



Corrective feedback in the EFL classroom:  
*Grammar Checker* vs. teacher's feedback

**Doctoral Dissertation**

Presented by:

Tamara Hernández Puertas

Supervised by:

Dr. Patricia Salazar Campillo

January 2022





Programa de Doctorat en Llengües Aplicades, Literatura i Traducció.  
Escola de Doctorat de la Universitat Jaume I.

## Corrective Feedback in the EFL Classroom: *Grammar Checker vs. Teacher's Feedback.*

Memòria presentada per Tamara Hernández Puertas per a optar al grau de doctora per la Universitat Jaume I.

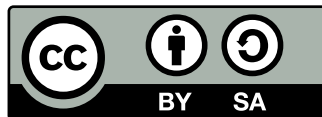
Tamara Hernández Puertas

Patricia Salazar Campillo

Firmado por PATRICIA  
SALAZAR CAMPILLO -  
NIF:18979168M el día  
15/01/2022 con un  
certificado emitido por  
ACCVCA-120

Castelló de la Plana, January 2022

**Creative Common Licence:  
Attribution  
NoCommercial  
ShareAlike  
(CC BY-NC-SA)**



No s'ha rebut finançament

## Acknowledgements

---

After 8 years of struggling, the day has arrived. Thinking it was never going to take place, after a lot of effort, today is the day. I cannot but thank those people who have made this moment possible. The first person and probably the most important one is my supervisor, Dr. Patricia Salazar, who has been all the way with me. Thanks, Patricia, for your support, encouragement, patience, and guidance throughout this process. The second one is my family, who have mentally supported me all these years: my parents, my sister, my husband and my two daughters. This dissertation would not have been possible if they had not helped me become the person I am. The third one is Luis, who assisted me in the statistical analysis along with Jim Lawley and Rubén Chacón who also helped me to understand how the computer program *Grammar Checker* works.

Needless to say, I am indebted to the 2B2 students who participated in the pilot study prior to this research and also the 2B2 students at *Escola Oficial d'Idiomes de Castelló* who took part in the final study. Without their participation, this Ph.D. Dissertation would not have been possible.

Thanks to all those who have given me the opportunity to fulfil this ambition

*To my family*





# Table Of Contents

|   |     |
|---|-----|
| List Of Figures.....  | 15  |
| List Of Tables .....  | 17  |
| List Of Abbreviations .....   | 19  |
| Introduction.....   | 23  |
| Part I: Literature review.....  | 25  |
| Chapter 1. Key Issues in Second Language Acquisition .....                  | 27  |
| 1.1 The Role of Input in Language Learning.....                             | 27  |
| 1.1.1 <i>Comprehensible input and language learning</i> .....               | 31  |
| 1.1.2 <i>Input and intake</i> .....   | 33  |
| 1.2 Output in the Language Learning Process .....                           | 35  |
| 1.3 Corrective Feedback in Language Learning.....                           | 40  |
| 1.3.1 <i>Definition of feedback</i> .....                                   | 41  |
| 1.3.2 <i>Error and mistake</i> .....  | 43  |
| 1.3.3 <i>Error analysis</i> .....   | 45  |
| 1.3.4 <i>Error treatment</i> .....  | 53  |
| 1.3.5 <i>Noticing and focus-on-form approach</i> .....                      | 58  |
| 1.4 L2 Acquisition of the Writing Skill .....                               | 63  |
| 1.5 L2 Acquisition of Grammar.....  | 65  |
| 1.5.1 <i>The acquisition of English articles</i> .....                      | 68  |
| 1.5.2 <i>The acquisition of English prepositions</i> .....                  | 72  |
| 1.5.3 <i>The acquisition of English simple and perfect tenses</i> .....     | 75  |
| 1.6 Chapter Summary .....   | 79  |
| Chapter 2. The Role of Feedback in Language Learning.....                   | 81  |
| 2.1 Types of Feedback and Uptake.....                                       | 82  |
| 2.2 Focused and Unfocused CF .....  | 89  |
| 2.3 Immediate vs Delayed CF.....  | 95  |
| 2.4 Implicit vs Explicit CF.....  | 98  |
| 2.5 Oral and Written CF.....  | 103 |
| 2.5.1 <i>Oral CF</i> .....  | 104 |
| 2.5.2 <i>Written CF</i> .....   | 108 |
| 2.5.2.1 <i>Direct and indirect WCF</i> .....                                | 110 |
| 2.6 Learners' and Teachers' Perceptions of Written Corrective Feedback..... | 117 |
| 2.7 Feedback Provider .....   | 121 |

|  |     |
|--|-----|
| 2.7.1 <i>Teacher's feedback</i> .....  | 121 |
| 2.7.2 <i>Peer feedback</i> .....   | 123 |
| 2.7.3 <i>Self-feedback</i> .....   | 126 |
| 2.7.4 <i>Automated feedback</i> .....  | 129 |
| 2.8 Chapter Summary .....  | 130 |
| Chapter 3. Computer Technology Applications in Language Learning.....        | 131 |
| 3.1 Computer-Assisted Language Learning .....                                | 132 |
| 3.1.1 <i>Electronic feedback</i> .....                                       | 135 |
| 3.1.2 <i>Computer-generated feedback</i> .....                               | 138 |
| 3.2 AWE Programs.....  | 143 |
| 3.2.1 <i>Microsoft Word</i> .....  | 144 |
| 3.2.2 <i>MyAccess!</i> .....   | 145 |
| 3.2.3 <i>GrammarCheck</i> .....  | 147 |
| 3.2.5 <i>Grammarly</i> .....   | 149 |
| 3.2.6 <i>Criterion</i> .....   | 151 |
| 3.2.7 <i>Grammar Checker</i> .....   | 153 |
| 3.2.8 <i>Intelligent Academic Discourse Evaluator</i> .....                  | 157 |
| 3.3 Learners' and Teachers' Perceptions on Computer-Generated Feedback ..... | 159 |
| 3.4 Chapter Summary .....  | 162 |
| Part III: The Study .....  | 165 |
| Chapter 4. Motivation for the Study .....                                    | 167 |
| 4.1 Purpose.....   | 168 |
| 4.2 Research Questions and Hypotheses.....                                   | 170 |
| Chapter 5. Method .....  | 173 |
| 5.1 Pilot Study.....   | 173 |
| 5.2 The Study.....   | 179 |
| 5.2.1 <i>Setting and participants</i> .....                                  | 179 |
| 5.2.2 <i>Targeted grammatical features</i> .....                             | 180 |
| 5.2.3 <i>Instruments</i> .....   | 181 |
| 5.2.3.1 <i>Grammar Checker</i> .....   | 181 |
| 5.2.3.2 <i>Comic strips</i> .....  | 181 |
| 5.2.3.3 <i>Tailor-made test</i> .....  | 182 |
| 5.2.3.4 <i>Questionnaire</i> .....   | 183 |
| 5.2.4 <i>Data collection procedure</i> .....                                 | 183 |
| 5.2.5 <i>Data analysis</i> .....   | 186 |

|  |     |
|--|-----|
| 5.3 Chapter Summary .....                                | 188 |
| Chapter 6. Results and Discussion .....                  | 189 |
| 6.1 Results and Discussion Related to Hypothesis 1 ..... | 190 |
| 6.1.1 <i>Prepositions</i> .....                          | 191 |
| 6.1.2 <i>Articles</i> .....                              | 199 |
| 6.1.3 <i>Tenses</i> .....                                | 206 |
| 6.1.4 <i>Summation</i> .....                             | 210 |
| 6.2 Results and Discussion Related to Hypothesis 2 ..... | 213 |
| 6.3 Results and Discussion Related to Hypothesis 3.....  | 221 |
| 6.4 Results and Discussion Related to Hypothesis 4 ..... | 226 |
| 6.5 Chapter Summary .....                                | 236 |
| Chapter 7. Conclusions.....                              | 239 |
| 7.1 Summary of the Main Findings.....                    | 239 |
| 7.2 Limitations of the Study.....                        | 246 |
| 7.3 Contribution of the Study and Further Research.....  | 247 |
| 7.4 Pedagogical Implications .....                       | 249 |
| References.....  | 251 |
| Appendices.....  | 285 |
| Appendix 1: Comic strips.....                            | 285 |
| Appendix 2: Questionnaire .....                          | 289 |
| Appendix 3: Examples of GC and teacher feedback .....    | 292 |
| Appendix 4: Tailor-made test .....                       | 314 |



## List Of Figures

|  |     |
|--|-----|
| Figure 1. A model of second language acquisition<br>(Gass & Selinker, 2008, p. 481) .....  | 30  |
| Figure 2. A computational model of L2 acquisition (adapted from Ellis, 1997, p. 35) .....  | 31  |
| Figure 3. Article continuum (Chesterman, 1991, p. 182).....  | 68  |
| Figure 4. Error treatment sequence (Lyster and Ranta, 1997, p. 44).....  | 84  |
| Figure 5. Uptake as a facilitator of language acquisition<br>(Hassanzadeh Nezami, 2012, p. 7) .....  | 88  |
| Figure 6. Contribution of self-assessment to learning (Ross, 2006, p. 6).....  | 128 |
| Figure 7. A simple conceptualisation of the CALL perspective.<br>(Levy & Hubbard, 2005, p. 146).....   | 134 |
| Figure 8. Microsoft Word ( <a href="https://www.microsoft.com/es-es/microsoft-365/word">https://www.microsoft.com/es-es/microsoft-365/word</a> ).....  | 145 |
| Figure 9. MyAccess!<br>( <a href="https://www.myaccess.com/myaccess/assets/newui/images/ma.promo-dashboard.png">https://www.myaccess.com/myaccess/assets/newui/images/ma.promo-dashboard.png</a> ) ..... | 146 |
| Figure 10. GrammarCheck ( <a href="https://www.grammarcheck.net/editor/">https://www.grammarcheck.net/editor/</a> ).....   | 148 |
| Figure 11. Ginger Grammar Checker ( <a href="https://www.gingersoftware.com/grammarcheck">https://www.gingersoftware.com/grammarcheck</a> ) .....  | 149 |
| Figure 12. Grammar Checker (Chacón, 2016).....   | 155 |
| Figure 13. Problem Words in Grammar Checker (Chacón, 2016).....  | 156 |
| Figure 14. Pairs Filters in Grammar Checker (Chacón, 2016).....  | 157 |
| Figure 16. Chronological order of the pilot study .....  | 175 |
| Figure 17. Number of errors and their correction in the pilot study.....   | 177 |
| Figure 18. Chronological order of the study.....   | 185 |
| Figure 19. Errors in prepositions before feedback (Group 1) .....  | 192 |
| Figure 20. Errors in prepositions after feedback (Group 1) .....   | 192 |
| Figure 21. Errors in prepositions before feedback (Group 2) .....  | 194 |
| Figure 22. Errors in prepositions after feedback (Group 2) .....   | 194 |
| Figure 23. Errors in prepositions before feedback (Group 3) .....  | 197 |
| Figure 24. Errors in prepositions after feedback (Group 3) .....   | 197 |
| Figure 25. Errors in articles before feedback (Group 1) .....  | 201 |
| Figure 26. Errors in articles after feedback (Group 1) .....   | 201 |
| Figure 27. Errors in articles before feedback (Group 2) .....  | 203 |
| Figure 28. Errors in articles after feedback (Group 2) .....   | 203 |
| Figure 29. Errors in articles before and after receiving feedback (Group 3) .....  | 205 |
| Figure 30. Errors in tenses (simple past-present/past perfect) before feedback. ....   | 208 |
| Figure 31. Errors in tenses (simple past-present perfect) after feedback .....   | 209 |
| Figure 32. Comparison of total number of errors between ABBY and PAM .....   | 212 |
| Figure 33. Most valued errors for correction (Group 1).....  | 229 |
| Figure 34. Technological challenge (Group 1).....  | 231 |
| Figure 35. Most valued errors for correction (Group 2).....  | 233 |
| Figure 36. Most valued errors for correction (Group 3).....  | 234 |



## List Of Tables

|  |     |
|--|-----|
| Table 1. Sample studies on focused WCF (Qui Guo, 2015, p. 45) .....  | 93  |
| Table 2. Main differences & similarities between oral and written CF (Tuzi, 2004, p. 219) .....                        | 103 |
| Table 3. Oral corrective feedback types (Lyster, Saito & Sato, 2012, p. 4).....  | 105 |
| Table 4. General differences between oral, written and e-feedback (Tuzi, 2004, p. 219).....                            | 136 |
| Table 5. Types of error comments provided by Criterion .....   | 152 |
| Table 6. Full and shortened names for the data used in statistical analysis .....                                      | 187 |
| Table 7. Kruskal-Wallis H Test among groups<br>for ABBY pre and effect size (partial eta square).....                  | 189 |
| Table 8. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 1 and for prepositions .....          | 191 |
| Table 9. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 2 and for prepositions .....          | 193 |
| Table 10. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 3 and for prepositions .....         | 196 |
| Table 11. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 1 and for articles .....             | 200 |
| Table 12. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 2 and for articles .....             | 202 |
| Table 13. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 3 and for articles .....             | 204 |
| Table 14. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 1 and for tenses.....                | 207 |
| Table 15. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 2 and for tenses.....                | 207 |
| Table 16. Wilcoxon signed-rank test for<br>Pre vs. Post and effect size for Group 3 and for tenses.....                | 208 |
| Table 17. Wilcoxon signed-rank test and effect size for<br>Group 1 for SUM_Pre vs. SUM_Post for ABBY and PAM .....     | 210 |
| Table 18. Wilcoxon signed-rank test and effect size for<br>Group 2 for SUM_Pre vs. SUM_Post for ABBY and PAM .....     | 211 |
| Table 19. Wilcoxon signed-rank test and effect size for<br>Group 3 for SUM_Pre vs. SUM_Post for ABBY and PAM .....     | 211 |
| Table 20. Kruskal-Wallis H test between groups for<br>ABBY_pre and PAM_pre and effect size (partial eta square).....   | 214 |
| Table 21. Kruskal-Wallis H test between groups for<br>ABBY post and PAM post and effect size (partial eta square)..... | 215 |
| Table 22. Wilcoxon test ABBY vs. PAM and Effect size for Group 1 .....   | 216 |
| Table 23. Wilcoxon test ABBY vs. PAM and Effect size for Group 2 .....   | 217 |
| Table 24. Wilcoxon test ABBY vs. PAM and Effect size for Group 3 .....   | 218 |
| Table 25. Wilcoxon test for TAILOR (Pre vs. Post) and Effect size for Group 1.....                                     | 222 |
| Table 26. Wilcoxon test for TAILOR (Pre vs. Post) and Effect size for Group 2.....                                     | 222 |
| Table 27. Wilcoxon test for TAILOR (Pre vs Post) and Effect size for Group 3.....                                      | 223 |
| Table 28. Summary of main findings of the study.....   | 237 |





