Chapter 2), (2) the nature of writing processes in a digital environment, and (3) the potential influence of WCF on L2 writing processes.

In the area of writing processes, the main contribution of this empirical study derives from the light our research has attempted to shed on the behavior of young children when composing in an L2, and more importantly, how text production develops in a fully digital environment. Our study confirms that model texts as WCF allow learners to place their attentional resources to local planning, i.e. to lexical chunks or reduced linguistic units, which goes in accordance with their nature as emergent planners (Cumming, 1989). Another contribution of this dissertation has been to clarify the role that models hold to foster new ideas. This, in turn, has contributed to adding new empirical evidence to the purported effect of this type of WCF on the process of L2 writing given the maintenance of words produced in sharp contrast to the group not receiving WCF.

Our study partially confirmed what previous research with adults and adolescents indicated on the predominant role of formulation while composing L2 texts (Roca de Larios et al., 2006, 2008; Gánem-Gutiérrez & Gilmore, 2018). Young learners in both the EG and the CG devoted more time to planning than what previous studies have shown in the case of adult and high school EFL learners. Similarly, our findings concerning revision revealed this process to be affected by the WCF greatly in terms of frequency. In this regard, our results indicated that revision took place fundamentally in the last stages of the composition process, while the absence of WCF affected the frequency of revisions to a great extent. Yet, young children did not initially display such a distribution pattern before receiving WCF, where revisions were mostly concentrated at the beginning and end of the writing process. After the provision of WCF, young children increased the time spent on revisions at larger scopes (macroscopic revisions) while maintaining revision values at small scopes (microscopic revisions). This indicates that young children were more prone to resorting to the revision process after having received WCF. Equally relevant, the greatest increase in the EG was observed in macroscopic revisions, hence suggesting that the influence of model texts was markedly visible in the incorporation of ideational content while the maintenance of the frequency of microscopic revision might be related to lexical and grammatical units (Coyle & Roca de Larios, 2014; García Mayo & Labandibar, 2017; Hanaoka, 2007; Martínez Esteban & Roca de Larios, 2010Roca de Larios & Martínez, 2010). Therefore, young children seem to have selected – either consciously or unconsciously – their strategic behavior as to when and where they opt for locating each of the writing processes (Roca de Larios et al., 2008).

Finally, our research contradicted some views on pausological behavior according to previous studies with adult writers (see Alves et al., 2007; Barkauoi, 2019; Medimorec & Risko, 2017; Wengelin et al., 2009) in that the text boundary effect did not seem to apply. Indeed, an interesting contribution to the disciplinary debate on pausing behavior is the great amount of time that young EFL learners in both the EG and CG devoted to smaller linguistic units. Along the same line, our results also confirm the purported cognitive load imposed on young learners as theorized in models of writing such as Flower & Hayes (1980), Hayes & Flower (1981), or the revisited model by Hayes (2012), as the EG paused less frequently but at a higher rate than the CG. Our research also contributed to shedding light on children's fluency measures when writing as: (i) the provision or absence of WCF did not seem to affect motoric fluency (i.e. interkeystroke intervals) since no changes were observed in either group, and (ii) model texts may have contributed to maintaining the pausing time in the EG, and thus a slightly less cognitive struggle might have been present.

7.3. Pedagogical implications

The findings in the present empirical study lead to put forward a series of pedagogical implications. The study of writing processes has been inherently linked to the concept of self-regulation (Roca de Larios et al., 2008), suggesting that learners model their compositional behavior in different ways depending on their L2 proficiency. Learners increase the degree of interaction between writing processes as their L2 proficiency increases. Our results revealed that young learners handle each writing process more recursively than what we expected following previous research with other populations (e.g. Roca de Larios et al., 2008; Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2019), especially bearing in mind the participants' age. Pedagogically speaking, a focus on processes has been greatly neglected in terms of integrating the empirical findings into writing instruction (e.g. Paulson Gjerde et al., 2017), in addition to the factual challenge in implementing it into the curriculum (Vandermeulen et al., 2020).

This recursiveness of writing processes is supported after observing the changes in the writing processes upon the provision of model texts as WCF. This might indicate that processes of attention to given aspects of text production are conducive to facilitating the development of the text, thus relating it to the problem-solving and complex nature of writing (Manchón, 2013, 2019). We have argued in previous sections how young learners have resorted more to planning locally, formulation, and microscopic revisions after receiving WCF, with varying frequency across all these writing processes. Given the interaction between writing processes, writing instruction should not only be made pivotal in the L2 language learning environment, but it should also take into consideration the strategic role of these processes, fostering learners' reflection on how they regulate them. The provision of feedback has demonstrated that young learners modify their attentional resources according to their needs. Gaining an understanding of where young EFL learners allocate the writing process both temporally and strategically ensures that teachers might direct their attention to very specific aspects while instructing L2 writing (see Carless & Boud, 2018). In sum, teachers are advised to gain awareness of where and how children allocate their writing processes both temporally and strategically. This awareness seems to be important for teachers so that they understand their students’ writing behavior when engaged in writing in regular class periods and exams. Moreover, teachers are advised to provide models as WCF to young learners. Indeed, models have allegedly been shown to generate a beneficial effect on students’ strategic modification of attentional resources in their writing processes according to their needs.

7.4. Limitations and suggestions for future research

In spite of the potential contribution to current disciplinary debates on children's L2 writing processes and pausological behavior, and the purported influence of model texts as WCF on such processes, there are several limitations to our study which should be acknowledged.

Firstly, the reduced number of participants (N = 18) hinders the extrapolation of findings. However, such a reduced number of participants could be explained by the reluctance of educational institutions to grant permission to researchers to carry out their investigation at schools, especially when it involves an activity other than the ones included in the curriculum. Yet, there are two aspects that are worth pointing out: (1) our study did not intend to extrapolate significant results to the population as our main aim was to gain exploratory insight into the situational background of L2 writing processes in young EFL learners, and (2) most of the studies dealing with L2 writing processes have not made use of larger numbers of participants, with numbers ranging from 20 to 30 (Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2019; Roca de Larios et al., 2008). However, it seems convenient to mention that pausological studies, conducted under keystroke-logging environments, have allowed for larger sample sizes ranging from 60 to more than 100 participants (e.g. Barkaoui, 2016, 2019; Medimorec & Risko, 2017) given the easiness of automatizing the processing of the data gathering and further analyses.

Secondly, the use of stimulated recall was meant to be used as a supplementary research tool to tap into writing processes besides the Inputlog data output. However, we were not able to use the data gathered in Writing 1 precisely because of the poor richness of the data obtained. In relation to this, memory decay (Gass & Mackey, 2017) has been mentioned as one of the potential issues of stimulated recall procedures, since participants are very likely to tiptoe around certain events or thoughts during the process of writing. Thus, the poor data in this stimulated recall data from Writing 1 prompted us to decide not to carry out a second stimulated recall procedure for Writing 2. Hence, despite our great efforts at understanding the rationale and type of modifications of writing processes before and after the provision or absence of WCF, human errors in the data analysis might be present. This is an important limitation that shall be acknowledged in this thesis since our results concerning writing processes were mostly based upon inferences made when considering a combination of the different Inputlog output analyses, as clearly stated in Chapter 4 (Method). The absence of a stimulated recall procedure is a clear limitation that future research should address. Equally relevant, more research is needed as regards the best way to obtain rich data from children's stimulated recall procedures. Finally, obtaining data from two stimulated recalls pertaining to pre- and post-tests after receiving WCF might allow for gaining deeper insight into how these processes are not only applied but also managed by the participants themselves.

Thirdly, concerning our research design, the inclusion of stimulated recall for both groups might have prompted participants to pay attention to more specific aspects, as stimulated recall has been reported to trigger noticing and language awareness (Lindgren & Sullivan, 2003; Sabbaghan, 2013). Thus, adding another group without stimulated recall in Writing 1 would have certainly contributed to neutralizing the purported influence of the stimulated recall on the second version of the writing task.

Fourthly, another limitation also relates to the area of methodology, more specifically to data coding, for which we adopted an eclectic view with a novel, but congruent theoretically-grounded taxonomy. Most of the previous studies (e.g. Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2019; Roca de Larios et al., 2008) have consistently relied on models of writing (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Kellogg, 1996) to code their data. Thus, a potential limitation lies in the challenging nature of comparing results across studies, which might not be as consistent as if the same taxonomies had been used. Nonetheless, most of these taxonomies conflate a series of common aspects which represent the same concept beyond terminology and, as stated in Chapter 4 (Method), our study has attempted to encompass certain characteristics in these taxonomies. We have endeavoured to neutralize this potential issue by leaning on the reliability of our study supported by the very specific data that Inputlog provides in terms of pausing and revision behavior together with how the text was developed. However, one must not forget that the present empirical research is an exploratory study.

As to the direction of future research, there are several perspectives which shall be taken into consideration. Firstly, there is a lack of empirical research dealing with young children's writing processes since studies have mainly focused on adult L2 writing (e.g. Gánem-Gutiérrez

& Gilmore, 2018; Roca de Larios et al., 2008; Révész et al., 2017). Notwithstanding the contribution that this doctoral dissertation has attempted to provide to the field of young children's L2 writing processes, the strategic behavior of young learners when engaging in writing, needs to be further investigated with a larger pool of participants to add empirical evidence to this body of research and thus allowing for a comparison with adult L2 writing. Second, some models of writing such as Kellogg's (1996) have placed great emphasis on the relevance of working memory capacity (WMC) (Ahmadian, 2012). The inclusion of a working memory variable in our study went beyond the purposes of this thesis, but some very recent studies (e.g. Michel et al., 2019) have explored how young learners' WMC has a bearing on written performance. On this basis, future research should explore the importance of WMC in the different writing processes, and the way in which these are moulded or cognitively selected by the learner (see Cumming, 2016). Thirdly, the study of writing processes in young EFL learners needs further investigation in terms of how cognitive operations such as setting goals or the organization of the text occurs. Previous studies have explored the role of, for instance, planning and formulation in adult L2 writing by adopting very fine-grained taxonomies (e.g. Manchón & Roca de Larios, 2007; Roca de Larios et al., 2006; Lindgren & Sullivan, 2006; Xu, 2018), but there is an existing need to explore how processes such as monitoring are recursively utilized by children, and more specifically, in a digital setting. This way, the data from such studies can help to build or refine pre-existing taxonomies of adults for young learners.

Concerning models as WCF, and their effect on the distribution of writing processes, our PhD did not take a product-oriented perspective, but rather a process-oriented one since the main objective, as it has been mentioned throughout this empirical study, was to observe the purported effect of the provision of WCF on writing processes. Thus, our empirical research only took into consideration one type of feedback, that is, model texts, building on previous studies accounting for their validity (e.g. Coyle et al., 2019; García Mayo & Labandibar, 2017, among others). In light of this, further research should look into how other types of feedback might affect the different writing processes and other higher-level operations, as well as pausological behavior with the aid of KSL software such as Inputlog. Finally, as pointed by previous research (e.g. Révész et al., 2017), process-oriented studies should look further into how text quality may correlate with writing processes as well as fluency and pausing behaviors.

References

Abbuhl, R. (2011). Using models in writing instruction: A comparison with native and nonnative speakers of English. *SAGE Open*, *1*(3), 2158244011426295.

Abe, M. (2008). *Exploring the role of model essays in the IELTS writing test: A feedback tool*. (Unpublished MA thesis), University of Queensland, Australia.

Adams, A. M., Simmons, F., Willis, C., & Pawling, R. (2010). Undergraduate students' ability to revise text effectively: relationships with topic knowledge and working memory. *Journal of Research in Reading*, *33*(1), 54-76.

Adams, R. (2003). L2 output, reformulation and noticing: implications for IL development. *Language Teaching Research*, 7(3), 347-376.

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing: with commentaries by Ronald T. Kellogg & John R. Hayes* (Vol. 9). Springer Science & Business Media.

Alamargot, D., Dansac, C., Chesnet, D., & Fayol, M. (2007). Parallel processing before and after pauses: A combined analysis of graphomotor and eye movements during procedural text production. In M. Torrance, L. van Waes, & D. Galbraith (Eds.), *Writing and cognition: Research and applications* (pp. 13–29). Elsevier Science.

Aljaafreh, A.L.I., & Lantolf, J.P. (1994). Negative feedback as regulation and second language learning in the Zone of Proximal Development. *The Modern Language Journal*, 78(4), 465-483

Allaw, E. (2019). A learner corpus analysis: Effects of task complexity, task type, and L1 & L2 similarity on propositional and linguistic complexity. *International Review of Applied Linguistics in Language Teaching*. DOI:10.1515/iral-2018-0294.

Allwright, R. L., Woodley, M. P., & Allwright, J. M. (1988). Investigating reformulation as a practical strategy for the teaching of academic writing. *Applied linguistics*, *9*(3), 236-256.

Alvès, R. A., Castro, S. L., de Sousa, L., & Strömqvist, S. (2007). Influence of typing skill on pause-execution cycles in written composition. In M. Torrance, D. Galbraith & L. Van Waes (Eds.), *Recents developpements in writing-process research* (Vol. 20, pp. 55-65). Kluwer.

Amelohina, V., Nicolás-Conesa, F., & Manchón, R. M. (2020). Effects of task repetition with the aid of direct and indirect written corrective feedback. In R. M. Manchón (Ed.), *Writing and Language Learning: Advancing research agendas* (Vol. 56), (pp. 145-181). John Benjamins Publishing.

Ames C. (1992). Classrooms: goals, structures, and student motivation. *J Educ Psychol, 84*, 261–271.

Anderson, J. (1985). *Cognitive psychology and its implications* (2nd ed.). New York, NY: Freeman.

Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, *89*(4), 369-406.

Anderson, J.R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Andersson, B., Dahl, J., Holmkvist, K., Holsanova, J., Johansson, V., Karlsson, H., & Wengelin, A. (2006). Combining keystroke-logging with eye-tracking. In L. van Waes, M. Leijten & C. Neuwirth (Eds.), *Writing and digital media* (pp. 166-172). Elsevier.

Ashwell, T. (2000). Patterns of teacher response to student writing in a multiple-draft composition classroom: Is content feedback followed by form feedback the best method? *Journal of Second Language Writing*, 9(3), 227-258.

Baddeley, A. (2013*). Essentials of human memory (classic edition).* Psychology Press.

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

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Barkaoui, K. (2007). Revision in Second Language Writing: What Teachers Need to Know. *TESL Canada Journal*, 25(1), 81–92. https://doi.org/10.18806/tesl.v25i1.109

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. (2019). What Can L2 Writers' Pausing Behavior Tell Us About Their L2 Writing Processes? *Studies in Second Language Acquisition*, 41(3), 529–554. https://doi.org/10.1017/s027226311900010x

Barlett, E. J. Learning to revise: some component processes. In: M. Nystrand. (Ed.) *What writers know: the language, process, and structure of written discourse* (pp. 345-363). Orlando: Academic Press, Inc., 1982.

Batstone, R. (1996). Noticing. *ELT Journal*, 50 (3), 273. https://doi.org/10.1093/elt/50.3.273

Bawarshi, A. S., & Reiff, M. J. (2010). *Genre: An introduction to history, theory, research, and pedagogy*. Parlor Press.

Benesch, S. (2009). Theorizing and practicing critical English for academic purposes. *Journal of English for Academic Purposes,* 8(2), 81−85.

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

Bernstein, B. (1990). *Class, Codes and Control* (vol. 4). Routledge.

Bhatia, V. K. (2008). Genre analysis, ESP and professional practice. *English for Specific Purposes*, 27(2), 161−174.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers* (Vol. 23). John Benjamins.

Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, *17*(2), 102-118.

Bitchener, J. (2012). A reflection on 'the language learning potential'of written CF. *Journal of Second Language Writing*, *21*(4), 348-363.

Bitchener, J. (2019). The intersection between SLA and feedback research. In K. Hyland and F. Hyland (Eds), *Feedback in second language writing: Contexts and Issues* (2nd edition) (pp. 85-105). Cambridge University Press.

Bitchener, J., & Ferris, D., (2012) *Written Corrective Feedback in Second Language Acquisition and Writing*. Routledge.

Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language teaching research*, *12*(3), 409-431.

Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, *19*(4), 207-217.

Bitchener, J., & Storch, N. (2016). *Written Corrective Feedback for L2 Development* (Vol. 96). Multilingual Matters.

Bloom, B. S. (1954). The thought process of students in discussion. In S. J. French (Ed.), *Accent on teaching: Experiments in general education.* Harper.

Bosher, S. (1998). The composing processes of three Southeast Asian writers at the post-secondary level: An exploratory study. *Journal of Second Language Writing*, 7(2), 205–241.

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, 29(5), 591–620.

Bowles M.A. (2018). Introspective Verbal Reports: Think-Alouds and Stimulated Recall. In: Phakiti A., De Costa P., Plonsky L., Starfield S. (Eds), *The Palgrave Handbook of Applied Linguistics Research Methodology*. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_16

Breetvelt, I., Van den Bergh, H., & Rijlaarsdam, G. (1994). Relations between writing processes and text quality: when and how? *Cognition and Instruction*, 12(2), 103–123. https://doi.org/10.1207/s1532690xci1202_2

Bridwell, L. S. (1980). Revising strategies in twelfth grade students' transactional writing. *Research in the Teaching of English*, 14(3), 197–222.

Bridwell, L., Sirc, G., & Brooke, R. (1985). Revising and computing: Case studies of student writers. In S.W. Freedman (Ed.), *The acquisition of written language: Response and revision* (pp. 172-194). Ablex.

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 University Press.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language Pedagogy*. Longman, Inc.

Bruton, A. (2009). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing*, 18(2), 136-140.

Cánovas, J., Roca de Larios, J. and Coyle, Y. (2015). The use of models as a written feedback technique with young EFL learners. *System*, *52(1*), 63-77.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83-1 15). Newbury.

Chan, S. H. C., Wu, R. Y. F., & Weir, C. J. (2014). Examining the context and cognitive validity of the GEPT Advanced Writing Task 1: A comparison with real-life academic writing tasks. *LTTC-GEPT Research Reports*, RG-03, 1–89.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*(3), 267-296.

Chanquoy, L. (2001). How to make it easier for children to revise their writing: A study of text revision from 3rd to 5th grades. *British Journal of educational psychology*, *71*(1), 15-41.

Chanquoy, L. (2009) Revisions Processes. In R. Beard, D. Myhill, J. Riley & M. Nystrand (Eds), *The Sage Handbook of Writing Development* (p. 80-97). Sage.

Cheng, Y.-S. (2004). EFL students' writing anxiety: Sources and Implications. *English Teaching & Learning*, 29(2), 41−62.

Chukharev-Hudilainen, E. (2019). Empowering automated writing evaluation with keystroke logging. In E. Lindgren and Sullivan (Eds.) (pp. 135-142), *Observing Writing*. Brill. doi:10.1163/9789004392526_007

Chukharev-Hudilainen, E., Feng, H.-H., Saricaoglu, A., & Torrance, M. (2019). Combined deployable keystroke logging and eyetracking for investigating cognitive processes that underlie L2 writing. *Studies in Second Language Acquisition*, 41(3), 583–604.

Cohen, A. D., & Dörnyei, Z. (2002). Focus on the language learner: Motivation, styles, and strategies. In N. Schmitt (Ed.), *An introduction to applied linguistics* (Vol. 14) (pp. 170-190). Arnold.

Conijn, R., Roeser, J., & Van Zaanen, M. (2019). Understanding the keystroke log: the effect of writing task on keystroke features. *Reading and Writing*, *32*(9), 2353-2374.

Coyle, Y., & Guirao, J. C. (2019). Learning to write in a second language: The role of guided interaction in promoting children's noticing from model texts. CLIL. *Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 2(1), 21-30.

Coyle, Y., & Roca de Larios, J. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task. *Studies in Second Language Acquisition, 36(03),* 451-485.

Coyle, Y., & Roca de Larios, J. (2020). Exploring young learners' engagement with models as a written corrective technique in EFL and CLIL settings. *System*, 95, 102374. https://doi.org/10.1016/j.system.2020.102374

Coyle, Y., Cánovas Guirao, J., & Roca de Larios, J. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, 42(March), 25–43. https://doi.org/10.1016/j.jslw.2018.09.002

Coyle, Y., Guirao, J. C., & de Larios, J. R. (2018). Identifying the trajectories of young EFL learners across multi-stage writing and feedback processing tasks with model texts. *Journal of Second Language Writing*, *42*, 25-43.

Cumming, A. (1990). Metalinguistic and Ideational Thinking in Second Language Composing. *Written Communication*, *7*(4), 482-511.

Cumming, A. (2013). Multiple dimensions of academic language and literacy development. *A Journal of Research in Language Studies*, 63(s1), 130-152.

Cummins, J. (1989). A theoretical framework for bilingual special education. *Exceptional children*, *56*(2), 111-119.

Daiute, C. (1986). Physical and cognitive factors in revising: Insights from studies with computers. *Research in the Teaching of English*, 20, 141–159.

Dam-Jensen, H. & Heine, C. (2013). Writing and translation process research: Bridging the gap. *Journal of Writing Research*, 5(1), 89-101.

de Smet, M. J., Leijten, M., & Van Waes, L. (2018). Exploring the process of reading during writing using eye tracking and keystroke logging. *Written Communication*, *35*(4), 411-447.

Deane, P., & Zhang, M. (2015). Exploring the feasibility of using writing process features to assess text production skills. *ETS Research Report Series*, *2015*(2), 1-16.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, *19*(2), 109-134.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, *11*(4), 227-268.

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19, 167-189.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125–151). Cambridge University Press.

DeKeyser, R. (2003). Explicit and Implicit Learning. In C. Doughty, & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 313-348). Blackwell.

DeKeyser, R. (2007). Introduction: Situating the context of practice. In R. DeKeyser (Ed.). *Practicing in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1-18). Cambridge University Press.

DeKeyser, R. (2015). Skill Acquisition Theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (2nd ed., pp. 94-112). Routledge.

DeKeyser, R. M. (2014). Research on language development during study abroad. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 313-325). John Benjamins Publishing Company.

DeKeyser, R. M., & Botana, G. P. (Eds.). (2019). *Doing SLA research with implications for the classroom: Reconciling methodological demands and pedagogical applicability* (Vol. 52). John Benjamins Publishing Company.