# List Of Abbreviations

ACP: Article Choice Parameter  
AH: Aspect Hypothesis  
AI: Artificial Intelligence  
+ART: a language with an article system  
-ART: a language without an article system  
AWE: Automated Writing Evaluation  
CA: Contrastive Analysis  
CAH: Contrastive Analysis Hypothesis  
CAI: Computed-Aided instruction  
CALL: Computer-Assisted Language Learning  
CF: Corrective Feedback  
CG: Control Group  
CL: Contrastive Linguistics  
DWCF: Direct Written Corrective Feedback  
EA: Error Analysis  
EFL: English as a Foreign Language  
ESL: English as a Second Language  
FL: Foreign Language  
FonF: Focus on Form  
GC: Grammar Checker  
H1: Hypothesis 1  
H2: Hypothesis 2  
H3: Hypothesis 3  
H4: Hypothesis 4  
IADE: Intelligent Academic Discourse Evaluator  
ICALL: Intelligent Computer Assisted Language Learning  
ID: Individual Difference  
IWCF: Indirect Written Corrective Feedback  
LAD: Language Acquisition Device  
L1: First language  
L2: Second Language  
NLP: Natural Language Processing  
NNS: Non-native speaker  
NT: Native Language  
RQ1: Research Question 1  
RQ2: Research Question 2  
RQ3: Research Question 3  
RQ4: Research Question 4  
SLA: Second Language Acquisition  
TL: Target Language  
UG: Universal Grammar  
WCF: Written Corrective Feedback



# **Introduction**



# Introduction

From the earliest use of basic technology (e.g., film-strip projector, overhead projector, and photocopier) in the classrooms in the 50s to the first drilling and practice programs in the 80s, computer-aided instruction (CAI, henceforth) has gained acceptance in educational contexts over the years. However, despite foreign language learners' main goal of achieving a high proficiency level (in English, in the present study) and the need to apply new technologies to language learning, few studies have investigated the effects of using computer programs as a tool to reduce grammar errors in L2 writing, even in the 21st century.

The internalisation and application of the knowledge students learn in the foreign language classroom in the written and spoken form is a concern for both teachers and students, as the latter keep repeating the same errors even at higher levels. When dealing with effective second language writing, regular and proper feedback on the students' drafts is necessary (Bitchener & Ferris, 2012; Burke, 2014; Ferris, 2003; Hu, 2007). However, daily practice with large classes is an arduous task for teachers, who may find real difficulties to meet deadlines and provide on-time and accurate feedback. Thus, tools that ease both the teachers' and the students' workload are a need and Automated Writing Evaluation (AWE, henceforth) software is gaining more and more relevance in this respect.

Computer programs are able to provide micro-level (grammar, spelling, etc.) and macro-level (content, organisation, etc.) feedback as well as holistic scores. Even so, there is still debate on the reliability and effectiveness of these programs (Burstein & Chodorow, 1999; Elliot, 2003; Ramineni, 2013; Rudner, Garcia & Welch, 2006). To bridge this gap, the present study aims at shedding light on the topic.

Our research has been conducted in a real context where the use of computer-generated feedback by the program *Grammar Checker* (GC, henceforth) and the teacher's feedback were implemented. We aimed at exploring the suitability of GC in the EFL classroom, to observe the participants'

development from the first writing to the last one, and to compare the effectiveness of GC with the teacher's feedback.

The present dissertation tackles the different questions related to the learning process of a language through computer programs. In order to do that, the study has been divided into two parts: Part I embodies the theoretical framework and is divided, in turn, into three chapters, and Part II constitutes the study and consists of four chapters. After each chapter, a summary is included to highlight the most relevant points.

Chapter 1 deals with key issues in language learning which constitute the basis of the dissertation. Concepts such as input, intake, output, feedback and error are explained and analysed in order to introduce the study. The skill of writing is also addressed as participants in the study used this skill to widen their language proficiency. In addition, a section is devoted to second language acquisition (SLA) and, more specifically, to the three targeted grammatical features analysed in this study.

Chapter 2 focuses on the role of feedback in language learning and specifically addresses issues such as uptake, corrective feedback (CF) effectiveness and its implementation in writing development. Different types of CF (focused/unfocused, immediate delayed, implicit/explicit, oral/written, teacher/peer/self/computer) as well as the teacher's and students' perceptions about feedback are examined. In this chapter, the term computer-generated feedback is already introduced and explained. Chapter 3 introduces CF into Computer-Assisted Language Learning (CALL) and Automated Writing Evaluation programs. It focuses on the use of computer programs as a tool to provide feedback and support SLA, and some of the most relevant options available in the market.

Taking into account the theoretical framework of the previous chapters, Chapter 4 in Part II examines the motivation for the study, the research questions and hypotheses proposed. Chapter 5 comprises all the relevant information of the study design. That is to say, we will discuss the method applied, an overview of the pilot study conducted before the final study, the setting and participants who took part in the research, the instruments employed, the data collection procedure and the data analysis. Results are presented in Chapter 6, where they are discussed in light of previous literature on computer programs and traditional CF, i.e., teacher's feedback. Finally, as concluding remarks, Chapter 7 centres on the summary of the main findings as well as the limitations of the study, further research and some pedagogical implications deriving from the study.

## **Part 1: Literature review**





## Chapter 1. Key Issues in Second Language Acquisition

The conditions for language acquisition to occur differ from theory to theory, since not all the authors agree on a fixed set. Behaviourists, nativists, cognitivists and interactionists emphasise different elements as necessary for language learning. In addition, different theories of language acquisition (e.g., Swain's (1985) *Comprehensible Output Hypothesis*, Krashen's (1981) *Input Hypothesis*, Selinker's (1972) *Interlanguage Hypothesis*, etc.) explain the learning process in a different way. However, different theories agree on the fact that "the learning of a second language is a multifaceted endeavour" (Gass & Selinker, 1994, p. 398), where distinct theories complement each other at times and even overlap to some extent.

Since 1989, when Spolsky identified a wide array of conditions leading to a general theory of L2 learning, the field of SLA has witnessed a dramatic increase of research. The human ability to use language is a learned behaviour surrounded by certain environmental conditions which contribute to its acquisition. Without those conditions, the development of language can be affected. For these reasons, in the next sections we will analyse crucial aspects in the language learning process.

### 1.1 The Role of Input in Language Learning

The importance of input and its close relationship with SLA has been under research over the last 40 years (Doughty & Long, 2003; Ellis, 1994, 1997; Gass & Selinker, 1994; Grady et al., 2011; Hart & Risley, 1995; Long, 1982; Nassaji & Fotos, 2011; VanPatten & Benati, 2010; Pica, Young & Doughty, 1987; VanPatten & Williams, 2007; Williams, Ritchie & Taj, 1999). Along these years, input has been defined as "what the learner hears and attempts to process" (Hatch, 1983, p. 81), "the language that a learner hears (or reads)" (VanPatten, 2003, p. 25), and above all, the "sine qua non of acquisition" (Gass & Mackey, 2007, p. 177).

Taking into account these definitions, as Carroll (2001, p. 2) explains, "one point on which there is consensus is that SLA requires exposure to second language". Van Patten (2000) also claims that "we seem to concur that [input] is somehow central to SLA; that without it successful SLA is not possible" (p. 295).

Input is a relevant element for the different schools, behaviorists, mentalists and interactionists (Ellis, 2008), as well as for many theories such as Krashen's *Input Hypothesis*, Long's *Input-interaction Model*, Chomsky's *Universal Grammar* (UG), and so on. SLA can be explained through different perspectives but all of them take language input as a key element. The behaviourist perspective views SLA as a process in which learners are constantly exposed to input in the form of stimulus and feedback. This approach does not consider the mental process occurring but the reaction to the stimuli received as a process of habit formation. Input comes into a black box (a metaphor for the human brain) and comes out. This view was very influential in the 60's and 70's and was supported by Skinner.

In the case of the mentalists (nativists), they believe people rely on their innate knowledge and the only thing they need is minimal exposure to input to activate acquisition (Ellis, 1997). Chomsky is one of the supporters of this view and explained that children are born with a fixed set of rules and an innate capacity to process the language, which is called the Language Acquisition Device (LAD)<sup>1</sup>, and is capable of handling first and second language acquisition. In turn, interactionists support that language acquisition occurs as a result of interaction between input and language processing, that is to say, between the learners' mental abilities and the linguistic environment. Its main figure is Piaget, who defended the idea that language is acquired through practice after having paid attention to language, made the effort to understand it and produced it. In the end, all the three perspectives of language acquisition aim at explaining the same concept but from different positions, which agree on the fact that it is the complex functioning of the brain in its different ways which works on the information provided and reacts to it.

With reference to the different theories, Long (1981), in his *Interaction Hypothesis* (also called the *Oral Interaction Hypothesis* by Ellis (1991), the *Interaction Theory* by Carroll (1999), and the *Input, Interaction, and Output Model* by Block (2003)), stated that interaction was the necessary condition for SLA. As mentioned before, for interaction to occur, input is key. Likewise, in Krashen's (1981) *Input Hypothesis*, input is the major causative factor for SLA. As for the

---

1. The Language Acquisition Device was proposed by Chomsky in the 1960s. It is a system of principles that children are born with that helps them learn language, and accounts for the order in which children learn structures, and the mistakes they make as they learn. (<https://www.teachingenglish.org.uk/article/language-acquisition-device>)

*Universal Grammar Model*, (Chomsky, 1964), although nativists believe that any speaker has an innate set of principles common to all human languages, input is “the evidence out of which the learner constructs knowledge of language” (Cook, 1991, p. 117).

Some of the models proposed for SLA which include the relevant role of input are the ones by Gass and Selinker (1994) and Ellis (1997). Figure 1 represents Gass and Selinker’s (1994) model. In this model there are five main stages: *apperceived input*, *comprehended input*, *intake*, *integration* and *output*. The first stage is called *apperceived input* and it is defined as the cognitive act “of understanding by which newly observed qualities of an object are related to past experiences. In other words, past experiences relate to the selection of what might be called noticed material” (Gass and Selinker, 1994, p. 482). Factors such as frequency, affect, prior knowledge and attention may be the reason why some aspects are noticed and some others are not (Gass and Selinker, 1994). Some researchers (Chapelle, 1998; Ellis, 1994; Lai et al., 2008) compared apperception with the concept of noticing explained in Schmidt’s (1990) *Noticing Hypothesis* (see Section 1.3.5). The second stage is *comprehended input*<sup>2</sup>, that is to say, the input apperceived in the previous stage and which is controlled by the learner after having done the work to understand it. The third stage is *intake*, and it refers to the mental activity “of assimilating linguistic material” (Gass and Selinker, 1994, p. 486). That is to say, what has been used to form the internal memory of L2. The fourth stage is *integration* and refers to the development of one’s second language and storage. *Output* is the last step and constitutes the “active part of the entire learning process [...] and represents more than the product of language knowledge” (Gass & Selinker, 1994, p. 490).

---

2. A different concept from *comprehensible input*, which is controlled by the person providing input.

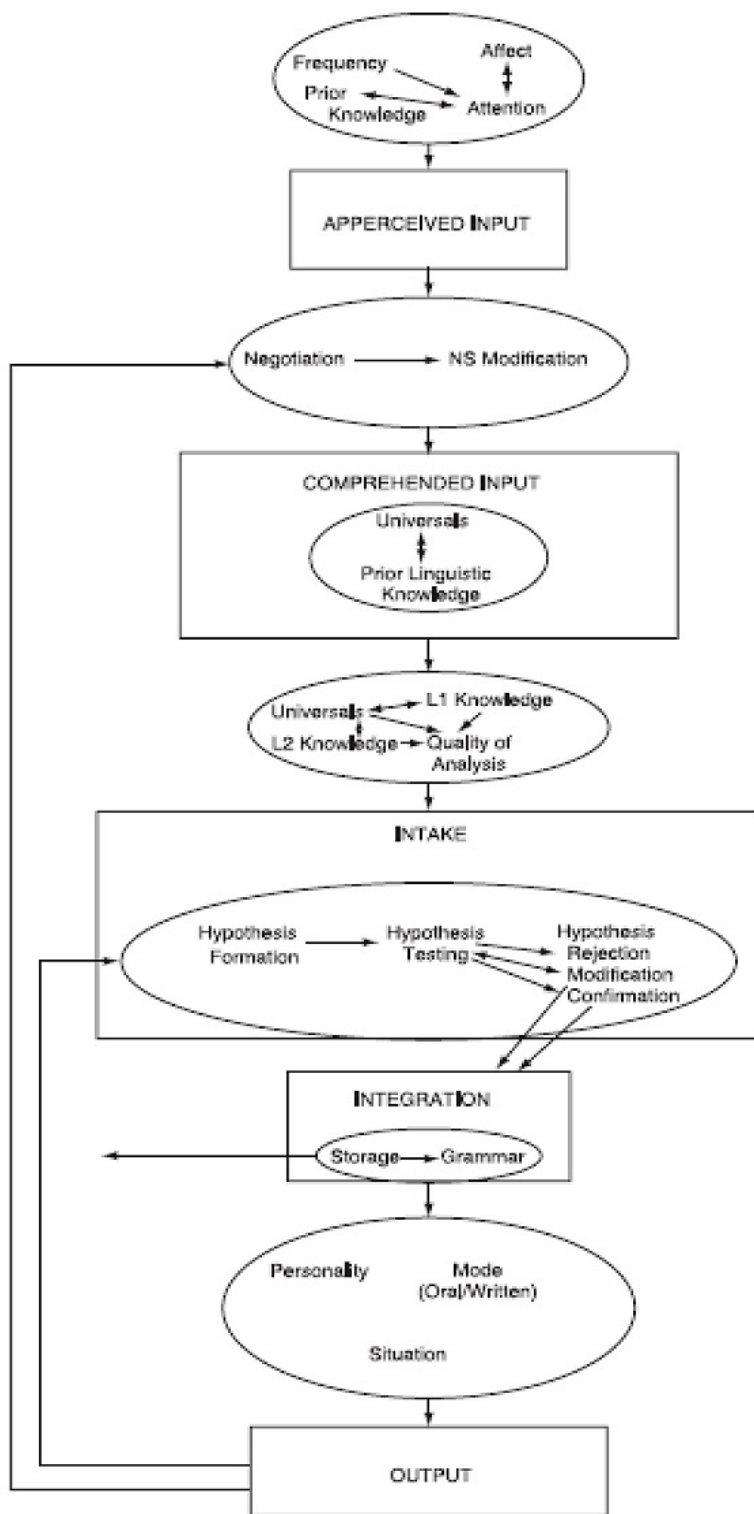


Figure 1. A model of second language acquisition  
(Gass & Selinker, 2008, p. 481)

The second model is the one by Ellis (1997), as presented in Figure 2. In this model, the learner is exposed to input, and part of this input is taken into short-term memory, what is known as

*intake*. In turn, “some of the intake is stored in long-term memory as L2 knowledge” (p. 35). Both intake and L2 knowledge take place in the ‘black box’<sup>3</sup> of the learner’s mind. Lastly, L2 knowledge is responsible for the production of spoken and written output.