Dörnyei, Z. (2010). Researching motivation: From integrativeness to the ideal L2 self. *Introducing applied linguistics: Concepts and skills*, 3(5), 74-83.

Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. Cambridge University Press.

Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 187-261). Cambridge University Press.

Dragsted, B., & Carl, M. (2013). Towards a classification of translation styles based on eye-tracking and keylogging data. *Journal of Writing Research*, *5*(1), 133-158.

Dulay, H. & Burt, M. (1973). Should we teach children syntax? *Language Learning*, *24*, 245–258.

Dulay, H. S., & Burt, M. K. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24(1), 37-53.

Dulay, H., & Burt, M. (1977). Remarks on creativity in language acquisition. *Viewpoints on English as a second language*, *2*, 95-126.

Dweck, C. S. (2000). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Edelsky, C. (1986). *Writing in a bilingual program: Habia una vez*. Norwood, NJ: Able

Egi, T. (2007). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A series of empirical studies* (pp. 249–267). Oxford: Oxford University Press.

Egi, T. (2008). Investigating stimulated recall as a cognitive measure: Reactivity and verbal reports in SLA research methodology. *Language Awareness*, 17(3), 212–228.

Ehrman, M. E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *The modern language journal*, 79(1), 67-89.

Ellis, N. C., Simpson-Vlach, R. I. T. A., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *Tesol Quarterly*, *42*(3), 375-396.

Ellis, R. (1991). The interaction hypothesis: a critical evaluation. In Sadtono, E., (Ed.), *Language acquisition and the second/foreign language class- room*. (pp. 179-211). SEAMEO Regional Language Centre**.**

Ellis, R. (1995). Interpretation tasks for grammar teaching. *Tesol Quarterly*, 29(1), 87-105.

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

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal*, 1(1), 3-18.

Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, *32*(2), 335–349.

Ellis, R. (2020). Task-based language teaching for beginner-level young learners. *Language Teaching for Young Learners*, 2(1), 4-27.

Ellis, R. Sheen, Y. Murakami, M. & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36, 353-371.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in second Language acquisition*, *26*(1), 59-84.

Engeström, Y., (2001). Expansive Learning at Work: Toward an Activity Theoretical Conceptualization. *Journal of Education and Work*, 14 (1): 133-156

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (p. 102–134). Cambridge University Press. https://doi.org/10.1017/CBO9781139174909.007

Eschholz, P. A. (1980). The prose models approach: Using products in the process. In T. R. Donovan & B. W. McClelland (Eds.), *Eight approaches to teaching composition* (pp. 21-35). National Council of Teachers of English.

Esteban, N. M., & de Larios, J. R. (2010). The use of models as a form of written feedback to secondary school pupils of English. *International Journal of English Studies*, *10*(2), 143-170.

Eveyik, E. (1999). *Development of an attitude scale to investigate Turkish EFL teachers' attitudes towards communicative language teaching*. Unpublished Master's Thesis. Bogazici University. https://doi. org/10.1046/j, 1469-1809.

Fagan, W. T., & Hayden, H. M. (1988). Writing pro- cesses ofFrench and English offifth grade French immersion students. *Canadian Modern Language Review*, 44, 653–668.

Faigley, L., & Witte, S. (1981). Analyzing revision. *College Composition and Communication*, 32(4), 400–414. https://doi.org/10.2307/356602.

Fathman, A. (1975) The Relationship between Age and Second Language Productive ability. *Language Learning, 25 (2)*, 245-253.

Fathman, A., & Whalley, E. (1990). Teacher response to student writing: Focus on form versus content. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (pp. 178-190). Cambridge University Press.

Fazio, L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority- and majoritylanguage students. *Journal of Second Language Writing*, 10, 235–249.

Ferris, D. (1995). Student reactions to teacher response in multiple-draft composition classrooms. *TESOL quarterly*, *29*(1), 33-53.

Ferris, D. (2004). The "grammar correction" debate in L2 writing: Where are we and where do we go from here? (and what do we do in the meantime…?). *Journal of Second Language Writing*, 13, 49-62.

Ferris, D. (2010). Second language writing research and written corrective feedback in SLA. Intersections and practical applications. *Studies in Second Language Acquisition*, 32, 181-201.

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *Tesol Quarterly*, *31*(2), 315-339.

Ferris, D. R. (1999). The case for grammar correction in L2 writing classes. A response to Truscott (1996). *Journal of Second Language Writing*, 8, 1–10.

Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81-104). Cambridge University Press.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be?. *Journal of Second Language Writing*, *10*(3), 161-184.

Fitzgerald, J. (1987). Research on revision in writing. *Review of educational research*, *57*(4), 481-506.

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication*, 31(1), 21–32.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second language acquisition*, *18*(3), 299-323.

Galbraith, D. (2009). Cognitive models of writing. *German as a foreign language*, (2-3), 7-22.

Gánem-Gutiérrez, G. A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning*, *68*(2), 469-506.

García Mayo, M. D. P., & Labandibar, U. L. (2017). The use of models as written corrective feedback in English as a foreign language (EFL) writing. *Annual Review of Applied Linguistics*, *37*(2017), 110-127.

Gardner, H. (1985). *Frames of mind: The theory of multiple intelligences*. Hachette UK.

Gardner, H. (2006). *Changing minds: The art and science of changing our own and other peoples minds*. Harvard Business Review Press.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and Motivation in Second-Language Learning*. Newbury House.

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

Gass, S. M. & Mackey, A. (2000). *Stimulated-Recall Methodology in Second Language Research*. New Jersey (LEA).

Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Psychology Press.

Gass, S. M., & Mackey, A. (2016). *Stimulated recall methodology in applied linguistics and L2 research*. Taylor & Francis.

Gathercole, S. E., Lamont, E. & Alloway, T. P. (2006). Working memory in the classroom. In S. J. Pickering (ed.), *Working memory and education* (pp. 219-240). Academic Press.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and speech*, *15*(2), 103-113.

Goldstein, L. (2006). Feedback and revision in second language writing: Contextual, teacher, and student variables. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 185-205). Cambridge University Press.

Goldstein, L. M., & Conrad, S. M. (1990). Student input and negotiation of meaning in ESL writing conferences. *Tesol Quarterly*, *24*(3), 443-460.

Grabowski, J. (2008). The internal structure of university students' keyboard skills. *Journal of Writing Research*, 1(1), 27-52.

Grombczewska, M. (2011). The relationship between teacher's feedback and students' motivation. *Humanising Language Teaching*, 3(3), 27-38.

Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24(4), 443-461.

Guenette, D. (2007). Is feedback pedagogically correct?: Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, *16*(1), 40-53.

Guirao, J. C. (2017). *The use of written models in the teaching of english in primary./la utilización de modelos escritos en la enseñanza del inglés en primaria* (Doctoral dissertation, University of Murcia).

Hamidun, N., Hashim, S. H. M., & Othman, N. F. (2012). Enhancing Students' Motivation by Providing Feedback on Writing: The Case of International Students from Thailand. *International Journal of Social Science and Humanity*, *2*(6), 591.

Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom. *Journal of Second Language Writing*, *30*, 31-44.

Hanaoka, O. (2006). Exploring the Role of Models in Promoting Noticing in L2 Writing. *JACET Bulletin, 42*, 1–13.

Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 459–479.

Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, 21(4), 332–347.

Harklau, L. (2002). The role of writing in classroom second language acquisition. *Journal of Second Language Writing*, 11(4), 329-350.

Harwood, N. (2009). An interview-based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics*, 41(3), 497–518.

Hasan, R. (1996). Literacy, everyday talk and society. In Ruqaiya Hasan and Geoffrey Williams (Eds.), *Literacy in society*, 377−424. Longman.

Hayes, J. R. (2004). What triggers revision? In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (Vol. 13) (pp. 9–20). Kluwer.

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

Hayes, J. R., & Flower, L. S. (1980). Identifying the organisation of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Lawrence Erbaum Associates.

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 29–55). Erlbaum.

Henderson, L., & Tallman, J. (2006). *Stimulated recall and mental models*. Scarecrow Press, Inc.

Hirvela, A., Hyland, K., & Manchón, R. M. (2016). Dimensions in L2 writing theory and research: Learning to write and writing to learn. In R. M. Manchón & P. K. Matsuda, (Eds.), *Handbook of second and foreign language writing* (pp. 45-64). De Gruyter Mouton. https://doi.org/10.1515/9781614511335-005

Ho, M., & Savignon, S.J. (2007). Face-to-face and Computer-mediated Peer Review in EFL Writing. *CALICO Journal*, *24*(2), 269-290.

Hoang, H. (2019). Metaphorical Language in Second Language Learners' Texts: Additional Baggage of the Writing Journey?. In E. Lindgren & K. Sullivan, (Eds.), *Observing Writing* (pp. 236-257). Brill.

Holmes, V. M. (1984). Sentence planning in a story continuation task. *Language and Speech*, 27(2), 115–134.

Holmes, V. M. (1988). Hesitations and sentences planning. *Language and Cognitive Processes*, 3(4), 323–361.

Hulstijn, J. H. (1995). Not all grammar rules are equal: Giving grammar instruction its proper place in foreign language teaching. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (Technical report No. 9) (pp. 359-386). University of Hawaii, Second Language Teaching and Curriculum Center.

Hyland, F. (2011). The language learning potential of form-focused feedback on writing: Students‟ and teachers‟ perceptions. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 159-180). John Benjamins.

Hyland, K. & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39 (2)*, 83-101.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of pragmatics*, *30*(4), 437-455.

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing*, *12*(1), 17-29.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16(3), 148-164.

Hyland, K. (2011). Learning to write: Issues in theory, research and pedagogy. In R. M. Manchón, (Ed.), *Learning-to-write and writing-to-learn in an additional language*, (vol. 35) (pp. 17-35). John Benjamins.

Hyland, K. (2016). Methods and methodologies in second language writing research. *System*, 59, 116–125. https://doi.org/10.1016/j.system.2016.05.002

Hyland, K., & Hyland, F. (Eds.). (2019). *Feedback in second language writing: Contexts and issues*. Cambridge University Press.

Hyönä, J., & Lorch, R. F. (2004). Effects of topic headings on text processing: Evidence from adult readers' eye fixation patterns. *Learning and instruction*, *14*(2), 131-152.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in second language acquisition*, 24(4), 541-577.

Izumi, S. (2003). Processing difficulty in comprehension and production of relative clauses by learners of English as a second language. *Language Learning*, 53(2), 285-323.

Izumi, S. (2013). Noticing and L2 development: Theoretical, empirical, and pedagogical issues. In J. M. Bergsleithner, S. N. Frota, & J. K . Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 25-38). Honolulu: University of Hawai'I, National Foreign Language Resource Center.

Izumi, S. (2014). Noticing and L2 development: Theoretical, empirical, and pedagogical issues. In J. M. Bergsleithner, S. N. Frot & J. K. Yoshioka (Eds.), *Noticing and second language acquisition:*

Studies in honor of Richard Schmidt (pp. 37-50). National Foreign Language Resource Center, University of Hawai'i at Manoa.

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in second language acquisition*, 21(3), 421-452.

Jakobsen, A. L. (2011). Tracking translators' keystrokes and eye movements with Translog. In C. Alvstad, A. Hild, & E. Tiselius (Eds.), *Methods and strategies of process research* (pp. 37-55). John Benjamins.

Johns, A. (1995). Teaching classroom and authentic genres: Initiating students into academic cultures and discourses. In D. Belcher, & G. Braine (Eds.), *Academic writing in a second language* (pp. 277–291). Ablex Publishing.

Johnson, K. (1996). *Language teaching and skill learning*. Basil Blackwell.

Johnson, K. E. (1996). The role of theory in L2 teacher education. *TESOL quarterly*, *30*(4), 765-771.

Johnson, M. D. (2020). Planning in L1 and L2 writing: Working memory, process, and product. *Language Teaching*, *53*(4), 433-445.

Jones, C. S. & Tetroe, J. (1987). Composing in a second language. In A. Matsuhashi (Ed.), *Writing in real time: Modeling production processes* (pp. 34-57). Ablex.

Kalaja, P., & Barcelos, A. M. F. (2003). Conclusion: Exploring possibilities for future research on beliefs about SLA. In P. Kalaja & A. M. Ferreira Barcelos, (Eds.), *Beliefs about SLA: New Research Approaches* (pp. 231-238). Springer.

Kang, E. Y. (2020). Using model texts as a form of feedback in L2 writing. *System*, *89*(June 2021), 102-196.

Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99(1), 1-18.

Karnan, M., Akila, M., & Krishnaraj, N. (2011). Biometric personal authentication using keystroke dynamics: A review. *Applied soft computing*, *11*(2), 1565-1573.

Kehagia, O., & Cox, M. (1997). Revision changes when using wordprocessors in an English as a foreign language context. *Computer Assisted Language Learning*, *10*(3), 239-253.

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

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1), 1-26.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research*, 5, 159–190.

Kennedy, G. D. (1987). Expressing temporal frequency in academic English. *TESOL Quarterly*, *21*(1), 69-86.

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

Khuder, B. (2015). L2 Writing in Test and Non-test Situations : Process and Product. *Journal of Writing Research*, 6(3), 233–278.

Khuder, B., & Harwood, N. (2015). L2 writing in test and non-test situations: Process and product. *Journal of Writing Research*, 6(3), 233–278. https://doi.org/10.17239/jowr- 2015.06.03.2

Kim, H. (2020). Profiles of undergraduate student writers: Differences in writing strategy and impacts on text quality. *Learning and Individual Differences*, *78*(2020), 101-823.

Kim, J. H. (2015). The role of models and error correction in L2 young learners' noticing and Output. *ENGLISH TEACHING (영어교육)*, *70*(2), 3-26.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, *21*(4), 390-403.

Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21(4), 390–403. https://doi.org/10.1016/j.jslw.2012.09.003

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Krashen, S.D. (1985). *The input hypothesis: Issues and implications*. Longman.

Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking. In P. Robinson, (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (vol. 2), (pp. 91-104).

Lalande, J. F. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, *66*(2), 140-149.

Lantolf, J. P. (2005). Sociocultural and second language learning research: An exegesis. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 359-378). Routledge.

Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.

Lantolf, P., & Thorne, S.L. (2007). Sociocultural theory and second language learning. In B.Van Patten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 693-701). Lawrence Erlbaum Associates, Inc.

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, *65*(S1), 127-159.

Latif, M. A. (2008) A state-of-the-art review of the real-time computer-aided study of the writing process. *IJES, International journal of English studies*, 8(1), 29-50.

Latif, M. A. (2019). Using think-aloud protocols and interviews in investigating writers' composing processes: combining concurrent and retrospective data. *International Journal of Research & Method in Education*, *42*(2), 111-123.

Lee, I. (1997). ESL learners' performance in error correction in writing: Some implications for teaching. *System*, *25*(4), 465-477.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, 13, 285–312.

Lee, S. M. (2014). The relationships between higher order thinking skills, cognitive density, and social presence in online learning. *The Internet and Higher Education*, 21(2014), 41–52. https://doi.org/10.1016/j.iheduc.2013.12.002

Lee, S. Y., & Krashen, S. (2002). Predictors of success in writing in English as a foreign language: Reading, revision behavior, apprehension, and writing. *College Student Journal*, *36*(4), 532-543.

Lee, S. Y., & Krashen, S. D. (1997). Writing apprehension in Chinese as a first language. *ITL-International Journal of Applied Linguistics*, *115*(1), 27-37.

Leijten, M., & Van Waes, L. (2005). *Inputlog: A logging tool for the research of writing processes*. Research Paper, University of Antwerp.

Leijten, M., & Van Waes, L. (2006). Inputlog: New perspectives on the logging of on-line writing processes in a Windows environment. In Sullivan, K. P.H, & Lindgren, E. (Eds). *Computer keystroke logging and writing: Methods and applications* (pp. 73-95). Amsterdam, Netherlands: Elsevier.

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. https://doi.org/10.1177/0741088313491692

Leijten, M., Macken, L., Hoste, V., Van Horenbeeck, E., & Van Waes, L. (2012, April). From character to word level: Enabling the linguistic analyses of Inputlog process data. In *Proceedings of the Second Workshop on Computational Linguistics and Writing (CL&W 2012): Linguistic and Cognitive Aspects of Document Creation and Document Engineering* (pp. 1-8).

Leijten, M. & Van Waes, L. (2015-2019). *Inputlog. Help documentation (in progress).* Self-publishing. University of Antwerp.

Leki, I., Cumming, A., & Silva, T. (2008). *A synthesis of research on second language writing in English*. Routledge.

Lens, W. (1994). Personal computers in the learning environment and student motivation. *Scandinavian Journal of Educational Research*, 38, 219-230.

Leow, R. & Suh, B. R. (2021). Theoretical perspectives on L2 writing, corrective feedback, and language learning in individual writing conditions. In R. Manchón & B. R. Suh (Eds.), *The Routledge Handbook of Second Language Acquisition and Writing* (pp. n/d). Routledge.

Leow, R. (2018). Noticing hypothesis. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching (1st Edition)*. John Wiley & Sons.

Leow, R. P. (2015). *Explicit Learning in the L2 Classroom: A Student-Centered Approach*. London and Routledge.

Leow, R. P. (2015). *Explicit learning in the L2 classroom: A student-centered approach*. Routledge.

Leow, R. P. (2020). L2 writing-to-learn. In R. M. Manchón (Ed.), *Writing and Language Learning: Advancing research agendas* (vol. *56*), 95-118.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation.* MIT Press.

Levy, M. C., Marek, P. J., & Lea, J. (1996). Concurrent and retrospective protocols in writing research. In G. Rijlaarsdam, H. van den Bergh, & M. Conzijn (Eds), *Theories, models and methodology in writing research* (pp. 542–556). Amsterdam University Press.

Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning, 31*(8), 882-904.

Lightbown, P. M., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in second language acquisition*, 12(4), 429-448.

Lindgren, E., & Sullivan, K. (Eds.). (2019). *Observing writing: Insights from keystroke logging and handwriting*. Brill.

Lindgren, E., & Sullivan, K. P. H. (2006a). Writing and the analysis of revision: An overview. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing*: *Methods and applications* (pp. 31-44). Elsevier.

Lindgren, E., & Sullivan, K. P. H. (2006b). Analysing on-line revision. In G. Rijlaarsdam (Series Ed.) & K.P.H. Sullivan & E. Lindgren (Vol. Eds.), *Studies in writing, computer keystroke-logging and writing: Methods and applications* (pp. 157–188). Elsevier.

Lo, J., & Hyland, F. (2007). Enhancing students' engagement and motivation in writing: The case of primary students in Hong Kong. *Journal of Second Language Writing*, 16(4), 219 – 237.

Loewen, S., & Plonsky, L. (2015). *An A–Z of applied linguistics research methods*. Macmillan International Higher Education.

Long, M. H. (1981). Questions in foreigner talk discourse. *Language learning*, 31(1), 135-157.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation ofcomprehensible input. *Applied Linguistics*, 4(2), 126-141.

Long, M. H. (2009). Methodological principles for language teaching. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 373-394). Wiley Blackwell.

Long, M. H.（2000). *Focus on form in Task-Based Language Teaching*. In Lambert,

Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. Cambridge University Press.

López, P., Torrance, M., & Fidalgo, R. (2019). The online management of writing processes and their contribution to text quality in upper-primary students. *Psicothema: Revista de Psicologia*, *31*(3), 311-318.

Luquin, M., & García Mayo, M. del P. (2021). Exploring the use of models as a written corrective feedback technique among EFL children. System, 98(2021), 102465. https://doi.org/10.1016/j.system.2021.102465

Lyster, R., & Sato, M. (2013). Skill acquisition theory and the role of practice in L2 development. In M. P. García de Mayo, M. J. Gutiérrez Mangado & M. Martínez Adrián (Eds.), *Contemporary approaches to second language acquisition* (pp. 71-92).

MacArthur, C. A. & Graham, S. (2016). Writing Research from a Cognitive Perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (2nd ed.) (pp. 11-23). The Guilford Press.

Makino, T. (1980). *Acquisition order of English morphemes by Japanese adolescents*. Shinozaki Shorin Press.

Manchón, R. M. (2009a). Broadening the perspective of L2 writing scholarship: The contribution of research on foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 1–19). Multilingual Matters.

Manchón, R. M. (2009b). Individual differences in foreign language learning: the dynamics of beliefs about l2 writing. *Revista Española De Lingüística Aplicada (RESLA (Revista Española De Linguistica Aplicada)*, *22*, 245-268.

Manchón, R. M. (2011a). Situating the learning-to-write and writing-to-learn dimensions of L2 writing. In R. M. Manchón (Ed.), *Learning to write and writing to learn in an additional language* (pp. 61 – 82). Benjamins.

Manchón, R. M. (2011b). The language learning potential of writing in foreign language contexts: Lessons from research. In M. Reichelt & T. Chimasko (Eds.), *Foreign language writing: Research insights* (pp. 44 – 64). Parlor Press.

Manchón, R. M., & De Larios, J. R. (2007). On the temporal nature of planning in L1 and L2 composing. *Language Learning*, 57(4), 549–593. https://doi.org/10.1111/j.1467-9922.2007.00428.x

Manchón, R. M., & Leow, R. P. (2020). An ISLA perspective on L2 learning through writing. In R. M. Manchón (Ed.), *Writing and Language Learning: Advancing research agendas* (vol. 56) (pp. 335-356).

Manchón, R. M., & Matsuda, P. K. (Eds.). (2016). *Handbook of second and foreign language writing (Vol. 11)*. Walter de Gruyter GmbH & Co KG.

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspec- tives and empirical evidence. In J. W. Schwieter & A. Benati (Eds.), *The Cambridge handbook of language learning* (pp. 341–362). Cambridge University Press. https://doi.org/10.1017/9781108333603.015

Manchón, R. M., & Williams, J. (2016). L2 writing and SLA studies. In R.M. Manchón (Ed.), *Handbook of second and foreign language writing* (pp. 567-586). De Gruyter Mouton.