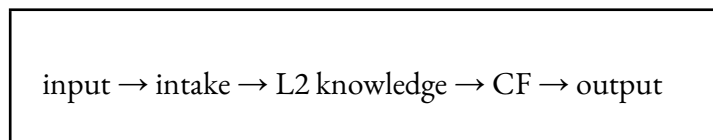


Figure 2. A computational model of L2 acquisition (adapted from Ellis, 1997, p. 35)

As we can notice in both models, the line between input and output is at least long. Only by presenting language, will it not become output. For input to become output different stages are necessary. That is why both models emphasise the importance of input as an initial stage but differ in the number of stages that lead to SLA. On the whole, input is seen as the primary condition in most approaches to SLA (Long, 1982; Grady et al., 2011; VanPatten & Benati, 2010; Pica et al., 1987; VanPatten & Williams, 2007); however, it is not so evident that only input leads to SLA. Thus, only by considering all the aspects and stages of SLA, will we be able to understand this process. For this reason, in the next section we are going to look further into the issue of input in language learning.

### 1.1.1 Comprehensible input and language learning

In an attempt to explain how a second language is acquired, Krashen formulated the Input Hypothesis, which assumes that the language input received is “the only causative variable in SLA” (Krashen, 1981, p. 57). Krashen goes on further to propose the concept of *comprehensible input* and explains its process leading to SLA as follows:

*We progress along the natural order (hypothesis 2) by understanding input that contains structures at our next ‘stage’-structures that are a bit beyond our current level of competence. (We move from  $i$ , our current level, to  $i+1$ , the next level along the natural order, by understanding input containing  $i+1$ ) (Krashen, 1985, p. 2)*

3. The Black Box was proposed by Chomsky (1964). He claimed it existed somewhere in the brain and contained all the principles which are universal to all human languages. (<https://www.kau.edu.sa/Files/0008718/Subjects/Chapter%202.pdf>)

In other words, language which contains structures that a learner already knows has no purpose if acquisition is the aim. On the contrary, if that language contains structures that are beyond the learner's actual knowledge, it leads to acquisition. Then, it can be inferred that speaking is the result of acquisition as it is created through comprehensible input.

As mentioned above, Krashen (1985) refers to comprehensible input. However, Gass (1988, 1997) claims that it is important to talk about comprehended input more than comprehensible input because learners can understand most or all the input they receive but only part of it is involved in the SLA process. That means that not all the language input needs to be  $i+1$  to be noticed, but sometimes prior knowledge or frequency make it possible too (Gass, 1997).

Criticisms of Krashen's *Input Hypothesis* are twofold: on the one hand, the way comprehensible input is addressed, and on the other, the  $i+1$  formulation. McLaughlin (1987) and Cook (1993) claimed that Krashen fails to define both  $i+1$  and comprehensible input. It is very difficult to determine each learner's level and then provide him/her with  $i+1$  since homogeneity in a classroom does not often occur. Hence, contrary to Krashen's assumption of comprehensible input, other researchers that support the *Input Hypothesis* (e.g., Ellis & He, 1999; Gass & Varonis, 1994; Long, 1982) state the need to include modified input, interactionally modified input, and modified output as part of the comprehensible input needed for SLA. That is to say, for input to become SLA, teachers need to adjust their output to the level of their students to facilitate comprehension (modified input), negotiate meaning and repair discourse to arrive at a mutual understanding and modify their own interaction to clarify meaning, for example.

Gregg (1984) also attacks Krashen's assumption that comprehensible input is "the only causative variable in SLA" (Krashen, 1981, p. 57) by stating that:

*There is no a priori reason to assume that a learner systematically ignores his own utterances. If output is available as input, and if Monitoring can increase the incidence of correct utterances of a given structure, then it would seem that output is being used to further acquisition, and thus that the Monitor can be used for acquisition (Gregg, 1984, p. 88).*

In the same vein, Swain (1985), in her *Output Hypothesis*, suggests that it is through language production that SLA is more likely to occur. Every time a learner produces language and notices a gap, modification of the same comes later. This modification may bring about new learning. Studies in the 1980s conducted by Pica, Young and Doughty (1987) and later by Romeo (2000) support Swain's *Output Hypothesis*. Output is necessary for teachers to check their students' level and improvement and for learners to face mistakes. In other words, both comprehensible input and comprehensible output are necessary elements for SLA to occur.

As Ellis (2002) states, exposure is necessary for learners to recognize linguistic features that have been previously repeated. "We may not be counting words as we listen or speak, but each time we process one there is a reduction in processing time that marks this practice increment, and thus the perceptual and motor systems become tuned by the experience of a particular language" (2002, p. 152). For this reason, the more input learners receive, the more easily they will recognize those linguistic structures. However, it does not mean that exposure alone is sufficient, but explicit instruction is sometimes needed, especially to speed the learning process (Ellis, 2002). It is important to mention that from this perspective, learners are the ones who deduce regularities for the input received instead of being imposed by UG.

When learners receive input, not all of it is processed. In Gass and Selinker's (1994) and Ellis' (1997) SLA models, both the terms input and intake are mentioned. As they are sometimes confused and cannot be interchanged because they have different qualities, they are going to be explained in the next section.

### **1.1.2 Input and intake**

After having reflected on the advantages of comprehensible input and the criticism of Krashen's *Input Hypothesis*, it is important to distinguish between two concepts: *input* and *intake*. In the 60s, Corder made the distinction between these two terms. As stated in Section 1.1, *input* refers to "potentially processible language data which is made available, by chance or by design, to the language learner" (Sharwood Smith, 1993, p. 167). On the contrary, *intake* refers to the input learners internalise, or "the information that can subsequently

be used for acquisition” (Truscott & Sharwood Smith, 2011, p. 498). That means that not all language exposure becomes intake; therefore, for input to be effective it must comply with some requirements. Input must be comprehensible, interesting and/or relevant, and sufficient (Krashen, 1982).

According to psychological findings (Carroll, 1990), input must be *comprehensible* because if the learner’s brain cannot filter the information provided, it cannot become part of intake, let alone part of long-term memory (Krashen, 1978). Moreover, the incomprehensible part may frustrate the learner when making the effort to decode the information and, therefore, hinder acquisition. It is the non-native speaker’s (NNS) ability which should indicate somehow (e.g., body language, backchannel cues, etc.) whether the information is being understood. In addition, if the input is not challenging or completely known, language development may also be missing. In this sense, according to Krashen (1982), input must be *interesting* and *authentic* in order to motivate students. Most textbooks are designed to pass an exam and include similar exam-based content leading to little creation, peer collaboration and use of authentic materials. For example, books try to reflect grammar in conversation by including full forms. However, in real life speakers tend to use short forms (e.g., *Do you like dancing?* vs. *Like dancing?*), that is why sometimes students sound too formal or even bookish. Finally, input must be *sufficient*. If learners are supposed to improve their language level, they need to be exposed to the language and receive the necessary quantity of input. It is important to take into account that learning a language in one’s own country is more difficult than in the target language (TL, henceforth) country, since in the first case the student’s exposure to the language is lower than in the second one.

Intake has also been long investigated as a component of the language learning process (e.g., Chaudron, 1985; Gass, 1997; Reinders, 2012; Schmidt, 2010). Some authors (e.g., Sato & Jacobs, 1992; Sharwood-Smith, 1993; Van Patten, 2002; Ying, 1995) have broadly defined intake as the part of input that is processed. In Schmidt’s (1994) words, intake is “that part of the input that the learner notices” (p. 139), or, according to Reinders (2012), intake is referred to as “a subset of the detected input (comprehended or not), held in short-term memory, from which connections with long-term memory are potentially created or



strengthened” (p. 28). From these definitions we may claim that, for input to become intake, there exists a continuum that starts from perception, passing through recoding and encoding in the long-term memory and finishing with integration or incorporation. Consequently, both input and noticing are key elements for intake to occur.

As a conclusion, for input to be effective and profitable, it must somehow be adapted so that learners can comprehend it if the objective is SLA. Gass and Selinker (1994) emphasise that “without understanding the language no learning can take place. Although understanding alone does not guarantee that learning will take place, it does set the scene for learning to take place” (p. 200).

After having discussed the role of input in SLA, we will now turn to consider a second key issue in language learning, that is, output.

## **1.2 Output in the Language Learning Process**

As explained in Section 1.1.1, although input is thought to be the only means for SLA and output is not influential at all (Krashen, 1981, 1994, 1998), others (e.g., Swain, 1985) argue that apart from comprehensible input, learners need to produce language so that they can move forward in their L2 knowledge. Output (i.e., language production) is an essential element that “may stimulate learners to move from the semantic, open-ended non-deterministic, strategic processing prevalent in comprehension to the complete grammatical processing needed for accurate production.” (Swain, 1995, p. 128). Output ensures mental language processing and represents the opportunity “to express a particular meaning by retrieving a particular form or structure.” (Benati, 2017, p. 3). Although the notion of output was first viewed as the production of language that had been already learnt (Gass & Selinker, 2001), after Swain’s Output Hypothesis in 1985 the idea of output as part of the learning process started to carry great weight. It became a useful tool to test the learners’ knowledge, to attempt new forms, to notice the gaps they encounter when attempting to produce the target language and to elicit new input.

Swain's *Comprehensible Output Hypothesis* appeared in response to Krashen's *Comprehensible Input Hypothesis*. Swain explains that

*simply getting one's message across can and does occur with grammatically deviant forms and sociolinguistically inappropriate language. Negotiating meaning needs to incorporate the notion of being pushed toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately (Swain, 1985, p. 249).*

Swain admits that output is not the sole element that leads to SLA, but she emphasises its relevant role. The need to deliver a message makes learners realise whether they can express what and how they wish to do so. If they cannot, then they may modify their output and learn something new about the target language. This process helps learners in the following ways: by providing opportunities for meaningful practice, by making them switch from semantic mental processes to syntactic ones, by testing hypotheses and withdrawing their own conclusions and, finally, by having the opportunity to receive feedback on their output (Swain, 1993).

Hence, taking into account that “the output hypothesis claims that the act of producing language (speaking and writing) constitutes, under certain circumstances, part of the process of second language learning” (Swain, 2005, p. 471), it is of utmost importance to describe the functions output can develop. Swain (1985, 1995, 2005), in her *Output Hypothesis*, describes three functions of output:

- a. **Noticing Function.** When learners produce language and have problems expressing themselves, they notice their linguistic gaps and, therefore, try to guess ways to do it or even ask for help. In this way, output helps learners to activate their cognitive mechanisms that foster SLA.
- b. **Hypothesis Testing Function.** By producing output, learners test the target language so as to confirm or modify structures according to the learners' interaction and the feedback received.
- c. **Metalinguistic Function.** This function is a way of “building knowledge about language” (Swain, 1995, p. 478). When linguistic problems arise, learners try to solve them by means of different strategies aimed at negotiating and reflecting about the forms and rules of the TL.

Some years later, based on Swain's hypothesis, Gass (2001) referred to four functions of output in L2 learning:

- a. *Feedback*. Interactional feedback provides learners with relevant information about the success or lack of success of their utterances. It also gives additional opportunities to focus on production or comprehension.
- b. *Hypothesis testing*. Interaction helps learners to test hypotheses. That is to say, when learners produce language, and negotiation and feedback occur, they are consciously or unconsciously checking language hypotheses.
- c. *Automaticity*. While some processes are more challenging as they require more time and memory capacity, some others are less demanding due to the fact that they happen automatically. Therefore, the repetitive practice of grammar results in automatic processing. This automaticity may facilitate learning since attentional resources are focused on other aspects of the language. That is to say, if a student is better able at fluency and syntax, more attention can be paid to other fields such as pragmatics or sociolinguistics. This automaticity contributes to facilitating fluency since students process information with little effort and quickly without having to think about every single step they take. It cannot be stated that automatization itself accounts for SLA, but it helps perform more accurately and consistently.
- d. *Meaning-based to grammar-based processing*. Output and feedback help learners to notice any mismatches or deficiencies in their production. That entails reassessment or longer-term complex thinking, which is the essence of the process of learning. Thus, processing language both at the level of meaning and syntax will help learners in the production of language.

The role of output has been investigated by a number of studies which have examined whether output contributes to SLA. The effectiveness of output to prompt SLA seems to make important progress after studies conducted by Izumi (2002), Izumi et al. (1999), Nobuyoshi and Ellis (1993) and Wang and Castro (2010). Nobuyoshi and Ellis' study, for example, "provides some support for the claim that 'pushing' learners to improve the accuracy of their production results not only in immediate improved performance but also in gains in accuracy over time"

(p. 208). Six adult EFL low-level students took part in a jigsaw where the use of past tenses was necessary. Three of the participants, the experimental subjects, received requests for clarification when they made mistakes in the use of past tenses, whereas the other three, the comparison subjects, received general requests for clarification when the teacher did not understand. After two sessions, clear gains in accuracy on the part of the experimental subjects were shown.

Izumi, Bigelow, Fujiwara and Fearnow (1999), in an attempt to confirm whether output promoted noticing and resulted in improvement of the target form, conducted a study where participants were divided into two groups: the experimental group, who received output opportunities and subsequent exposure to input, and the control group, who was exposed to the same input and had to answer comprehension questions instead of producing any output. Results show that both groups increased significantly in their noticing of the target form; however, the experimental group improved in the production test.

Izumi's (2002) study analysed whether output helped students notice formal elements in the TL and the subsequent learning of the form. In order to conduct this study, Izumi used a computer-assisted reconstruction and reading task as materials to present the target input. The results show that the students engaged in output-input activities outperformed the group which was exposed to the same input for the sole purpose of comprehension in learning gains. In addition, those who received visual input enhancement did not show improvement in language learning.

Paninos (2005) studied the role of L2 output for orienting L2 learners' attention to form in L2 input in SLA and she found a significant positive effect for output on orientation of attention to form in input and on the production of grammatical forms. Likewise, Morgan-Short and Bowden (2006) investigated the effects of meaningful input and output-based practice on SLA with a group of 45 Spanish students. Results showed that the experimental groups outperformed the control groups on immediate and delayed interpretation and production tasks. In light of these results, the authors suggest that not only input-based but also output-based instruction may lead to linguistic development.

Similar results were obtained in Song and Suh's (2008) study with fifty-two adult Korean EFL learners and which investigated the efficacy of two different kinds of output tasks (a

reconstruction task and a picture-cued writing task) in noticing and learning a type of English conditional. Students who were given opportunities to produce output showed greater noticing of the target form than non-output groups. Likewise, Wang and Castro's (2010) study with L1 Chinese learners of English as a foreign language suggests that classroom interaction and output result in noticing the target form under analysis and in learning gains. Recently, Zalbidea (2020), in an attempt to study the extent to which producing L2 impacts learner-generated noticing and L2 development of grammatical structures, showed that output led to greater noticing as well as L2 grammar development compared to no output.

Although the above studies seem to support the fact that output contributes to SLA, there are others that consider the possibility of acquisition without output. Early studies such as the one described in Fourcin (1975) confirm that literacy competence can be developed from input alone, without any language production at all. This is the case of Richard Boydell, who despite suffering from cerebral palsy, developed very high levels of competence from input alone (listening and reading).

Ellis (1995), for instance, compared a group that did no speaking at all (premodified group), with a group that interacted with a native speaker. The premodified group outperformed the group that interacted with a native speaker in vocabulary. Similarly, Izumi and Izumi (2004), in a study on the acquisition of relative clauses with twenty-four adult ESL learners divided into three groups (output group, nonoutput group, and control group) found that the output group did not outperform the non-output group. The authors report that "the output task failed to engage learners in the syntactic processing that is necessary to trigger L2 learning, while the task for the non-output group appeared to promote better form-meaning mapping" (p. 587).

There are also studies that show that acquisition of vocabulary and spelling are possible without the need of output. For example, Nagy, Herman, and Anderson (1985) conducted a study with different subjects and different comprehension abilities who were tested on their vocabulary gains. After a checklist and a reading period, the students completed a multiple choice vocabulary test. Results showed that all the groups improved in vocabulary, which suggests that reading triggers incidental vocabulary growth. These findings seem to support Krashen's (1989)

view that “the best hypothesis is that competence in spelling and vocabulary is most efficiently attained by comprehensible input in the form of reading, a position argued by others” (p. 440).

In sum, although research has shown that SLA can be achieved by means of input and/or output, by producing language learners may be obliged “to move from semantic processing to syntactic processing” (Swain, 1985, p. 249). In fact, as Swain and Lapkin (1995) suggest, “sometimes, under some conditions, output facilitates second language learning in ways that are different from, or enhance, those of input” (p. 371).

After having examined two key concepts in SLA, we will now move on to consider a third related element, that is, feedback.

### **1.3 Corrective Feedback in Language Learning**

There is general agreement that accuracy plays an important role in SLA and CF is the common instrument adopted by teachers to treat learners’ errors. However, there has also been avid debate on the role and use of CF from those who reject it (e.g., Krashen, 1985; Truscott, 1996) to those who see its potential in language acquisition (e.g., Dekeyser, 2003; Gass & Mackey, 2006). Despite countless research into the provision of feedback (e.g., Bitchener & Knoch, 2008; Black & William, 1998; Ellis, Loewen & Erlam, 2006; Ferris, 2002; Gibbs & Simpson, 2004; Hyland & Hyland, 2006; Jiang & Chen, 2013; Lyster, 2004; Van Beuningen, de Jong & Kuiken, 2012), there is still uncertainty about the effect that different CF strategies have on SLA.

One of the functions of output analysed previously is the hypothesis testing function (Swain, 1985), and it is here where CF comes into play. When students produce language, they test their hypotheses, and the feedback they receive is another way of testing. Yet, there are limitations to such assumptions. For instance, learners are not always aware they are being corrected or the feedback provided may not be specific enough. In addition, although feedback is “individualized attention that is otherwise rarely possible under normal classroom conditions” (Hyland & Hyland, 2006, p. xv), and it may affect students positively, it can be time-consuming for teachers and ensuring it is provided on a regular basis can be both difficult and tedious. In view of the many issues concerning corrective feedback, in the next section we address this term in more detail.

### 1.3.1 Definition of feedback

In the 1990s the concept of focus-on-form, an approach where learners' attention is attracted to linguistic forms, started to receive special attention and, as a consequence, special emphasis was also placed on CF. Whether to employ CF or not, as well as problems with regard to its use, have been the focus of vigorous debate for some decades now, for example, the ambiguity or even ineffectiveness of teachers' corrections (Lyster & Mori, 2006; Truscott, 1996).

First, what is called in general terms error correction has also been referred to as negative feedback (Ortega, 2009), negative evidence (Long, 1991), interactional feedback (e.g., Lyster & Mori, 2006), corrective feedback (Lyster, 1998) and negative input. These different terminologies are sometimes used as synonyms. However, a distinction is made between them, as in Saxton (2000), where negative evidence, negative feedback and negative input are considered different entities. Negative evidence refers to the corrective input provided after a student's error. On the contrary, negative feedback denotes a non-specific signal such as a clarification question about an error. The latter, negative input, is a more general term that denotes "any kind of adult response, contingent on child grammatical error, which embodies information conducive to the realignment of an overgeneralized grammar" (Saxton, 1997, p. 140). In light of the different definitions of CF, in this dissertation we will refer to CF as any teacher's reaction to a student's error.

An early definition of CF was provided by Chaudron (1977, p. 31), who referred to it as "any reaction of the teacher which clearly transforms, disapprovingly refers to, or demands improvement of the learner utterance". Good and Brophy (2000) defined it as a motivating tool for students to know how they are doing in the learning process. A few years later, Russell and Spada (2006, p. 135) defined it as "any feedback provided to a learner, from any source, that contains evidence of learner error of language form". Ellis (2006) described it as reactions to students' erroneous utterances. Although this term has been defined along the years in different ways, those definitions are similar in essence. CF can be provided at different stages but it can be most beneficial when the learner is cognitively prepared to receive it. Despite the variety of definitions, CF seems to be a useful tool for language learning by many authors

(e.g., Bitchener & Knoch, 2008; Bitchener, Young & Cameron, 2005; Ellis, Loewen & Erlam, 2006; Hattie & Timperley, 2007; Li, 2010; Sheen, 2006; Van Patten, 2003). However, the opposite view is defended, too. If CF is used incorrectly or very frequently, it can discourage language learning (Ayedh & Khaled, 2011; Ellis, 2009; Martínez, 2008; Storch, 2010). In light of opposing views, we turn to consider to what extent CF may be counterproductive or worthless.

On the one hand, Truscott (1996) claimed that “grammar correction has no place in writing courses and should be abandoned” (p. 328). In fact, he contended that error correction is harmful since it devoted time and energy to aspects that were less important. After analysing some studies by Fazio (2001), Kepner (1991), Semke (1984) and Sheppard (1992), this author found no convincing evidence that error correction was of any help to student’s writing accuracy. Those assumptions have become the basis of strong debate in subsequent articles (Ellis, 1998; Ferris, 2002, 2004; Ferris & Hedgcock, 2014; Truscott, 1999).

Another argument against corrective feedback is Dulay and Burt’s (1977) *Affective Filter Hypothesis*, later incorporated by Krashen as one of his five hypotheses in his Monitor Theory (the *Acquisition-Learning Hypothesis*, the *Monitor Hypothesis*, the *Natural Order Hypothesis*, the *Input hypothesis*, and the *Affective Filter Hypothesis*.) It suggests that CF can increase students’ anxiety and consequently increase this affective filter preventing them from acquiring language. When students have a high motivation, self-confidence and a low level of anxiety, they are more likely to succeed in second language acquisition. On the contrary, when students feel pressured to produce language or receive a composition with a lot of corrections, they are likely to become anxious and feel demotivated to continue producing language. For some researchers (Fazio, 2001; Kepner 1991; Polio et al., 1998; Sheppard, 1992) CF is discouraging and ineffective to improve subsequent writings. However, it must be noted that some of the studies (Polio et al., 1998; Robb et al., 1986; Sheppard, 1992) did not include a control group and, therefore, they lacked any comparison with students who received no feedback.

On the other hand, there are numerous studies that report improved accuracy after receiving CF (e.g., Ashwell, 2000; Bitchener & Ferris, 2012; Ellis, 2009; Ferris & Roberts,



2001; Sheen, 2007, 2010; Van Beuningen, De Jong & Kuiken, 2012). They claim that students' accuracy improves when they attend to feedback as they draw their attention to linguistic inconsistencies and/or errors. Fathman and Whalley (1990), for example, found that students who received feedback made fewer grammatical errors; however, this study did not examine new pieces of writing but text revisions which did not prove improvement over time.

In sum, despite mixed findings of studies showing more or less improvement, CF plays an important role to facilitate L2 development when students commit errors. Although the term 'error' is usually employed as a synonym of 'mistake' in language learning literature, there is a difference between them which is going to be tackled in the next section.

### 1.3.2 Error and mistake

When dealing with errors, there has been a huge change of perspective between the period before and after the 1960s. Before the 1960s and under the behaviourist dominance, errors were seen as an undesirable product of an inconsistent methodology and, therefore, should not take place. Some scholars, such as Gass and Selinker (2008), consider errors should be avoided and eliminated. However, with the concept of UG suggested by Chomsky in 1965, the idea of seeking a perfect methodology started to lose impact and errors started to be regarded as part of the learning process and, therefore, as a representation of the learner's progress (cognitivist school). It is at this moment when the term *error* acquires a different view and becomes an interest of research. In 1967, Corder started analysing errors as a source of information of the strategies used during the learning process.

Regardless of these differences and the debate whether or not a mother tongue and a second language are learnt in the same way (Corder, 1967, 1971, 1974), making errors during the learning process provides evidence of the language that the learner has learned. According to Corder (1967, p. 167):

Errors are significant in three different ways. First to the teacher, in that they tell him, if he undertakes a systematic analysis, how far towards the goal the learner has progressed and, consequently, what remains for him to learn. Second, they provide to the researcher evidence

of how language is learned or acquired, what strategies or procedures the learner is employing in his discovery of the language. Thirdly (and in a sense this is their most important aspect) they are indispensable to the learner himself, because we can regard the making of errors as a device the learner uses in order to learn. It is a way the learner has to test his hypotheses about the nature of the language he is learning.

Therefore, taking into account their nature, it is important to make a difference between *error* and *mistake*. We can find several definitions of *error* over the years. Lennon (1991) described it as “a linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers’ native speaker counterparts” (p. 182). James (1998) defined it as “an unsuccessful bit of language” (p. 1).

Committing errors is a frequent situation but should not be seen as something negative. When learners make an error, teachers receive useful information about the learner’s knowledge and the forms which still have not been acquired. According to Stenson (1974), learners may commit an error because the target form has not been acquired, the learning/teaching situation may be demanding or because they are common language performance errors. Whatever its nature, learners tend to receive corrective feedback, which offers them the possibility to notice the gap and make a new attempt. On the contrary, *mistakes* are slips or random ungrammaticalities that can be self-corrected. They are mismatches in the speakers’ language caused by external factors such as distraction or boredom despite having the knowledge of the correct linguistic form and being familiar with the rule.

Errors and mistakes are different in nature but sometimes it is not so easy to make a difference between them. According to James (1998), errors cannot be self-corrected, whereas mistakes can if the person who makes them is aware of the erroneous form. However, the fact of not self-correcting makes researchers wonder whether it is a mistake or an error. In this case, two options are possible to distinguish an error from a mistake (Ellis, 1997). The first one is to check the frequency of the incorrect utterance. If a student says ‘She love animals’ twice and then ‘My mother speaks Italian’, it is difficult to determine if it is an error or a mistake. However, if this student keeps omitting the -s third person singular, it reveals it is

an error. The second option is asking the learner to correct the utterance. If he/she comes up with the correct form, it is a mistake; if he/she cannot, it is an error.

In short, taking into account the previous definitions, errors are systematic deviations of their linguistic knowledge that the learners themselves cannot correct because of their underlying competence. On many occasions, errors are the language transfer's outcome (Scovel, 2001). Unlike errors, mistakes are unsystematic and identifiable flubs that can be easily and immediately corrected on the part of the learner since they are not the cause of incomplete knowledge.

Errors are part of the learning process, and their analysis and classification may reveal interesting information which may be helpful to understand the SLA process. This fact motivated researchers to study learners' errors, the so-called Error Analysis (EA) approach. We turn to consider this topic in the next section.

### **1.3.3 Error analysis**

Many teachers do not understand why learners keep repeating the same errors and are unable to implement the input provided in the classroom. In order to provide an answer, Lengo (1995) stated that it was due to the teacher's misconception that output should be an authentic copy of input. If all the input were acquired, then, intake would play no role. As explained in Section 1.1.2, not all input is internalised. That is why, by means of analysing the errors learners make, valuable insights may be gained in order to better understand language acquisition. For this reason, the Contrastive Analysis Hypothesis (CAH) became so predominant in the 1950s when dealing with learners' errors.