Manchon, R. M., Murphy, L., & de Larios, J. R. (2007). Lexical retrieval processes and strategies in second language writing: A synthesis of empirical research. *International Journal of English Studies*, 7(2), 149-174.

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 *Writing in foreign language contexts* (pp. 102-129). Multilingual Matters.

Manchón, R.M. & Polio, C. (Eds.). (2021). *The Routledge Handbook of Second Language Acquisition and Writing*. Routledge.

Manchón, R.M. & Roca de Larios, J. (Eds.). (forthcoming/2021). *Research methods in the study of L2 writing processes*. John Benjamins.

Marsden, E., & Torgerson, C. J. (2012). Single group, pre-and post-test research designs: Some methodological concerns. *Oxford Review of Education*, *38*(5), 583-616.

Martínez, N., & Roca de Larios, J. R. (2010). The Use of Models as a Form of Written Feedback to Secondary School Pupils of English. *International Journal of English Studies, 10(2),* 143-170.

Matsuhashi, A. (1981). Pausing and planning: The tempo of written discourse production. *Research in the Teaching of English*, 113-134.

McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educational Psychology Review*, 8(3), 299e325.

McLaughlin, B. (1978). The Monitor Model: Some methodological considerations. *Language Learning*, 28(2), 309–332.

McLaughlin, B. (1980). Theory and research in second language learning: An emerging paradigm. *Language Learning*, *30*, 331–350.

McLaughlin, B. (1987). Reading in a Second Language: Studies with Adult and Child Learners. In S. i. Goldman, H. T. Trueba (Eds.), *Becoming Literate in English as a Second Language* (pp. 57-70). Ablex.

McLaughlin, B. (1987). *Theories of second language learning.* Edward Arnold.

McLaughlin, B. (1990). Restructuring. *Applied linguistics*, *11*(2), 113-128.

McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 213-228).

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information processing perspective. *Language Learning*, 33(2), 135-158. https://doi.org/10.1111/j.1467-1770.1983.tb00532.x

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective, *Language learning*, 33(2), 135-158.

Medimorec, S., & Risko, E. F. (2017). Pauses in written composition: On the importance of where writers pause. *Reading and Writing*, *30*(6), 1267-1285.

Meisel, J. M., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in second language acquisition*, *3*(2), 109-135.

Michel, M., Kormos, J., Brunfaut, T., & Ratajczak, M. (2019). The role of working memory in young second language learners' written performances. *Journal of Second Language Writing*, 45, 31-45.

Michel, M., Révész, A., Lu, X., Kourtali, N. E., Lee, M., & Borges, L. (2020). Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research*, *36*(3), 307-334.

Michel, M., Stiefenhöher, L., Verspoor, M. & Manchón, R. M. (2021). L2 writing processes and language learning in individual and collaborative conditions. In R. Manchón & B. R. Suh (Eds.), *The Routledge Handbook of Second Language Acquisition and Writing* (pp. n/d). Routledge.

Miller, K. S. (2005). Second language writing research and pedagogy: A role for computer logging? *Computers and Composition*, 22(3 Spec. Issue.), 297–317. https://doi.org/10.1016/j.compcom.2005.05.003

Myles, J. (2002). Second language writing and research: The writing process and error analysis in student texts. *Tesl-Ej*, *6*(2), 1-20.

Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59(2), 411-452.

Nassaji, H., & Sato, K. (2019). The effects of feedback explicitness and type of target structure on accuracy in revision and new pieces of writing. *System*, 81, 135–145. https://doi.org/10.1016/j.system.2018.12.017

Neill, J. T. (2008). *Enhancing life effectiveness: The impacts of outdoor education programs*. Unpublished doctoral dissertation, Faculty of Education, University of Western Sydney, NSW, Australia.

Nicolás–Conesa, F., Manchón, R.M., & Cerezo, L. (2019). The effect of unfocused direct and indirect written corrective feedback on rewritten texts and new texts: Looking into feedback for accuracy and feedback for acquisition. *Modern Language Journal*, 103(4), 848-873.

Nicolás-Conesa, F., Roca de Larios, J., & Coyle, Y. (2014). Development of EFL students' mental models of writing and their effects on performance. *Journal of Second Language Writing*, 24(1), 1–19. https://doi.org/10.1016/j.jslw.2014.02.004

Olive, T., & Cislaru, G. (2015). Linguistic forms at the process-product interface: Analyzing the linguistic content of bursts of production. In G. Cislaru (Ed.), *Writing(s) at the crossroads: The process-product interface* (pp. 99–123). John Benjamins.

O'Malley. J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge University Press.

Ong, J. (2014). How do Planning Time and Task Conditions Affect Metacognitive Processes of L2 Writers? *Journal of Second Language Writing*, 23(1), 17–30. https://doi.org/10.1016/j.jslw.2013.10.002

Ong, J., & Zhang, L. J. (2013). Effects of manipulation of cognitive processes on English-as- a-foreign-language (EFL) writers' text quality. *TESOL Quarterly*, 47(2), 375–398.

Ortega, L. (2009). Studying writing across EFL contexts: Looking back and moving forward. In R. M. Manchón (Ed.), *Writing in foreign language contexts. Learning, teaching, and research* (pp. 232 – 255). Multilingual Matters.

Ortega, L. (2012). Epilogue: Exploring L2 writing–SLA interfaces. *Journal of Second Language Writing*, 21(4), 404–415. https://doi.org/10.1016/j.jslw.2012.09.002

Pajares, F., & Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary educational psychology*, *26*(3), 366-381.

Palviainen, Å., Kalaja, P., & Mäntylä, K. (2012). Development of L2 writing: Fluency and proficiency. *AFinLA-e: Soveltavan kielitieteen tutkimuksia*, (4), 47-59.

Park, E. S. Song, S., & Shin, Y. K. (2016). To what extent do learners benefit from indirect written corrective feedback? A study targeting learners of different proficiency and heritage language status. *Language Teaching Research*, 20(6), 678-699.

Paulston, C. (1972). Teaching writing in the ESOL classroom: Tech- niques of controlled composition. *TESOL Quarterly*, 6, 33-59.

Pellicer Sánchez, A., & Siyanova-Chanturia, A. (2018) Eye movements in vocabulary research: Current findings and future directions. *ITL - International Journal of Applied Linguistics*, *169*(1), 5–29.

Pennington, M. C. & So, S. (1993). Comparing writing process and product across two languages: A study of 6 Singaporean student writers. *Journal of Second Language Writing*, 2 (l), 41-63.

Phinney, M., & Khouri, S. (1993). Computers, revision and ESL writers: The role of experience. *Journal of Second Language Writing*, 2, 257–277.

Pienemann, M. & Lenzing, A. (2015). Processability Theory. In B. VanPatten & J. Williams (Eds.), *Theories in Second Language Acquisition* (pp. 159-179). Routledge.

Pienemann, M. (1981). *Der Zweitspracherwerb ausländischer Arbeiterkinder* (Vol. 4). Bouvier.

Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienneman (Eds.), *Modelling and assessing second language acquisition* (pp. 23-75). John Benjamins.

Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied linguistics*, *10*(1), 52-79.

Pienemann, M. (1998). *Language processing and second language development: Processability theory* (Vol. 15). John Benjamins Publishing.

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

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during Computerized Dynamic Assessment (C-DA). *Language Teaching Research*, 17(3), 323-342.

Polio, C. (2012). The relevance of second language acquisition theory to written error correction debate. *Journal of Second Language Writing*, 21, 375-389.

Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, *21*(4), 375-389.

Polio, C., Fleck, C., & Leder, N. (1998). "If I only had more time:" ESL learners‟ changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7, 43-68.

Porte, G. (1996). When writing fails: How academic context and past leaming experiences chape revision. *System*, 24 (l), 107-1 16.

Prior, P. A. (1995). Redefining the task: An ethnographic examination of writing and response in graduate seminars. In *Academic writing in a second language: Essays on research and pedagogy* (pp. 47-82). Ablex.

Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, *10*(4), 277-303.

Ranalli, J., Feng, H.-H., & Chukharev-Hudilainen, E. (2018). The affordances of process-tracing technologies for supporting L2 writing instruction. *Language Learning & Technology,* 23(2), 1–11.

Rankin-Brown, M. S. (2006). Addressing Writing Apprehension in Adult English Language Learners, *Proceedings of CATESOL State Conference*, 2006.

Ransdell, S. E., & Levy, M. (1998). Writing in a second language and the limits of working memory. *Paper presented at the 1998 European Writing Conference of the EARLI Special Group on Writing*, Poitiers, France.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7(1), 65–81. https://doi.org/10.1016/0010-0285(75)90005-5

Reinders, H., Erlam, R., Philp, J., Loewen, S., Elder, C., Ellis, R., & Philp, J. (Eds.) (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. Multilingual Matters.

Révész, A. J. (2007). *Focus on form in task-based language teaching: Recasts, task complexity, and L2 learning*. Doctoral dissertation, Columbia: Columbia University.

Révész, A., & Michel, M. (2019). Methodological Advances in Investigating L2 Writing Processes: Introduction. *Studies in Second Language Acquisition*, *41*(3), 491-501.

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., Kourtali, N. E., & Mazgutova, D. (2017). Effects of Task Complexity on L2 Writing Behaviors and Linguistic Complexity. *Language Learning*, 67(1), 208–241. https://doi.org/10.1111/lang.12205

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors: a mixed-methods study. *Studies in Second Language Acquisition*, *41*(3), 605-631.

Révész, A., Michel, M., & Lee, M. (2019). Exploring Second Language Writers' Pausing and Revision Behaviors. Studies in Second Language Acquisition, 41(3), 605–631. https://doi.org/10.1017/s027226311900024x

Reynolds, T. H., & Bonk, C. J. (1996). Facilitating college writers' revisions within a generative-evaluative computerized prompting framework. *Computers and Composition*, *13*(1), 93-108.

Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL quarterly*, *20*(1), 83-96.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies*, 21(2), 45–105.

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

Roca de Larios, J. (1996). Linearization strategies in ELF writing: some observations. *Lenguaje y Textos*, 8, 191–208.

Roca de Larios, J., Manchón, R. M., & Murphy, L. (2006). Generating text in native and foreign language writing: A temporal analysis of problem-solving formulation processes. *Modern Language Journal*, 90(1), 100–114. https://doi.org/10.1111/j.1540-4781.2006.00387.x

Roca de Larios, J., Manchón, R., & Murphy, L. (2007). Componentes básicos y evolutivos del proceso de formulación en la escritura de textos en lengua materna y lengua extranjera. *Revista Española de Lingüística Aplicada*, 20(20), 159–184.

Roca de Larios, J., Murphy, L., & Manchón, R. (1999). The use of restructuring strategies in EFL writing: A study of Spanish learners of English as a Foreign Language. *Journal of Second Language Writing*, 8(1), 13–44.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Russo, J. E., Johnson, J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory and Cognition*, 17(6), 759–769.

Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on an L2 writing revision task. *Studies in Second Language Acquisition*, *29*(1), 67-100.

Salmeron-Majadas, S., Santos, O. C., & Boticario, J. G. (2014). An evaluation of mouse and keyboard interaction indicators towards non-intrusive and low cost affective modeling in an educational context. *Procedia Computer Science*, *35*, 691-700.