Contrastive Analysis (CA) is based on the comparison of two languages to identify their structural differences and similarities and be able to predict the linguistic difficulties experienced during the learning process. Although learning a second language is different from learning the mother tongue, when learning a second language, the rules of the native language (NT) inevitably influence the rules of the target language (TL). Due to the differences in both systems, interferences may arise.

The CAH exists in strong and weak versions (Wardhaugh, 1970). The strong version refers to the contrast between “one language—the grammar, phonology and lexicon—with the system of a second language in order to predict those difficulties which a speaker of the second language will have in learning the first language and to construct teaching materials to help him learn that language” (p. 4). The drawback of this version is that it only explains interlingual errors as they are the ones caused by transfer from L1 to L2.

On the contrary, the weak version only requires the use of the best linguistic knowledge available in order to explain the observed difficulties in second language learning (Wardhaugh, 1970). Instead of being based on prediction, it is based on the evidence resulting from interference to explain the similarities and differences between the two languages. Whereas the strong version is based on predictive power, the weak version is based on a posteriori observation to explain the linguistic difficulties. Therefore, the strong version focuses on negative transfer as the main cause of L2 errors whereas the weak version denies interference from the L1 defending the idea that the L1 is not an inhibitor of language learning but it is the lack of knowledge that may cause errors.

A third version in between the weak and strong versions was proposed by Oller and Ziahosseiny (1970), the moderate version. It is based on both predicting and interpreting errors, even those caused by overgeneralization, and regards L1-L2 differences as something positive rather than interfering and easier to learn since students do not generalise a specific response to one stimulus. In addition, it can explain both interlingual errors related to the NL and intralingual errors related to the TL. Thus, whereas the strong version states that the greatest difficulty is where more differences between NL and TL exist, the moderate version suggests the opposite.

Even though the moderate version seemed to be more effective, CAH fails in reaching its target as it suffers from both under prediction and over prediction. According to the CAH, difficulties in learning the TL are due to language interference but there are many other factors which can be the cause (e.g., aptitude, motivation, lack of concentration, age, etc.) In sum, CA is partially ineffective in explaining students' failure as it only focuses on the linguistic aspect leaving others such as the psychological one aside.

Due to the weaknesses of CA, EA gained ground in view of the unsuccessful attempt of CA to predict learner errors and explain why some TL features are more difficult to acquire than others (Keshavarz, 1999). As explained above, CA compared a pair of languages (i.e., the learner's L1 and the TL) and by analysing their similarities and differences tried to predict the errors learners were likely to commit. However, EA examined both the impact of transfer errors as well as those related to the target language, including overgeneralization (Schackne, 2002). From that moment on, recognising the origin of the learners' errors has gained more and more relevance.

According to Richards et al. (1992), the study of errors could contribute to “identify strategies which learners use in language learning, try to identify the causes of learner errors and obtain information on common difficulties in language learning, as an aid to teaching or in the preparation of teaching materials” (p. 184). In this sense, EA set out in the 1960s to show that learner errors were not only caused by their mother tongue, but they could also reflect universal learning strategies. One of the earliest studies on EA was Richards (1971), who tried to explain the reasons why people who learned an L2 did not speak and write it in a native way. In his study, Richards analysed errors relating to production and distribution of verb groups, prepositions, articles, and the use of questions by learners from different language backgrounds (Japanese, Chinese, Burmese, French, Czech, Polish, Tagalog, Maori, Maltese, and Indian and West African Languages) and distinguished three different sources of error: interference, intralingual and developmental errors. Some years later, (Richards, 1974) reduced these three categories to two: interlingual and intralingual errors.

Corder, in turn, classified errors into four categories: omission of some required element; addition of some unnecessary or incorrect element; selection of an incorrect element; and misordering of the elements (1973, as cited in Erdogan, 2005). Corder also referred to overt and covert errors. Overt errors are ungrammatical inconsistencies at the sentence level whereas covert errors are grammatically correct forms but inconsistent at the discourse level.

Burt (1975) also classified errors into global or local. Global errors are those “that significantly hinder communication [...] that affect overall sentence organisation” (p. 6),

whereas local errors “affect single elements (constituents) in a sentence [and] do not usually hinder communication significantly” (p. 7).

According to their source, errors can be due to *interlingual transfer* and *intralingual transfer* (Richards, 1974). The former refers to language transfer caused by mother tongue interference. This transfer can be positive or negative. When the structure is similar in both languages and the transfer may be somehow justified, it is called positive transfer. However, if the transfer is not justified because the structure is different in the two languages, it is called negative transfer (Wilkins, 1972). In turn, intralingual errors occur during the learning process due to “lack of knowledge or difficulty” (Keshavarz, 2003, p. 62). Richards (1985) divided intralingual errors into the following categories:

- a. overgeneralizations: errors caused by extension of target language rules to inappropriate contexts,
- b. simplifications: errors resulting from learners producing simpler linguistic rules than those found in the target language,
- c. developmental errors: errors reflecting natural stages of development,
- d. communication-based errors: errors resulting from strategies of communication,
- e. induced errors: errors resulting from transfer of training,
- f. errors of avoidance: errors resulting from failure to use certain target language structures because they are thought to be too difficult,
- g. errors of overproduction: structures being used too frequently.

Some other researchers, taking into account that it is sometimes not easy to distinguish between interlingual and intralingual errors, developed different categories. For example, Dulay and Burt (1974) referred to developmental errors (those similar to L1 acquisition), interference errors (mother tongue errors) and unique errors (those different from developmental and interference). In turn, Stenson (1974) included induced errors (those due to incorrect instruction). Some years later, Brown (2007) classified errors into four categories: interference transfer (mother tongue errors), intralingual transfer (those similar to L1 acquisition), context of learning (those errors induced by the teacher or the textbook), and communi-

cation strategies (strategies for communicating when linguistic forms are not available to the learner). Some years later, James (1998) classified errors into interlingual errors (mother tongue errors), intralingual errors (caused by the target language: false analogy, overgeneralization, exploiting redundancy, etc.), communication strategy-based errors (holistic strategies or approximation and analytic strategies or circumlocution), and induced errors (caused by the teaching and learning process).

The nature and cause of learners' errors have aroused a lot of interest and the literature on this topic is extensive. Whereas at the beginning most errors were thought to be caused by the L1 transfer, over time, researchers found that L1 influence on L2 was quite insignificant (Sattayatham & Honsa, 2007). In order to shed more light on interlingual and intralingual errors, both views are going to be illustrated through some studies over the last decades.

As for interlingual studies, Ying (1987), for instance, investigated the relationship between students' L1 (Taiwanese) and their EFL writings. From the errors they made, 13.6% were overgeneralization, 7.5% simplification and 78.9% language transfer errors. Kim (1989) cited in Lee (2001) also analysed the influence of L1 using the translation of a set of sentences into English by two hundred Korean EFL learners. Transfer errors were slightly higher (24%) than overgeneralization (23%), for example. Likewise, Jiang (1995) investigated errors in English prepositions by Taiwanese learners. Results showed that the fact that the use of prepositions in Mandarin and English is different influenced L2 production negatively. That is why a great deal of errors are derived from language transfer. Articles and tenses were also used in a different way in Mandarin since there was no equivalent element for articles and no inflections in verbs.

Another study conducted by Alonso-Alonso (1997) analysed the writings of 28 Spanish EFL elementary students aged fourteen and fifteen to identify the most common interlingual errors. Results revealed that the most common type was transfer of structure (deviant form when mother tongue rules are followed). Orthographic, syntactic and semantic similarity to the L1 were also a source of concern for Spanish students. On the contrary, vocabulary and grammar seemed not to be so. Similarly, Horney (1998, as cited in Chen, 2006) analysed the compositions written by 80 Taiwanese EFL students and the findings showed that L1

transfer errors accounted for the highest percentage: articles were the most common error followed by tenses and prepositions.

In the 21st century, studies on the relevance of EA and L1 transfer have continued in the same vein. Some students commented on the difficulty of mastering the use of articles because the L1 and the L2 do not share an equivalent system (e.g., Chen, 2000; Ionin & Montrul, 2010; Romaine, 2003). Lee's (2001) study aimed at identifying and classifying errors committed by 35 intermediate to advanced Korean students in their English essays. Results showed that 26% of the errors were rooted in L1 transfer followed by wrong words (16%), prepositions (15%), and articles (14%). Similarly, Hsin's (2003) study revealed that the different syntactic presentations for indefinite subject noun phrases between English and Chinese resulted in learning difficulties for Chinese students when producing EFL writings.

Farsi has also been a language studied from the point of view of EA. Koosha and Jafarpour (2006) tried to "determine the extent to which Iranian EFL learners' knowledge of collocation of prepositions is affected by their L1" (p. 192). 200 English senior university students underwent two completion tests on collocation of prepositions. Results illustrate the influence of the participants' L1 in their collocation patterns of prepositions in English. Likewise, in Khodabandeh's (2007) research, 58 graduate students of English were asked to translate 30 headlines from Persian to English. The findings supported the idea that native language interference was the most immediately noticeable source of error. Similarly, Ridha's (2012) study, in an attempt to examine the nature of errors of 80 EFL Arabian students' writings, showed that most of the students' errors (mainly grammar and mechanical) were due to L1 transfer because students relied on their mother tongue to express themselves.

Ibáñez-Moreno (2011) analysed the texts produced by 53 secondary and high school Spanish students. Results revealed that L1 interference in all linguistic levels was the most recurrent error. Still, the secondary school students committed fewer transfer errors than the high school group. In an attempt to explain the findings, this researcher claimed that the younger students "are still working with the TL and are trying to communicate with its linguistic elements, not through the L1" (p. 16). In contrast, the study conducted by Ciesielkiewicz and Márquez (2015) with Spanish students of 1st and 2nd year of high school



showed quite balanced results. The authors found that, in the students' writings in English, 44% of the total errors in both groups were caused by L1 transfer.

Finally, Murtiana's (2019) study analysed the occurrence of interlingual and intralingual errors by undergraduate EFL Indonesian learners through their writings. The findings demonstrated that although both types of errors were committed, "first language interference caused more errors in writing than learners' incomplete process of acquiring second language rules" (p. 204).

The above studies suggest that, regardless of the L1, all learners are likely to make a transfer from their mother tongue to their second or foreign language. However, there are also opposing views, claiming that the influence of the L1 on the L2 seems to be quite minimal (e.g., Bataineh, 2005; Sattayatham & Honsa, 2007). Already in 1971, Richards challenged the fact that the L1 was the cause of learners' errors by stating that "many of the learners' errors came from the strategies that they use in language acquisition and the reciprocal interference of the target language items" (p. 208, cited in Heydari & Bagheri, 2012). Thus, it is worth exploring research on the opposite perspective, that is, the role of intralingual errors.

H. Kim (1987, cited in Lee, 2001), in his study on 12<sup>th</sup> grade Korean English compositions, observed that intralingual errors exceeded transfer errors. In another study one year later, I. Kim (1988) investigated errors in English verbs with reference to tense, mood, and voice. By analysing the translations of the 120 EFL Korean subjects, the author found that overgeneralization errors (65%) occurred the most while L1 transfer (22%) and simplification (13%) occurred the least. In 2001, S. Kim conducted another study with Korean university students and analysed the errors of 30 written essays to classify them into interlingual or intralingual errors. Intralingual errors outnumbered by far L1 interference errors. The same was observed by Bataineh (2005), who after analysing nine types of errors in the use of the indefinite article of Jordanian EFL university students, discovered that all errors, except one, had no connection with their L1. Sattayatham and Honsa (2007) also demonstrated that mother tongue interference errors happened the least. The 237 participants translated some sentences from Thai into English and after analysing the type of errors, both interlanguage and intralanguage errors occurred but mother tongue interference errors obtained a low

percentage. Similarly, Ahmadvand (2008) analysed 40 pre-intermediate to intermediate Iranian EFL learners' errors in their written productions and among the different errors committed, omissions, additions, and regularizations were among the most frequent types of errors and L1 transfer represented only 30%.

Pastor and Selistean (2015) conducted a study with A1 and A2 EFL Spanish secondary students with the aim of analysing the students' writing errors (transfer or intralingual errors) to gain insights that may "help teachers plan future lessons, design class materials, and make decisions on correction techniques more attuned with learners' needs" (p. 120). The findings showed that intralingual errors (mainly morphological) occurred more frequently than transfer errors (syntax and lexis), which may demonstrate that students rely on their L1 in syntax and lexis much more than in morphology.

These studies, then, raise several concerns. First, if learners make more intralingual than interlingual errors, then, the assumption that L1 and L2 share some patterns in the way they are learnt gains support (i.e., children's L1 acquisition). Second, the higher the level of proficiency, the fewer interlingual errors are committed as they become intralingual. That is to say, the lower the level, the more the students need to reflect on the systematic rules of the TL and transfer errors are more likely to happen.

In conclusion, whether the errors are intralingual or interlingual, EA plays a key role in analysing learners' errors that can help both teachers and learners to know their causes and how to handle them. As Lodoño-Vásquez (2008) stated, "indisputably, error analysis is a fundamental and relevant tool in language teaching, in order to reorganise and transform the teacher's point of view and readdress his/her methodology, with the aim of fixing and filling the students' gaps" (p. 144). EA is still an appeal to teachers and researchers since "the fact of making errors is an important concern" (James, 1998, p. 120).

After having dealt with EA, the nature of errors and some error classifications, we will now review the way errors can be treated.

### 1.3.4 Error treatment

The correction of errors is a question that has always worried teachers and played an important role in past methods such as Grammar Translation and Audiolingualism. However, when the Communicative Approach emerged, error correction faded and meaning was emphasised over form (Harmer, 2001; Richards & Rodgers, 2001).

As mentioned in Section 1.2, the testing hypothesis function of output (Swain, 1985, 1995, 2005) serves to test out the hypothesis of the target language either consciously or unconsciously. That is to say, through interaction and negotiation learners notice gaps that they try to solve by testing target language forms until they come across the correct form. However, correcting errors can have negative effects on the learner (Horwitz et al., 1986; Truscott, 1996, 1998, 1999, 2007, 2009; Young, 1991) if we take into account Krashen's (1985) *Natural Order Hypothesis*. It explains that rules of language are acquired in a predictable order, which is independent of the order in which they are taught formally. As a result, correction may have negative emotional reactions which raise the student's 'affective filter', in Krashen's (1985) terms. In Walker's (1973) study, for instance, learners preferred not to be corrected for each error because it undermined their confidence and motivation, and interfered with the flow of conversation. If interruptions only happen when an error is corrected, the students' motivation may decrease and they may not be willing to take risks in future situations.

On the other hand, if error correction is nonexistent, false hypotheses can be created and mistakes can be fossilized. Hence, the right balance needs to be sought. There are numerous studies (e.g., Ashwell, 2000; Doughty & Varela, 1998; Iwashita, 2003; Long et al., 1998; Lyster, 2001; Panova & Lyster, 2002; Russell & Spada, 2006; van Beuningen, 2010; van Beuningen et al., 2008) that support the beneficial effects of both oral and written CF in SLA. As already mentioned in the previous section, despite the effectiveness of error correction, inaccurate or overcorrection can be a hindrance. For this reason, some errors such as the ones that are beyond the students' level should not be corrected as they are beyond their capacity.

According to Akay and Akbarov (2011), there are some important aspects that teachers need to take into account in error correction in the language classroom:

- a. Be aware of the goals of the lesson, and the students' levels. Positive gains are reported when there is a main objective. When the focus of correction is a specific goal, students tend to remember their inconsistencies better.
- b. Encourage self-correction. Giving learners the feeling of freedom and autonomy to correct their mistakes makes them feel responsible for their learning.
- c. Be aware of timing and how to correct. Some mistakes can be later included in sentences on the board and corrected by students.
- d. Do not waste time correcting mistakes. During the learning process mistakes are inevitable. Therefore, teachers should not waste time overcorrecting and repeating the correct form.

In relation to the above aspects, knowing who, when and how to treat errors is a very important task since if done incorrectly, students' anxiety and frustration can increase. The questions about who can provide CF, when, and how to provide it are going to be briefly reviewed now, since a thorough analysis will be presented in Chapter 2. As for the 'who', feedback can be provided by different entities. Students always have in mind the figure of the teacher as the provider of corrective feedback, but it is not always the case. Feedback can also be provided by a peer, by oneself, or by a computer. The teacher's feedback is a way of receiving individualized responses to one's production and it is part of the instruction process (Ferris & Roberts, 2001). In the case of written CF (WCF, henceforth), which is the focus of the present dissertation, it represents a facilitating tool in improving students' writing skill (Ferris & Hedgcock, 2014; Goldstein, 2004; Leki, 1990). However, taking into account the number of students in the classrooms and the workload (Tuck, 2012; Winstone & Carless, 2019), students sometimes report that CF is insufficient, confusing, discouraging and even at the wrong time (Evans, 2013; Nicol, 2010; Winstone et al., 2017). What is more, teachers complain that learners do not make use of the feedback provided (Orrella, 2006). Thus, to ensure the effectiveness of CF, both students and teachers need to engage in the process and

make the most of it (Nash & Winstone, 2017). Otherwise, as mentioned before, feelings such as frustration, anxiety and discontent by both parties may arise and interfere with the learning process negatively (Handley, Price & Millar, 2011).

Apart from the teacher's feedback, peers can also be a source of help. Peer feedback is the process by which learners engage in dialogue with the objective of improving language production. In addition, peer feedback can also have other benefits for both parties, such as developing objectivity of other's work that can later affect their own work positively and providing alternative strategies in an accessible language (Nicol & MacFarlane-Dick, 2006). For instance, some researchers such as Scharle and Szabo (2000) and Rollinson (2005) have a preference for student-student feedback to teacher feedback because students feel more comfortable with peers. The class becomes student-centred letting them be autonomous and responsible for their own learning. Peer-correction engages critical reading (Berg, 1999), fosters confidence (Byrd, 2003; Min, 2006) and diminishes the anxiety level as it is a way of supporting each other (Mangelsdorf & Schlumberger, 1992; Schmid, 1999; Stanley, 1992). In Diab's (2010) study, those students who engaged in peer-editing were able to reduce their language errors more than those who were given the possibility to self-edit their essays.

On the contrary, peer feedback can have some risks and has in fact been questioned by both teachers and students. Some students may not like to be corrected by peers as they think they are not proficient enough to do so (Allaei & Connor, 1990; Diab, 2010), they may not trust each other's comments (Carson & Nelson, 1996) and may even feel inferior (Harmer, 2004).

Another source of feedback is the students themselves. Self-assessment is not an easy job since the learner has to distance himself/herself from his/her composition, try to adopt an objective stance and criticise it. Judging one's own work allows learners to monitor their progress and improve future work (Rolheiser & Ross, 2000). It is a way of engaging and motivating students in the learning process by reducing dependence on the teacher (Munns & Woodward, 2006; Schunk, 1996). Despite these benefits, self-assessment can also be troublesome if learners are unrealistically optimistic by "believing they are above average, [...] by overestimating their ability to bring about personally desirable events, [...] by

overestimating how easily they can complete tasks and by overestimating the chances that their decisions about the present are sound and that their predictions about the future will prove correct” (Dunning, Heath & Suls, 2004, pp. 72-73).

The last source of CF, which is gaining importance over the years, is computer-generated feedback. With the advent of new technologies and their constant development, language classrooms have adapted to the times with tools such as whiteboards, beamers, computers, etc. However, computer-generated feedback is not a widespread practice at schools. Computer-generated feedback, commonly known as *Automated Writing Evaluation* or AWE (Warschauer & Ware, 2006), is feedback provided by a program which can also score the writing. AWE is a way of saving time, especially with large groups, providing feedback immediately and encouraging revision. As occurs with teacher, peer and self-assessment, computer-generated feedback can work in combination with other types of feedback and it is not meant to substitute the teacher. Despite its benefits, AWE reliability is still a debatable topic. Whereas some studies report positive results (Coniam, 2009; Hutchison, 2007), others reveal unfavourable or mixed results (Lai, 2010; Lee et al., 2009; Tuzi, 2004). A deeper analysis of computer-generated feedback will be provided in Chapter 3.

In regards to when CF should be provided, feedback can be immediate (at an early stage of a task performance) or delayed (some time after having completed the task). Chaudron (1977) used the terms delayed and postponed correction. Delayed refers to the correction provided once the speaker has finished the sentence and postponed refers to the correction provided in the future. According to some researchers (e.g., Hunter, 2007; Quinn, 2014), as little time as possible should pass between error and feedback. Doughty and Williams (1998) refer to it as the ‘window of opportunity’, a short slot of time (around 40 seconds) in which the speaker still holds the output in his/her memory and the input received. For this reason, Doughty (2001, p. 257) explains that:

*... one of the most promising kinds of intervention is an immediate contingent recast, which can easily fit into working memory along with the original utterance to which it is to be compared. Results of recasting studies suggest that such cognitive comparison does lead to forms-function-meaning mapping and, hence, can be considered a successful means to promote processing for language learning.*

In this way, as explained in the quotation above, learners undergo a cognitive process which gives them the chance to compare their linguistic attempt and the correct form. However, the opposite view is also supported by some authors (e.g., Ellis, 2015; Ortega, 2009). If the inter-language system is developing and errors are inevitable, by means of input and practice some errors will disappear progressively. Then, delayed feedback should be provided and only on target remaining errors.

Finally, as for how errors should be treated, Ellis et al. (2006) refer to two ways of providing feedback: explicitly and implicitly. Although both kinds of feedback will be analysed in detail in Section 2.4, a brief description of explicit and implicit feedback is now convenient for the sake of clarity. Whereas explicit feedback indicates that there is an error, in implicit feedback there is no indicator. As an illustration, Lyster and Ranta (1997) proposed four different types of implicit feedback: elicitation, recasts, repetition and clarification requests, and two types of explicit feedback: explicit correction and metalinguistic feedback. Sometimes explicit and implicit feedback can also be combined.

Apparently, the studies showing that explicit feedback results in greater gains than implicit feedback outnumber those that prove evidence on the effectiveness of implicit correction over explicit correction (e.g., Ellis et al., 2006; Havranek & Cesnik, 2003; Lyster, 2004; Muranoi, 2000; Zohrani & Ehsani, 2014). However, as some other research proves (e.g., Azizollah, 2008), it is likely that the way errors are addressed may not just depend on which type of CF is more beneficial but factors such as the teacher's methodology, students' proficiency level, oral or written mode, students' preferences, etc. may also play a role.

In conclusion, Corder, already in 1967, noted the importance of errors in language learning and their analysis as a way to infer L2 learners' strategies in the L2 acquisition process. It is important to treat errors so that they do not fossilise and feedback should be provided taking into account different aspects such as nature of the error, timing, explicitness, learners' affective filter, etc. to make the most of it. Treating errors either by the teacher, the peers, oneself or the computer is not a random task but a thorough one, as will be explained in Section 2.7. What is more, as Leeman (2003) and Doughty and Varela (1998) concluded,

corrective feedback has to be sufficiently explicit to be noticed by learners if any effect in the classroom is desired. This will be tackled in the next section.

### 1.3.5 Noticing and focus-on-form approach

The dominant theories over the past decades were those which advocated the unconscious processes of learning. However, in the 1980s Schmidt started to question these assumptions since after a personal experience in Brazil, he realised “adults do seem to have lost the still mysterious ability of children to acquire the grammatical forms of language while apparently not paying attention to them” (Schmidt, 1983, p. 172). During a five-month stay in Brazil, Schmidt and Frota (1986) conducted a study in which they realised that although some forms were frequently repeated in the input, only until some kind of focus-on-form (FonF) happened, did they acquire them. What is more, despite being corrected when speaking, those corrections were not effective since they were unnoticed. This fact might mean that some level of conscious attention to linguistic features is required if the aim is language learning.

It was after Schmidt and Frota’s (1986) study that the term *noticing* started to gain importance in the SLA process, culminating in 1990 when Schmidt elaborated his *Noticing Hypothesis*. This hypothesis emphasises the importance of noticing the target features while receiving input as a key condition for language learning (Schmidt, 1990, 1993, 1995, 2001, 2010). According to Schmidt (1990), “the subjective experience of ‘noticing’ is the necessary and sufficient condition for the conversion of input to intake” (p. 209). Nevertheless, the fact of noticing the target feature does not mean it has been understood. For this reason, Schmidt attempted to categorise different levels of awareness: perception, noticing and understanding. As for perception, it is “not necessarily conscious, and subliminal perception is possible” (1990, p. 132); in turn, noticing refers to the focal awareness or *apperceived input* (Gass, 1988) already mentioned in Section 1.1. Focal awareness is more than simple perception, it is the verbal report that is used to verify or falsify claims concerning the role of noticing in cognition. Finally, understanding is the mental activity by which we analyse and compare what is noticed to what has been noticed in previous occasions.



Authors like Gass (1991, p. 136) also claim that “nothing in the target language is available for intake into a language learner’s existing system unless it is consciously noticed”. That is to say, some level of noticing is necessary if acquisition is the goal. As Ellis (1994) explains, noticing is essential in order to conduct a cognitive comparison. Some researchers have also suggested that students need an explicit focus on language to facilitate acquisition (Ellis, 2001; Norris & Ortega, 2000). This led researchers to reflect on the effects of language focus in instruction. For example, Leow’s (1997, 2000) studies show that those learners with higher levels of awareness (understanding) improved the most and those learners unaware that they were being corrected showed no improvement. How can learners become aware of L2 forms when receiving input? Mackey (2006) found out that when students received some kind of feedback, it triggered noticing, which contributed to language development. However, Izumi’s (2002) research showed that learners who received enhanced input showed more noticing but not more learning. Thus, although it is true that previous research lends support to implicit learning, the idea that implicit learning is the default type of learning in both children and adults (Reber & Allen, 2000) seems to be contradictory to Mackey’s (2006) and Izumi’s (2002) conclusions.

Schachter’s (1998, p. 547) words below claim that not all learning requires attention:

*... although I (among many others) am perfectly willing to agree that learning individual words (the lexicon), individual sounds (the phonetic inventory), and writing systems must be via attentional focus, I am not the least willing to say that learning phonological, morphological and syntactic rules requires this attentional focus.*

In addition, DeKeyser (2005) argues that exposure alone is not sufficient to learn some structures or linguistic aspects when they are not salient enough. For this reason, he advocates grammar practice in the classroom. However, a study conducted by Perruchet and Pacton (2006) shows the opposite view: “research from the past few years has made it increasingly clear that participants in artificial grammar learning experiments do not need to extract the rules to perform well, even in situations involving transfer across surface forms” (p. 233).

Despite vigorous debate, probably the answer is not an emphatic yes or no since some forms may require more attention than others due to complexity or speed, for example.

Every learner has a different predisposition for language learning due to variables such as age, mental ability, knowledge of other languages, etc., and the level of noticing may be influenced by different factors. For instance, studies by Creel, Newport and Aslin (2004), Gebhart, Newport and Aslin (2009), Granena (2013), Hunt and Aslin (2010), Kaufman et al. (2010), and Saffran, Johnson, Aslin and Newport (1999) gather experimental evidence that adults can learn aspects such as distributional patterns and tone implicitly. However, regarding the EFL context, teachers play the dominant role and are the ones who must set a good example. For this reason, they are expected to raise learners' awareness to promote ways in which the input delivered becomes intake (Wong, 2005). In addition, with the emergence of the communicative teaching method and the importance of fluency and communication over form, the student's attention needs to be drawn to some linguistic elements.

According to Schmidt (2001) and other authors (e.g., Leow, 2000; Robinson, 2003), there is general agreement that noticing contributes to language learning as it draws students' attention to linguistic elements while receiving input, helping, thus, to transform this input into intake (Schmidt, 1990, 2001). Although there is a wide range of definitions, they agree on its essence:

*Focus on form overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication (Long, 1991, p. 45-46)*

*Form-focused instruction is any pedagogical effort which is used to draw the learners' attention to language form either implicitly or explicitly (Spada, 1997, p. 73)*

*Focus on form involves an occasional shift in attention to linguistic code features—by the teacher and/or one or more students—triggered by perceived problems with comprehension or production (Long & Robinson, 1998, p. 23)*

Focus on form in the classroom is an approach controlled by the teacher who tries to direct the students' attention to form (e.g., a repetitive mistake) with the aim of correcting and avoiding it in the future. It is of paramount importance to push learners to create target-like language (Doughty, 2003; Doughty & Williams, 1998; Long & Robinson, 1998). However, the way to do it depends to a great extent on the teachers' beliefs about how a language is

learnt and taught (Williams & Burden, 1997). Yet, those beliefs are not static but change in the course of time and according to the context (Barcelos & Kalaja, 2003).