Santos, M., López-Serrano, S., & Manchón, R. M. (2010). The differential effect of two types of direct written corrective feedback on noticing and uptake: Reformulation vs. error correction. *International Journal of English Studies 10*(1), 131−154.

Sasaki, M. & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning*, 46(1), 137–174.

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9(3), 259–291.

Sasaki, M. (2002). Building an empirically-based model of EFL learners' writing processes. In S. Ransdell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 49–80). Kluwer.

Sasaki, M., Mizumoto, A., & Murakami, A. (2018). Developmental trajectories in L2 writing strategy use: A self‐regulation perspective. *The Modern Language Journal*, *102*(2), 292-309.

Schilperoord, J. (1996). The distribution of pause time in written text production. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), Theories, models and meth- odology in writing research (pp. 21–35). Amsterdam University Press.

Schmidt, D. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11-26.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14(4), 357–385.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. *Attention and awareness in foreign language learning*, 9, 1-63.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). Cambridge University Press. doi:10.1017/CBO9781139524780.003

Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. In *Perspectives on individual characteristics and foreign language education* (pp. 27-50). De Gruyter Mouton.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11(2)*, 129-158.

Schmidt, R. W. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206-226.

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press. DOI: 10.1007/979-0-387-21591-4_6.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. *Talking to learn: Conversation in second language acquisition* (pp. 237-326).

Schoonen, R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts. Learning, teaching and research* (pp. 77-101). Multilingual Matters. http://dx.doi.org/10.21832/9781847691859-007

Severinson Eklundh, K., & Kollberg, P. (1996). Computer tools for tracing the writing process: From keystroke records to S-notation. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn, M. (Eds.), *Current Research in writing: Theories, models and methodology* (pp. 526–541). Amsterdam University Press.

Shaw, P., & Liu, E. T. K. (1998). What develops in the development of second-language writing. *Applied Linguistics*, 19(2), 225–254.

Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283.

Silva, T. (1992). L1 vs. L2 writing: ESL graduate students' perceptions. *TESL Canada Journal*, 10, 27-46.

Silva, T. (1993). Toward an understanding of the dis- tinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, 27(4), 657-677.

Skehan, P. (1989). *Individual Differences in Second-language Learning*. Edward Arnold.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (vol. 2), 69-94.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge University Press.

Skibniewski, L. (1988). The writing processes of advanced foreign language leaniers in their iiative and foreign languages: Evidence from think-aloud protocols. *Studia Anglica Posnaniensa*, 2(1), 177-186.

Smagorinsky, P. (1994). Think-aloud protocol analy- sis: Beyond the black box. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (pp. 3–19). Sage.

Spelman Miller, K. (2000). *Writing on-line: Temporal features of fi rst and second language written text production*. Unpublished PhD thesis, the University of Reading.

Spelman Miller, K. (2005). Second language writing research and pedagogy: A role for computer logging? *Computers and Composition*, *22*(3), 297-317.

Spelman Miller, K. (2006a). Pausing, productivity and the processing of topic in online writing. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging: Methods and applications* (pp. 131–155). Elsevier.

Spelman Miller, K. (2006b). The pausological study of written language production. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging: Methods and applications* (pp. 11–30). Elsevier.

Spelman Miller, K.; Sullivan, K.P.H. (2006) Keystroke logging – an introduction. In: K.P.H. Sullivan & E. Lindgren (Eds.) *Computer Keystroke Logging and Writing: Methods and Applications*. Elsevier.

Stapleton, P. (2010). Writing in an electronic age: A case study of L2 composing processes. *Journal of English for Academic Purposes*, *9*(4), 295-307.

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*(3), 201-233.

Storch, N. (2016). Collaborative writing. In R. M. Manchón & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 387-406). Walter de Gruyter.

Storch, N. (2018). Written corrective feedback from sociocultural theoretical perspectives: A research agenda. *Language Teaching*, *51*(2), 262-277.

Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, 32(2), 303-334. https://doi.org/10.1017/S0272263109990532

Sullivan, K., & Lindgren, E. (2002). Self-assessment in autonomous computer-aided second language writing. *ELT journal*, *56*(3), 258-266.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125-144). Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom secondlanguage acquisition* (pp. 64-81). Cambridge University Press.

Swain, M. (1998). Focus on form through conscious reflection. *Focus on form in classroom second language acquisition*, 64-81., 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). Newbury House.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. *Sociocultural theory and second language learning*, *97*, 114.

Swain, M. (2005). The output hypothesis: Theory and research. In *Handbook of research in second language teaching and learning* (pp. 495-508). Routledge.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16(3), 371-391.

Swain, M., 1997. Collaborative dialogue: its contribution to second language learning. *Revista Canaria de Estudios Ingleses*, 34, 115–132.

Swain, M., Kinnear, L., & Steinman, L. (2011). *Sociocultural theory in second language education: An introduction through narratives*. Multilingual Matters.

Swales, J. (1990). *Genre analysis.* Cambridge University Press.

Swales, J. (2004). *Research genres*. Cambridge University Press.

Swanson, H. L., & Berninger, V. W. (1996). Individual differences in children's working memory and writing skill. *Journal of experimental child psychology*, *63*(2), 358-385.

Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning- Insights from and for L2 writing* (pp. 217–236). John Benjamins.

Teddlie, C. & Tashakkori, A. (2006). A General Typology of Research Designs Featuring Mixed Methods. *Research in the Schools*, *13*(1), 12–28.

Tillema, M., Van den Bergh, H., Rijlaarsdam, G. & Sanders, T. (2011). Relating self reports of writing behaviour and online task execution using a temporal model. *Metacognition Learning*, 6, 229–253. http://dx.doi.org/10.1007/s11409-011-9072-x

Tiryakioglu, G., Peters, E., & Verschaffel, L. (2019). The effect of L2 proficiency level on composing processes of EFL learners: Data from keystroke loggings, think alouds and questionnaires. In E. Lindgren, & K. P. H. Sullivan (Eds.), *Observing Writing: Insights from Keystroke Logging and Handwriting* (pp. 212-235). Brill.

Tocalli-Beller, A., & Swain, M. (2005). Reformulation: The cognitive conflict and L2 learning it generates. *International Journal of Applied Linguistics*, *15*(1), 5-28.

Tomlin, R. S. & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. A. MacArthur, S. Graham,&J.Fitzgerald(Eds.), *Handbook ofwriting research* (pp. 67–80). Guilford Press.

Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition. *Language Learning*, 56, 319-385.

Towell, R., & Hawkins, R. D. (1994). *Approaches to second language acquisition*. Multilingual matters.

Triandis, H. C. (1971). *Attitude and attitude change* (Vol. 8). Wiley.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327–369.

Truscott, J. (1999). The case for “the case for grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, *8(1)*, 111–122.

Truscott, J., & Sharwood Smith, M. (2004). Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition*, 7(1), 1-20.

Tsao, J. J., Tseng, W. T., & Wang, C. (2017). The effects of writing anxiety and motivation on EFL college students' self-evaluative judgments of corrective feedback. *Psychological Reports*, *120*(2), 219-241.

Turner, H. M. I., & Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary issues in communication science and disorders*, *33*(Spring), 42-55.

Uggen, M. S. (2012). Reinvestigating the noticing function of output. *Language Learning*, 62(2), 506-540.

Uzawa, K. (1996). Second language learners' processes of L1 writing, L2 writing and translation from L1 into L2. *Journal of Second Language Writing*, 5(3), 271–294. http://dx.doi.org/10.1016/S1060-3743(96)90005-3

Van Beuningen, C. (2010). Corrective feedback in L2 writing: Theoretical perspectives, empirical insights, and future directions. *International Journal of English Studies*, *10*(2), 1-27.

Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, *62*(1), 1-41.

Van Den Bergh, H., & Rijlaarsdam, G. (2001). Changes in cognitive activities during the writing process and relationships with text quality. *Educational Psychology*, *21*(4), 373-385.

Van Gelderen, A., & Oostdam, R. (2002). Improving linguistic fluency for writing: Effects of explicitness and focus of instruction. *L1-Educational Studies in Language and Literature, 2*(3), 239-270.

Van Horenbeeck, E., Pauwaert, T., Van Waes, L., & Leijten, M. (2012). S-notation markup rules (Technical Description) (D. o. Management, Trans.) (pp. 7). University of Antwerp.

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(6), 829–853. https://doi.org/10.1016/S0378-2166(02)00121-2.

Van Waes, L., Leijten, M., & Van Weijen, D. (2009). Keystroke logging in writing research: Observing writing processes with Inputlog. *GFL-German as a foreign language*, *2*(3), 41-64.

Van Waes, L., Leijten, M., Wengelin, A., & Lindgren, E. (2011). 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). Psychology Press.

van Weijen, D., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2008). Differences in process and process-product relations in L2 writing. *ITL-International Journal of Applied Linguistics*, *156*(1), 203-226.

Vandermeulen, N., Leijten, M., & Van Waes, L. (2020). Reporting Writing Process Feedback in the Classroom Using Keystroke Logging Data to Reflect on Writing Processes. *Journal of Writing Research*, *12*(1), 109-139.

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Routledge.

Vygotsky, L. S. (1981). Pensamiento y palabra. *Infancia y Aprendizaje*, *4*(sup1), 15-35.

Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, *11*(3), 225-246.

Warschauer, M. (1996). Motivational aspects of using computers for writing and communication. In M. Warschauer (Ed.), *Telecollaboration in foreign Language Learning* (pp. 29-46). Second Language Teaching & Curriculum Center, University of Hawai'i at Manoa.

Wengelin, A. (2006). Examining pauses in writing: Theories, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer keystroke logging and writing: Methods and applications* (Vol. 18) (pp. 107-130). Elsevier.

Wengelin, A., Leijten, M., & Van Waes, L. (2010). Studying reading during writing: New perspectives in research. *Reading and Writing*, 23(7), 735-742. doi:10.1007/ s11145-009-9187-5

Wengelin, A., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eye-tracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods*, 41(2), 337-351.

Wertsch, J.V. (1991) A socio-cultural approach to socially shared cognition. In L. Resnick, J. Levine & S. Teasley (Eds). *Perspectives on Socially Shared Cognition* (pp. 85-100). American Psychological Association.

Whalen, K., & Ménard, N. (1995). L1 and L2 writers' strategic and linguistic knowledge: A model of multiple-level discourse processing. *Language Learning*, 44(3), 381–418.

White, R. (1988). *The ELT curriculum: Design, management, innovation*. Blackwell.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7, 133–161.

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, *21*, 321–331.

Xu, C. (2017). A temporal investigation of L2 writing process—A pausological perspective. *Foreign Language Teaching and Research*, 49(4), 583-595.

Xu, C. (2018). Understanding online revisions in L2 writing: A computer keystroke-log perspective. *System*, 78, 104–114.

Xu, C., & Qi, Y. (2017). Analyzing pauses in computer-assisted EFL writing: A computer-keystroke-log perspective. *Journal of Educational Technology and Society*, 20(4), 24–34.

Yang, L., & Zhang, L. (2010). Exploring the role of reformulations and a model text in EFL students' writing performance. *Language Teaching Research*, 14(4), 464–484.

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, *86*(1), 54-66.

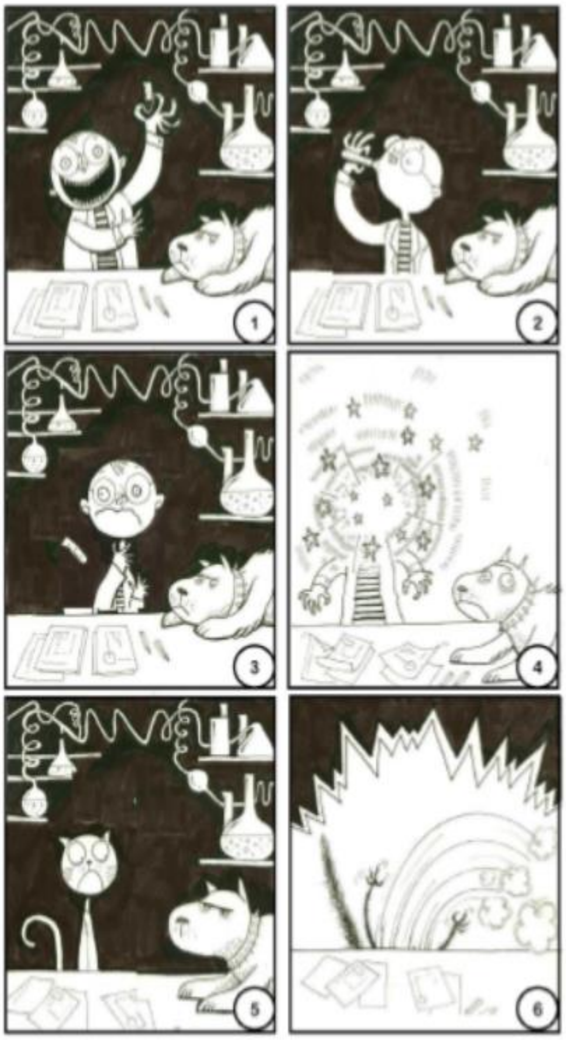
Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, *36*, 90-102.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.

Appendices

Appendix 1. "The Scientist" Six-picture frame story (Cánovas, 2017).

Observe the pictures below. Write a **brief story** which reflects what happens in the pictures. Please do not forget to write about every single picture.



Appendix 2. Socio-demographic questionnaire.

Cuestionario Socio-demográfico

TD_Q2

*Obligatorio

1. Nombre completo *

____

2. Edad *

____

3. Sexo *

Marca solo un óvalo.

- ( ) Chico
- ( ) Chica

4. Clase *

Marca solo un óvalo.

- ( ) 5º PRIMARIA
- ( ) 6º PRIMARIA

5. Fecha de HOY *

____

Ejemplo: 7 de enero del 2019

1. ¿A qué edad empezaste a estudiar inglés? *

Marca solo un óvalo.

- Antes de los 5 años
- Entre los 5 y los 7
- A los 8 años
- Entre los 9 y los 11

2. ¿Asistes a clases extraescolares de inglés? *

Marca solo un óvalo.

- Sí
- No

3. Si recibes clases de inglés extraescolares, ¿dónde?

Marca solo un óvalo.

- Clases particulares
- Academia de inglés
- Otro: ____________________

4. ¿Cuábnto tiempo llevas asistiendo?

Marca solo un óvalo.

- 0-2 años
- 2-4 años
- 4-6 años
- 6-9 años
- 10-12 años

1. ¿Cuántas horas a la semana?

Marca solo un óvalo.

- 1
- 2
- 3
- 4

2. ¿Te gusta estudiar inglés? *

Marca solo un óvalo.

- Sí
- No

3. ¿Has vivido alguna vez en un país donde se hable inglés? *

Marca solo un óvalo.

- Sí
- No

4. Si la respuesta anterior es "sí", ¿cuánto tiempo?

Marca solo un óvalo.

- Menos de 1 año
- De 2 a 3 años
- Más de 3 años

5. ¿Escuchas música en inglés? *

Marca solo un óvalo.

- No, nunca
- De vez en cuando
- A menudo
- Todos los días

1. ¿Te gusta leer en inglés? *

Marca solo un óvalo.

- Sí
- No

2. ¿Te gusta hablar en inglés? *

Marca solo un óvalo.

- Sí
- No

3. ¿Te gusta escribir en inglés? *

Marca solo un óvalo.

- Sí
- No

4. Escribir en inglés te resulta... *

Marca solo un óvalo.

- Muy fácil
- Fácil
- Difícil Muy
- difícil

1. Ordena del 1 al 4 tu preferencia en las siguientes destrezas en inglés *

1= la que menos te gusta; 4= la que más te gusta

Marca solo un óvalo por fila.

	1	2	3	4
Readiing ((lleerr))	⬭	⬭	⬭	⬭
Wrriittiing ((escrriibiirr))	⬭	⬭	⬭	⬭
Liistteniing ((escucharr))	⬭	⬭	⬭	⬭
Speakiing ((habllarr))	⬭	⬭	⬭	⬭

2. ¿Cómo consideras que se te da escribir con un teclado? *

Marca solo un óvalo.

- ⬭ Muy mal
- ⬭ Mal
- ⬭ Regular
- ⬭ Bien
- ⬭ Muy bien

3. ¿Como crees que escribes utilizando un teclado en el ordenador? *

Marca solo un óvalo.

- ⬭ Muy lento
- ⬭ Lento
- ⬭ Normal
- ⬭ Rápido
- ⬭ Muy rápido

Appendix 3. Exit questionnaire (students' perceptions on the WCF).

Cuestionario

TD_Q1

*Obligatorio

1. Nombre completo *

2. Edad *

3. Sexo *

Marca solo un óvalo.

- Chico
- Chica

4. Clase *

Marca solo un óvalo.

- 5º PRIMARIA
- 6º PRIMARIA

5. Fecha de HOY *

Ejemplo: 7 de enero del 2019

6. ¿Te ha resultado útil el modelo de texto propuesto para cometer menos errores? *

Marca solo un óvalo.

Sí

1. ¿Por qué? *

2. Si ha habido algún error que no habías corregido anteriormente, explica por qué. Si puedes, indica el tipo de error (gramática, vocabulario, ortografía o puntuación). *

Resumen de la tesis (Español)

INTRODUCCIÓN, OBJETIVOS Y PREGUNTAS DE INVESTIGACIÓN

Esta tesis doctoral se centra en la escritura digital por niños, cuyo fin último es contribuir con evidencia empírica en dos áreas disciplinares relacionadas: (i) el estudio de los procesos de escritura en una segunda lengua (L2), y (ii) los efectos del feedback correctivo escrito (WCF, en inglés). Así, el principal objetivo de esta tesis doctoral se enmarca en arrojar luz acerca de la caracterización de la dimensión temporal de los procesos de escritura en niños dada la ausencia de estudios previos relacionada con dichos procesos, y por otro lado, como estos procesos y el comportamiento pausológico se ven modificados por la provisión del feedback correctivo escrito con modelos.

La escritura en una L2 ha recibido atención considerable en las últimas décadas (Leki et al., 2008; Manchón & Polio, próximamente/2021). Así, la investigación en la escritura en L2 se ha centrado fundamentalmente en el producto escrito y en los beneficios de los distintos tipos de WCF (cf. Bitchener & Storch, 2016; Manchón, 2009a, 2009b), incluyendo variables relacionadas con la provisión o ausencia de diferentes tipos de WCF. Así, el estudio de la escritura en L2 se ha abordado también desde la perspectiva del proceso de escritura, lo que ha permitido identificar y caracterizar los diferentes procesos de escritura durante la producción de un texto en L2 (Roca de Larios, Nicolás-Conesa & Coyle, 2016; Michel et al., 2019; Révész et al., 2017, 2019). Muchos de estos estudios han tenido como eje la exploración de los procesos de escritura, su asignación estratégica durante el proceso completo así como el análisis de la complejidad de la resolución de problemas en el mismo. Más recientemente la perspectiva del potencial para el aprendizaje de idiomas de la dimensión procesual de la escritura ha sido abordada empíricamente (véase Manchón, 2011; Manchón & Vasylets, 2019; Manchón & Williams, 2016). Así, esta linea de investigación se enmarca entre los estudios de escritura y adquisición de segundas lenguas (ASL), y en la perspectiva de *writing-to-learn* considerando la escritura como clave para la contribución en el aprendizaje de lenguas (Manchón, 2011a).

En linea con lo anterior, la investigación en procesos de escritura en L2 desde la perspectiva de la escritura y del procesamiento del WCF adquiere especial relevancia para esta tesis doctoral. Así, la descripción y análisis de los procesos de escritura así como de comportamientos pausológicos ha suscitado gran interés en las últimas décadas, mediante la verificación de los modelos de escritura (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; Hayes, 2012; Kellogg, 1996). La evidencia empírica ha evidenciado cambios estratégicos en el comportamiento de adultos y adolescentes en los procesos de escritura en L2 en planificación (Johnson, 2020; Manchón & Roca de Larios, 2007), formulación (Roca de Larios et al., 2006; Zimmerman, 2000) o revisión (Stevenson et al., 2006; Xu, 2018). Sin embargo, apenas se han analizado y estudiado los procesos de escritura en los estudiantes más jóvenes, como los niños de primaria.

De la misma forma, los trabajos empíricos que examinan los procesos de escritura en L2 han observado fundamentalment dichos procesos en entornos tradicionales (Roca de Larios et al., 2008). Sin embargo, el papel de la escritura digital ha supuesto un añadido importante en esta línea de investigación con la inclusión de nuevas herramientas metodológicas como el software de registro de pulsación de teclas (en inglés, *keystroke logging*). Estas herramientas, como Inputlog (Leijten & Van Waes, 2013) han abierto nuevas posibilidades para estudiar el proceso de escritura de una forma discreta sin que pueda afectar el diseño de la investigación ni lo que se pretende observar. Numerosos estudios han hecho uso de esta herramienta para observar los procesos de escritura de la L2 en adolescentes y adultos (por ejemplo, Chukharev-Hudilainen et al., 2019; Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2017). En este contexto, *keystroke logging* (es decir, un entorno digital) para el estudio del proceso de escritura en L2 apenas ha sido usado como herramienta metodológica con niños, por lo que requiere de más evidencia empírica.

Finalmente, como ocurre en otras areas de investigación en ASL y su orientación hacia la escritura en L2, la mayoría de la evidencia empírica deriva de estudios de L2 con adultos y adolescentes. Por ende, muy pocos estudios (Cánovas et al., 2015; Coyle & Cánovas, 2019; Coyle et al., 2014, 2018; Coyle & Roca de Larios, 2014; Roca de Larios et al., 2015) han investigado el papel de los modelos como WCF en niños. Estudios previos han sugerido que este tipo de feedback activa una comparación cognitiva, y da lugar a procesos como planificación y revisión durante la escritura (Schdmit, 2001). Por otro lado, no hay apenas investigación con respecto a cómo los procesos de escritura en L2 se ven modificados por la provisión de feedback, y su influencia en su distribución temporal de dichos procesos, así como el comportamiento pausológico.