In the same vein, not only Schmidt (2001) but also other researchers (e.g., Sharwood-Smith, 1993; Swain, 1985, 1995) believe that the input learners receive is more beneficial when their attention is drawn to it. For this reason, Schmidt makes reference to Tomlin and Villa's (1994) attentional functions, a work that has been inspired by the research of Posner and Petersen (1990) and Posner and Rothbart (1992) in neuroscience. Tomlin and Villa's model of attention includes three stages: *alertness*, *orientation* and *detection*. Learners need to be ready and *alert* to receive feedback and motivated to discover and correct mistakes. Otherwise, they may not process it in the same way. The *orientation* stage is in charge of guiding the attentional resources toward "some type or class of sensory information at the exclusion of others" (Tomlin & Villa, 1994, p. 191). This stage helps the next one to be completed, *detection*. It is "the process that selects, or engages, a particular and specific bit of information" (p. 192). That is to say, it transforms input into intake and is the most important stage of all three since neither of the previous attentional processes are a condition for detection to occur.

The three attentional functions are separable and are not a prerequisite for the others to occur. In addition, attention and awareness can be detached from one another. Tomlin and Villa's (1994) work has been key in subsequent studies dealing with input enhancement (e.g., Alanen, 1995; Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995; White, 1996, 1998) and has contributed to the reflection of attentional processes in SLA; however, among researchers there is no consensus of what Tomlin and Villa claimed. Leow (1998) implemented the analysis of attention in a SLA study which found empirical support for Tomlin and Villa's (1994) model, especially for the attention and detection stages. However, some authors such as Simard and Wong (2001) pose a word of caution when referring to that model and thus suggest that "a model of attention that more accurately reflects the complex nature of SLA is one in which awareness and attentional functions are viewed as being present in graded amounts, and whose degree of activation is influenced by the interactions among task

type, linguistic items, individual differences (such as processing capacity), and by any other concurrent cognitive activity competing for processing resources” (p. 118).

In this line, transferring Tomlin and Villa’s (1994) model of attention from the neuroscience research to the SLA field does not seem an easy task since “designing a task that could adequately examine the isolated effects of alertness and orientation during detection of L2 input seems virtually impossible” (p. 110). It would be idealistic to expect a single theory to apply to both fields, neuroscience and SLA, due to the wide array of variables to take into account, such as type of task, targeted linguistic item, individual differences, etc.

Schmidt (2001) explains that for learning to occur, learners need to attend to the input they receive. Thus, some kind of attention is necessary, but it is still a mystery the amount and type of attention needed as well as the way the learner’s attentional processes are influenced. Those reflections pave the path to two main approaches which focus on the creation of input to raise learners’ awareness of the forms to be acquired: *visual input enhancement* and *output*. In both cases their aim is to make the learner aware of possible mismatches. What makes them different is that in visual input enhancement attention is drawn by external methods such as highlighting, bolding or underlining, among others, and in output it is the student him/herself who notices the mismatch when producing language. There are many studies which are based on this attempt to highlight some specific forms and analyse the students’ reactions and progress (e.g., Ellis, 1994; Long, Inagaki & Ortega, 1998; Long & Robinson, 1998; Lyster, 1994; Lyster & Ranta, 1997; Mackey & Philp, 1998; van Patten, 1990, 1994; van Patten & Cadierno, 1993). However, there is still much research to be conducted on how, how much and “whether there are more propitious pedagogical moments to draw learners’ attention to language form” (Spada, 1997, p. 80). In Tomasello and Herron’s (1989) study, providing corrective feedback after a communicative task proved to be effective. On the contrary, other studies (e.g., Doughty, 2003; Williams, 2005) show the opposite case.

In conclusion, taking into account the different views on the role of noticing, it seems it plays a role in L2 acquisition. Schmidt’s *Noticing Hypothesis* has been an important contribution and has been widely supported by, for example, Batstone (1994), Ellis (1994, 1997), Gass (1988), Lewis (1993), Lynch (2001), Skehan (1998), among others, but it has also

been criticised by some authors who defend the idea that a) attention and awareness should be dissociated (Tomlin & Villa, 1994), b) noticing can be only necessary for the acquisition of metalinguistic knowledge (Truscott, 1998), c) noticing may depend on many factors apart from attentional elements such as input saliency, prior knowledge or task difficulty (Philp, 2003), and d) not all what is apperceived becomes intake automatically (Gass, 1997). Despite this disparity in opinion, noticing plays a key role in SLA, especially in the productive skill of writing, a topic we turn to in the next section of this chapter.

#### **1.4 L2 Acquisition of the Writing Skill**

Writing is a communicative process in which there is a sender who delivers some ideas in the written form and a receiver who reads them. Learning to write is difficult not only for language learners but for native speakers, too (Esmeralda, 2013). As attested by Barkaoui (2007) and Daud et al. (2016), the writing skill may be the most difficult one to be acquired by L2 learners as they are expected to master a variety of linguistic, cognitive, and sociocultural competencies (e.g., Ellis, 2015; de Oliveira & Silva, 2016; di Gennaro, 2016).

Writing entails a complex system of processes and sub-processes in which learners go back and forth within the text while implementing the different macro- (e.g., planning, drafting, revising) and micro-strategies (e.g., content, form, orthography, lexicon, syntax) (Cumming, 2001; Lee, et al., 2016). According to Hyland (2011), learning to write involves five kinds of knowledge: content (i.e., ideas and concepts in the topic area), system (i.e., syntax, lexis, formal conventions), process (i.e., preparation), genre (i.e., communicative purposes of the genre), and context (i.e., reader's expectations, cultural preferences, related texts).

Manchón (2011) refers to two general dimensions of L2 writing to better understand L2 written language learning. On the one hand, the learning-to-write dimension and, on the other hand, the writing-to-learn content and writing-to-learn language. The first one refers to “the manner in which L2 users learn to express themselves in writing” (p. 3), whereas the latter refers to the way in which the engagement with L2 writing tasks and activities can contribute to development in areas other than writing itself (p. 3), that is to say, writing

as a vehicle for language practice and development of communicative language proficiency (Shrum & Glisan, 2010).

Writing development refers to language progress in terms of complexity, accuracy, fluency, etc. (Norris & Manchón, 2012). It is multifaceted in nature as it is ruled by individual and social elements such as beliefs, goals, learner histories, etc. Teachers spend a lot of time correcting their students' written pieces but they may still remain poor (e.g., grammar errors, unnatural language, lack of language variety, etc.). It may be caused by the little time being devoted to the writing skill at school (Taggart, 2009) or the lack of learner involvement in their own learning, which hinders writing development.

Writing in FL instruction is sometimes viewed as a tool to induce pushed output since new grammar structures, vocabulary, expressions, etc. have to be produced (Brown, 2007; Edward & Willis, 2005). In fact, "composition writing elicits attention to form-meaning that may prompt learners to refine their linguistic expression—and hence their control over their linguistic knowledge—so that it is more representative of their thoughts and of standard usage" (Cumming, 1990, p. 483). The problem-solving inherent nature of writing can promote proficiency, as learners' control over their linguistic knowledge increases (Cumming, 1990; Swain & Lapkin, 1995).

As for the teacher's role in the writing process, we find him/her as a guide, modeller, supporter, corrector, motivator, etc. In this regard, the teacher's role is to guide students through the writing process, avoiding an emphasis on form to help them develop strategies for generating, drafting, and refining ideas (Hyland, 2011). Some years earlier, this author had argued that one of the main teacher's tasks was to raise learners' awareness about L2 writing processes. He also noted that "effective writing instruction involves guiding students to an awareness of their readers, and the interactional strategies, background understandings and rhetorical conventions these readers are likely to expect" (Hyland, 2002, p. 83).

L2 acquisition of the writing skill is not just a cognitive matter but also a motivational one (Hyland, 2002). Motivating students to write and teaching in a supportive atmosphere are positive factors for language development to occur. It is of vital importance that students consider writing as something positive and teachers should praise learners for their effort and

success (Williams, 2003). Studies which date back to the 60s already proved that by praising students and believing in them performance improved (e.g., Ashton-Warner, 1963; Carini, 1982; Curwin, 1992; Howard, 1990). Hence, attending to the students' IDs and boosting the students' motivation can contribute to the students' engagement in the writing process (Dörnyei, 2001).

Other factors such as allowing students to take part in the choice of topic, for example, presenting reachable goals and challenging but manageable tasks also contribute to learners' motivation (Cumming, 2002; Ferris & Hedgcock, 2014; Hyland, 2002; Raimes, 1998; Williams, 2003). In addition, another aspect that has been mentioned to contribute to the students' motivation is learner autonomy (e.g., Dörnyei, 2001; Foster, 1996; Hout, 2002; Myles, 2002; Ross, Rolheiser & Hogaboam-Gray, 1999). When students review their own compositions, they have to keep their distance and adopt a reflective and critical view. By engaging learners in self-assessment, they become conscious of their strengths and weaknesses and can check if they match the target (Huot, 2002; Myers, 2001).

In sum, writing is a complex process where the two dimensions (i.e., learning to write and writing to learn content, and writing to learn language) may overlap as they share some kind of mutual dependence. Thus, questions such as whether they should be taught in a symbiotic manner or separately are still a debate.

## **1.5 L2 Acquisition of Grammar**

Over the years, a hotly debated issue has centred on whether the L1 and the L2 are acquired in the same way. It is undeniable that L1 and L2 language acquisition processes differ in a number of aspects. When learning our mother tongue(s), we not only learn a language but knowledge about life and the world whereas when we learn a second/foreign language, students use the knowledge of the world as the means to language learning. Thus, L1 learners are constantly exposed to the language whereas L2 learners may not be so exposed. As a child, becoming a native speaker of the L1 is something expected, whereas learning an L2 may depend on a wide range of factors such as age, individual aptitude and skills, exposure to the

language, etc. Despite these differences, both L1 and L2 learners have to cope with the same difficulty, to produce meaningful language (Lieven & Tomasello, 2008).

According to Ellis (2008), the linguistic knowledge a speaker has can be referred to as implicit and explicit. On the one hand, implicit knowledge refers to the automatic and unconscious acquisition of knowledge and can only be acquired up to a certain age. On the other hand, explicit knowledge refers to the conscious acquisition of a set of rules and can be learnt at any age. The goal of teachers is implicit knowledge as they want their students to be natural and fluent when producing language and explicit knowledge can support implicit language development (Ellis, 2005). Understanding the set of rules can help notice the structures of the target grammar more easily, which contributes to implicit language development (Schmidt, 1990). In addition, higher noticing of the structures can lead to higher comprehensible input too, which, at the same time, facilitates form processing (VanPatten, 2004).

In the case of children, Chomsky's (1965) theory suggests they are born with what he termed 'Universal Grammar' and 'Language Acquisition Device' (LAD), two tools that, along with the input received, develop the L1. Some researchers (e.g., Krashen, 1981, 1982; Schwartz, 1993) see the acquisition of L1 and L2 as a similar process where language instruction or explicit feedback have no place in SLA. In fact, innatists compare the lack of correction of parents to children in the L1 with the learning of the L2, where they consider corrective feedback is unnecessary (Chomsky, 1980; Pinker, 1996). If it is not necessary in the L1, why should it be in the L2? The debate grows when despite being exposed to large amounts of input, acquisition does not occur because the input received is not understandable (van Lier, 1996). As a consequence, Krashen, after having created his *Input Hypothesis*, proposes his *Comprehensible Input Theory*, which explains that the input received must be comprehensible to learners but go beyond the learners' current stage of development. Nevertheless, sharing the view of the existence of UG does not imply rejecting negative feedback. In White's (1991) words, "L2 learners sometimes make incorrect generalizations (in many cases based on the mother tongue) that cannot be disconfirmed by positive evidence alone [...] negative evidence may play a more significant role than is the case in L1 acquisition" (p. 134). Indeed,



this author concluded that sometimes input alone is not enough and negative feedback is required to make students aware of inaccurate forms (White, 2003).

Learning a language involves different stages that complement each other. After being exposed to input, learners need to be provided with opportunities for interaction and meaning negotiation contributing, thus, to language acquisition. However, the input received makes no sense if learners do not pay attention to it.

While researchers such as Krashen (1982) state that abundant comprehensible input may result in second language fluency, current SLA theory contends that as L1 and L2 acquisition processes are not identical, abundant comprehensible input is not sufficient for linguistic accuracy (Polio, 2010). In the same vein, Van Beuningen (2010) holds that for an L2 speaker to develop proficiency at any level, attention to linguistic form is needed so that learners can move towards accuracy in their L2. As a result, all contemporary communicative methodologies incorporate some form of grammar instruction, if not, L2 acquisition could be expected to be slower, more difficult, and less successful (Doughty, 2003). With reference to grammar instruction, a distinction between grammatical knowledge and grammatical ability has to be made. While grammatical knowledge refers to knowledge of the rules to produce correct language at the sentence level, grammatical ability refers to the ability to use grammar at the text level in order to communicate orally or in the written form (Jones, 2012). Grammatical ability is context dependent and despite having a high understanding of grammar rules and having practised them, correct grammatical usage is not ensured. Although traditional approaches drilled grammar explicitly and saw it as an independent component, current communicative approaches take grammar as a component of communicative ability (Richards & Reppen, 2014).

In the next sections we turn our attention to three linguistic aspects which have proved to be harder to learn and are the focus of the feedback provided in the study of the present dissertation; these are a) articles, b) prepositions and c) simple past-perfect tenses difference.

### 1.5.1 The acquisition of English articles

Learning a language is not an easy task and some linguistic elements may be more difficult than others to acquire. It is the case of the English article system, which seems to be quite laborious to learn (Master, 1987). Articles are unstressed elements that may go unnoticed in the input and have different functions at the same time, which makes one-to-one form and meaning impossible (Master, 2002). The English and Spanish article systems share similarities and differences.