Teniendo en cuenta todo lo anterior, la presente tesis doctoral pretende aportar nueva evidencia empírica relacionada con la implementación y la distribución temporal del proceso de escritura en L2 en niños, una población inexplorada en estudios previos centrados en adultos (por ejemplo, Roca de Larios et al., 2008; Révész et al., 2019). Por otro lado, hemos hecho uso de una herramienta de *keystroke logging* como Inputlog (Leijten & Van Waes, 2013) como instrumento principal de recogida de datos respondiendo a la necesidad de aplicar nuevas herramientas metodológicas para observar el proceso de escritura en niños y el impacto de los modelos como WCF. Por otro lado, la decisión de utilizar modelos se basa en las predicciones de las propiedades lingüísticas de los modelos en cuanto a la selección léxica y la generación de ideas (Coyle et al., 2018), lo que puede influenciar la forma en la que los procesos de escritura y el comportamiento pausológico se ubican en la distribución temporal del proceso de escritura. Finalmente, nuestro estudio es fundamentalmente exploratorio dada la ausencia de estudios que traten la influencia de modelos como WCF en la dimensión temporal de la escritura y, más específicamente, la escasez de estudios que examinen los procesos de escritura en L2 de los niños en entornos digitales.

Con la finalidad de dar respuesta a estos interrogantes empíricos, las siguientes preguntas de investigación guiaron nuestro estudio:

RQ1. ¿Hasta qué punto hay diferencias en la distribución temporal de los procesos de escritura de niños en lengua extranjera (es decir, planificación, formulación, y revisión) cuando escriben y cuando revisan el mismo texto en el ordenador, con y sin recibir WCF en forma de modelos?

RQ2. ¿Hasta qué punto hay diferencias en el comportamiento pausológico de niños cuando escriben en lengua extranjera cuando escriben y cuando revisan el mismo texto en el ordenador, con y sin recibir WCF en forma de modelos?

RECOGIDA DE DATOS Y METODOLOGÍA

A partir de una muestra de conveniencia, un grupo intacto de niños de diez años pertenecientes a una clase de Educación Primaria participaron en el estudio. El diseño de la investigación es experimental, y se trata de un estudio de laboratorio basado en aula experimental. En total, dieciocho participantes tomaron parte en la investigación, los cuales fueron asignados aleatoriamente a dos grupos diferentes: grupo experimental (n= 10) y grupo control (n= 8). Todos los tutores legales firmaron un formulario de consentimiento en el que daban su autorización para que los niños tomaran parte en el estudio.

Así, el diseño de la investigación consistió en tres fases siguiendo estudios previos con modelos como WCF (Coyle & Cánovas, 2020; Luquín & García Mayo, 2020): fase 1, en la que los niños de ambos grupos completaron una tarea escrita por ordenador individualmente usando Inputlog como herramienta de captura de pulsaciones. La tarea consistió en escribir una historia a partir de seis imágenes organizadas de forma cronológica; fase 2, que constituyó el tratamiento. Así, el grupo experimental tuvo acceso al texto escrito en la fase 1, y al feedback en forma de un modelo de respuesta a la tarea que habían realizado. El grupo de control recibió acceso a su texto de la fase 1 pero sin feedback, y autoeditaron su texto. Tras esto, ambos participaron en un protocolo de entrevista de recuerdo estimulado, pero dichos datos no fueron tenidos en cuenta para la presente tesis doctoral dada la pobreza de los mismos; en cuanto a la fase 3, tanto el grupo experimental como el grupo control reescribieron los textos producidos en la fase 1, reproduciendo las mismas condiciones.

La variable independiente en nuestro estudio es el WCF con dos niveles, ausencia y presencia. Las variables dependientes son la dimensión temporal de los procesos cognitivos durante el proceso de composición, y el comportamiento pausológico.

Para analizar las diferencias entre la fase 1 y la fase 3, y observar la posible incidencia del tratamiento en las variables dependientes, se operacionalizaron los procesos de escritura en planificación, formulación y revisión. Planificación fue subdividido en dos procesos: planificación global y local; por otro lado, revisión fue subdividido en dos procesos: revisión microscópica y macroscópica. Estas variables fueron medidas con varios niveles: frecuencia (es decir, número de episodios), duración total (expresada en porcentajes), así como ediciones y palabras producidas (solo en el caso de formulación y revisión). Por otro lado, la segunda variable independiente fue operacionalizada en varios medidas pausológicas obtenidas de Inputlog: duración de las pausas (expresadas con los segundos), distribución de las pausas en intervalos, frecuencia de las pausas, ubicación de las pausas (es decir, dentro de la palabra, después de párrafo, entre otras), así como medidas de fluidez (pausas por minuto y intervalo entre pulsaciones).

Los análisis estadísticos que se llevaron a cabo comprendieron la descripción estadística de lo observado entre la fase 1 y fase 2 con medidas de tendencia central (media) y de dispersión (desviación estándar). Dado que el tamaño de la muestra fue demasiado pequeño, se utilizó el tamaño del efecto con la *g* de Hedges (Larson-Hall & Plonsky, 2015; Turner & Bernard, 2006) con el fin de averiguar la magnitud del efecto de la variable independiente en las diferentes variables dependientes.

RESULTADOS Y DISCUSIÓN

En cuanto a la primera pregunta de investigación (es decir, efecto de los modelos en los procesos de escritura), se encontraron efectos significativos en la frecuencia del proceso de planificación al principio y final del proceso escritor, en los que el grupo que recibió WCF redujo dicha frecuencia, pero utilizó más tiempo para este proceso. Por otro lado, los espisodios de planificación global se redujeron a mitad del proceso de escritura en el grupo con WCF, aunque el número fue mayor que en el grupo control. Por otro lado, se encontraron resultados significativos en planificación local, en el que el grupo con WCF recurrió mucho más a este tipo de proceso que el grupo control, si bien hubo reducción en ambos. Estos resultados indican que los modelos pueden haber contribuido a mantener la atención a la forma así como a expandir el repertorio léxico (Eschholz, 1980; García Mayo & Labandibar, 2017; Roca de Larios et al., 2015). Así, el hecho de que la planificación local se haya mantenido en términos de frecuencia tras recibir WCF frente al grupo control sugiere la naturaleza de los niños como planificadores emergentes (Cumming, 1989).

En cuanto al proceso de formulación, se confirmó que dicho proceso llegaba a ocupar la mitad del tiempo de la tarea, lo cual iba acorde con lo propuesto en estudios previos (Gánem-Gutiérrez & Gilmore, 2018; Roca de Larios et al., 2008). Aunque no de forma significativa, ambos grupos incrementaron el tiempo dedicado al proceso de formulación, lo cual puede explicarse por problemas de compensación (Roca de Larios et al., 2006) así como cuestiones de teclado (Kellogg et al., 2013). Adicionalmente, se encontraron diferencias significativas en las ediciones y el número de palabras producidas. Así, el grupo con WCF necesitó menos ediciones para completar tramos de texto, aunque el número fue mayor que en el grupo control, que lo redujo. Esto se explica por la reducción de los procesos ascendentes que requieren esfuerzo cognitivo (Chukharev-Hudilainen et al., 2019; Kormos, 2019). De la misma forma, el grupo con WCF redujo escasamente el número de palabras producidas durante este proceso aunque el número seguía siendo mayor que el grupo control, lo cual puede explicarse por el aumento del léxico o ideacional tras la recepción de WCF.

Los resultados del proceso de revisión indican que el grupo con WCF revisó más al final del proceso que en el resto de intervalos. No obstante, este grupo redujo considerablemente el tiempo dedicado al proceso de revisión a mitad del proceso, quizá relacionado con el incremento del proceso de formulación. En la misma linea, ambos grupos dedicaron más tiempo a las revisiones microscópicas a mitad y final del proceso, lo que puede indicar sustitución de aspectos gramaticales y léxico (Barkaoui, 2007; Stevenson et al., 2006), confirmando la atención de escrtores con menos habilidad en cambios a nivel superficial. Finalmente, el grupo con WCF revisó más a nivel macroscópico que el grupo de control, lo que sugiere que los modelos pueden haber contribuido a incrementar la atención del grupo con WCF hacia el contenido ideacional.

En cuanto a la segunda pregunta de investigación, se encontraron diferencias significativas en el tiempo de pausa durante todo el proceso, que fue mayor en el grupo con WCF frente al grupo control, que lo redujo. Esto refleja operaciones de reflexión que requieren de la activación de procesos cognitivos de alto nivel (Barkaoui, 2016, 2019; Chenu et al., 2014; Medimorec & Risko, 2017). El número de pausas también se vio reducido en el grupo con WCF. En lo que se refiere a la ubicación de las pausas, los cambios más significativos se encontraron en pausa a nivel de palabra, en las que el grupo con WCF pausó con menos frecuencia, probablemente como consecuencia de reducción de la frecuencia en planificación global (Alves et al., 2007). Por otro lado, el grupo con WCF mantuvo (aunque de forma más reducida) las pausas intra-palabra, probablemente relacionadas con problemas de teclado o de ortografía. Este incremento se vio reflejado también en las pausas después de palabra, reflejando monitoreo (Révész et al., 2017, 2019). De forma similar, el grupo con WCF incrementó las pausas a nivel de párrafo, lo que nos induce a pensar que el modelo indujo a planificar el contenido previamente a ser modificado de forma macroscópica. Finalmente, se encontraron diferencias significativas en otras pausas (por ejemplo, las concernientes a movimientos del cursor, o a operaciones de revisión), en las que el grupo con WCF redujo el número de pausas de este tipo pero mantuvo la duración media de las mismas, lo que indica un efecto muy débil de los modelos en estas pausas. Así, también se arrojó luz con respecto al efecto límite del texto, que no fue reproducido por nuestro estudio en niños frente a estudios previos con adultos (Chanquoy et al., 1996; Schilperoord, 2001; Wengelin et al., 2009), lo cual puede explicarse por la edad así como por el medio de escritura y el entorno de la tarea.

Finalmente, debemos mencionar ciertas limitaciones de la presente tesis doctoral. En primer lugar, el número limitado de participantes (N= 18) no permite obtener conclusiones extrapoladas ni concluyentes, aunque nuestro estudio sea puramente exploratorio. Por otro lado, limitaciones metodológicas como el uso de entrevistas de recuerdo estimulado deben ser reconocidas. Así, el uso de este procedimiento metodológico con niños ha probado no aportar datos ricos, por lo que no pudo ser utilizado para complementar los datos de Inputlog. De esta forma, la ausencia de un procedimiento concurrente de este tipo ha de ser abordado en investigación futura, pues el uso de entrevistas de recuerdo estimulado para el pre-test y post-test permitiría obtener datos ricos sobre el proceso antes y después del feedback. No obstante, primero se han de resolver aspectos relacionados con cómo los niños responden a dicho procedimiento, y qué se puede hacer para que los datos sean más funcionales. La investigación futura ha de intentar dar respuesta a cómo los procesos de escritura se desarrollan en niños no solo con más investigación con modelos, sino con diferentes tipos de WCF. De la misma forma, se deben explorar variables que incluyan la capacidad de la memoria de trabajo. Así, también se debe orientar la investigación – conjuntamente con la entrevista de recuerdo estimulado – para profundizar más en cuanto a las operaciones subyacentes a los procesos de escritura (por ejemplo, planteamiento de objetivos, organización de las ideas, entre otros).