The Spanish article system has definite (*el, los* -masculine- and *la, las* -feminine) and indefinite articles (*un, unos* -masculine- and *una, unas* -feminine), like English. Nouns in Spanish are preceded by an article that agrees in gender and number. The indefinite article is used to mention something for the first time whereas the definite article is used to refer to something that has already been mentioned, days of the week and times, when talking about a group of mixed gender, and with names of rivers, seas and oceans, mountain ranges and archipelagos as well as unique and well-known referents in the speech community. Unlike English, Spanish uses the definite article to talk about things in general. Consider the following examples:

1. Tengo **un** piano nuevo. (indefinite, masculine and singular)
2. **La** entrada fue muy cara. (definite, feminine and singular)
3. **Los** bancos siempre ganan. (definite, masculine and plural–general)

Both the English and the Spanish article systems use definiteness to distribute the articles whereas other languages like Samoan are based on specificity (Ionin et al., 2004). Chesterman (1991) described the article continuum as follows:

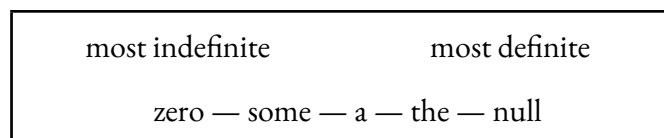


Figure 3. Article continuum (Chesterman, 1991, p. 182)

This continuum includes the indefinite article *a/an*, the definite article *the*, the zero and the null articles. The article *the* occurs with definite singular and plural count nouns and mass nouns as there is shared knowledge between the speaker and the hearer; the indefinite article *a/an* occurs with indefinite singular count nouns as this knowledge is only held by the speaker; the zero article is frequently used with indefinite noncount (e.g., water) and plural count nouns (e.g., cars); the null article is generally used with bounded singular proper nouns (e.g., Spain), and certain singular count nouns (e.g., “They always have lunch at 1”).

Due to the similarities and disparity between both article systems, L2 learners of English are likely to commit different sorts of errors when they learn the English article system. After extensive research on the acquisition of articles (e.g., Avery & Radišić, 2007; Butler, 2002; Ekiert, 2004; García-Mayo, 2009; Goad & White, 2004, 2009; Hawkins et al., 2006; Ionin, 2003; Ionin & Montrul, 2010; Ionin et al., 2004, 2008; Jarvis, 2002; Snape et al., 2009; Tryzna, 2009; White, 2008; Zdorenko & Paradis, 2007, 2008), it seems errors mainly include article omission and article misuse or substitution. However, there is no consensus on the reason why they are committed. Avery and Radišić (2007) suggest three main causes: L1 influence, UG and L2 influence. In turn, Master (2002, p. 1-2) claims the difficulty stems from three facts about the article system:

- a. articles [...] are among the most frequently occurring function words in the language, making continuous conscious rule application difficult over an extended stretch of discourse;
- b. function words are normally unstressed and consequently very difficult if not impossible for a NNS to discern, thus affecting the availability of input in the spoken mode;
- c. the article system stacks multiple functions onto a single morpheme, a considerable burden for the learner, who generally looks for a one-form-one-function correspondence in navigating the labyrinth of any human language until the advanced stages of acquisition. Moreover, as Master (1988) claimed, L1 influence

can be added as a fourth difficulty since “English article usage, especially at the beginning levels, is clearly influenced by the first language” (p. 2).

To explain the origin of such difficulties Ionin (2003) proposes the *Article Choice Parameter* (ACP) and the *Fluctuation Hypothesis* (Ionin et al., 2004). The ACP states that article-based languages are based on specificity or definiteness. For example, English, French, and Arabic are languages that belong to different families (Germanic, Romance and Semitic, respectively) but whose article system is based on definiteness. However, Samoan and Turkish are languages that are based on specificity. As for the *Fluctuation Hypothesis*, “L2 learners have full access to UG principles and parameter-settings [and] fluctuate between different parameter-settings until the input leads them to set the parameter to the appropriate value” (Ionin et al. 2004, p. 16).

The ACP predicts that L2 learners whose L1 has an article system (+ART) will transfer the properties of L1 to L2. It is the case of EFL Spanish learners. Ionin et al.’s (2008) study shows that EFL Spanish learners are able to use articles accurately even from the beginning because they do not need to fluctuate between both parameters, as both languages are similar in their article systems.

Contrarily, for learners whose L1 does not have an article system (-ART) and the L2 does have it, the ACP predicts there will not be any transfer and learners will fluctuate between definiteness and specificity until the L2 input helps them to distinguish both parameters. It may be the case of Russian and Korean students learning English, where there is no L1 transfer since both languages are articleless. For instance, Ko and Wexler’s (2004) study with intermediate and advanced Korean and Russian EFL learners found that the errors this type of students committed were mainly the overuse of *the* with indefinites and *a* with definites, which confirms the fluctuation predictions. Fluctuation can continue until advanced levels of proficiency but it will depend on the learners’ accuracy to continue with it and even to misuse articles. What the ACP and *Fluctuation Hypothesis* suggest is that errors on article choice in L2 English follow some pattern (Ionin et al. 2004). However, authors like Tryzna (2009) are uncertain about this theory in lower levels since L2 English “article use is better characterised by variability rather than fluctuation” (p. 67). In her study, the advanced groups fluctuated

as predicted but intermediate learners failed to fluctuate and showed more variability in the article use. Thus, Tryzna (2009) suggests that “there is a developmental path in the acquisition of the English article system, initially characterized by widespread optionality of article use and finally constrained by the fluctuation between the two parameter settings” (p. 84).

In general, speakers from an L1 that contains an article system (e.g., Spanish) tend to overgeneralise *the*, what Huebner (1983) and Master (1997) term “the-flooding”. *A* seems to be acquired after *the* is somehow mastered (Chaudron & Parker, 1990; Huebner, 1983; Master, 1997; Parrish, 1987; Thomas, 1989). On the contrary, other authors (e.g., Liu & Gleason, 2002; Young, 1996) adopt the opposite stance. In view of such inconclusive findings, some studies are going to be discussed to illustrate this disparity.

For example, in Park’s (1996) study, French and German students, two languages with an article system, obtained better results than Japanese and Korean learners whose L1 does not contain an article system. Even so, the acquisition order of English articles was similar in both groups independent of the L1. In turn, Snape (2005) studied the acquisition of articles by intermediate and advanced Japanese and Spanish learners of English. Taking into account that Japanese lacks an article system and Spanish does not, results show that Spanish learners outperformed Japanese learners and fluctuated less between definiteness and specificity. Japanese learners overused *the* in singular and plural [-definite, +specific] contexts but not in the indefinite mass (non-count) noun contexts. Likewise, Ionic and Montrul (2010) examined the acquisition of plural noun phrases of Spanish and Korean English language learners. Results prove that Spanish participants “overaccepted the generic interpretation of English definite plurals to a greater extent than proficiency-matched speakers of Korean, an articleless language” (p. 877).

The same results were obtained by Reid et al. (2006, as cited in Le Bruyn, 2019), who tested 14 intermediate Japanese and 9 intermediate Spanish learners of English in the use of articles. The results showed that Spanish learners did not fluctuate, whereas Japanese learners fluctuated between definiteness and specificity by overusing *the* in [-definite, +specific] contexts and overusing *a* in [+definite, -specific] contexts.

Another study with intermediate Japanese and Greek learners of English conducted by Hawkins et al. (2006) showed the same fluctuation between definiteness and specificity by overusing the definite article *the* in [-definite, +specific] singular and plural count noun contexts. It was not the case of Greek learners whose L1 has an article system. Zdorenko and Paradis (2007) also conducted a study comparing child learners of English whose L1 either had or lacked an article system. Results are similar to previous studies, because the -article learners omitted articles in obligatory contexts more often than the +article group and overused the article *the* in [-definite, +specific] contexts.

Finally, García-Mayo (2008), Snape, García-Mayo and Gurel (2013) studied L1 transfer on the use of English generic noun phrases by Spanish (a language with article system), Turkish (a language with an indefinite article but no definite article), and Japanese (a language with no article system) learners. Results show that article choice was mostly determined by students' L1.

In conclusion, the acquisition of the English article system seems to be easier when the native language and the TL are similar, since positive transfer occurs (e.g., Ekiert, 2004; Park, 1996; Romaine, 2003; Snape et al., 2013). However, it seems to be more complex when the L1 lacks an article system since negative transfer may take place (e.g., Avery & Radišić, 2007; Chen, 2000; Ionin, Ko & Wexler, 2004; White, 2008; Zdorenko & Paradis, 2007, 2008, 2012).

### 1.5.2 The acquisition of English prepositions

Similar to the use of English articles above, the use of prepositions in English is another complex system to master, especially for learners whose L1 is non-Indo-European. According to Fang (2000), if we take a corpus of one million English words, one in ten is a preposition, that is, the English language has approximately 70 simple prepositions, many more than other languages (Koffi, 2010).

Prepositions express a relation between two entities, one represented by the prepositional complement and the other by another part of the sentence. Sometimes they are formed by prefixing a noun or an adjective such as *inside=in+side*, *along=a+long* and their meaning

can be inferred. However, some other times they are not so obvious (Rowe & Webb, 2000). English prepositions usually come before the noun, and in some languages such as Albanian, Hungarian and Finnish, prepositions are inflected. Hence, some of the causes for such complexity lie in the fact that prepositions behave differently in different languages (James, 2007; Jie, 2008), are ruled by the inconsistency of the English language itself, (i.e., the arbitrary semantic characterisation of some prepositions such as *over* and *above*) (Evans & Tyler, 2005), and learners may not recognize them in natural speech because they are frequently monosyllabic and unstressed (Lam, 2009). That is the reason why they cannot always be translated literally (Celce-Murcia & Larsen-Freeman, 1999). In fact, many prepositional errors stem from making assumptions of semantic equivalence between L1 and L2 (Lam, 2009). As a result, systematising English prepositions is “nearly impossible” (Catalan, 1996, p. 171).

Prepositional errors have been traditionally described as omission, addition and substitution (Dulay, Burt & Krashen, 1982). Omission refers to the absence of the preposition, addition refers to the unnecessary use of the preposition, and substitution refers to the use of a wrong preposition instead of the correct one. As explained before, prepositions express a relation between two entities. For this reason, elements such as verbs play an important role in the correct choice and use of prepositions (Celce-Murcia, 2001). Some errors may be attributed to interlingual transfer, as in the case of the present study where both languages (i.e., English and Spanish) have similar prepositional systems, or to intralingual transfer depending on the learners' proficiency level. Although there are some studies that attribute a minimal role to L1 interference in prepositional errors (e.g., Ahmadvand, 2008; Castro, 2013; Chua et al., 2015; Foo, 2007; Zurniati, 2018), there are others that see interlingual transfer as one of the main causes (e.g., Asma, 2010; Gass & Selinker, 1992; Gvarishvili, 2012; Jansson, 2006; Kirkgöz, 2010; Khodabandeh, 2007; Pindo, 2016; Suzanne, 2017; Tahaineh, 2010; Yildiz, 2015). In addition, there are also studies that attribute intralingual errors to poor teaching, inefficient presence in textbooks or little attention in the classroom (e.g., Kim, 2001; Pinto & Rex, 2006).

Kim (2001), for example, analysed the written work of 150 Upper Secondary Malay students with the aim of identifying common prepositional errors in English and their cause. Students were required to write an essay and complete some gap-filling tasks which showed that the subjects made the most number of errors in using the prepositions of time, place and direction in the category of wrong selection. The researcher suggested that an efficient method to teach English prepositions is basic to facilitate acquisition.

Similarly, Pinto and Rex (2006) analysed the use of the prepositions *por* and *para* of four English-speaking undergraduate groups of L2 Spanish at four stages. Results show that, in spite of explicit grammar explanations and continuous exposure to the target language, the learners' improvement during the four-year university program only amounts to 8%. The authors come to the conclusion that the way prepositions are presented in textbooks may not be in line with the difficulties learners have or with the prepositions' functions in discourse.

Two more examples of research on English prepositions are Hayashi (2008) and Bong (2010). Hayashi (2008) focused on the semantic relatedness of three English prepositions (*at*, *on*, *in*) in terms of prototypicality and L1 transfer by 129 Japanese learners of English and 40 native English speakers. The study included three experiments whose results revealed that participants showed strong preferences for spatial instances of the prepositions when writing a free sentence. Moreover, the author found that there was a relation between the increase of the learners' proficiency and the growth of lexical networks of the three prepositions differently.

Bong's (2010) study also investigated the acquisition of *in*, *on* and *at* by Japanese speaking learners and Korean speaking learners of English with the aim of shedding further light on the influence of prototypicality and L1 transfer. Results indicated some inconsistencies regarding prototypicality, since Bong claimed that prototypical senses were not necessarily easier to acquire than less prototypical ones. For instance, Korean learners performed better with some *at* instances assumed to be less prototypical than more prototypical ones. Thus, regardless of their L1, with increasing proficiency both groups "show a gradual but extremely moderate development of the uses of the three prepositions" (Bong, 2012, p. 153). It seems L1



transfer was not evident and no significant differences in the degree of difficulty in acquiring the prepositions *at*, *in* and *on* were found.

There have been a number of studies aiming to account for the acquisition and variability of prepositions (e.g., Bong, 2005, 2010, 2011, 2012; Hayashi, 2001, 2008; Yamaoka, 1995, 1996). These studies propose two main hypotheses to explain the hierarchy and variability in the acquisition of L2 prepositions: the *Prototypicality Hypothesis* proposed by Yamaoka (1995, 1996) and further developed by Hayashi (2001, 2008), and the *Economy-Driven Development Hypothesis*, suggested by Bong (2005, 2009, 2010, 2011, 2012, 2013). The latter assumes that acquisition of both L1 and L2 “undergo the feature selection-construction process in forming lexical items and setting or resetting parameters through testing hypotheses, [...] which in effect give rise to ‘parametric differences’ between languages” (Bong, 2013, p. 143). As for the Prototypicality Hypothesis, it refers to the categorisation in the semantics of prepositions. They are organised hierarchically according to their prototypical sense (e.g., locative, literal, abstract, etc.). Thus, prototypical senses are easier to acquire, whereas less prototypical senses are more difficult (Hayashi, 2008).

### **1.5.3 The acquisition of English simple and perfect tenses**

Research on tense/aspect acquisition is also a relevant field of interest (e.g., Bardovi-Harlig, 2000; Comajoan, 2006; Klein, 1994; Shirai, 2004), since noticing the difference and mastering both (tense and aspect) is not easy for EFL learners. As Halliday (1994) claimed, if learners do not have a clear notion of these linguistic issues, advancing towards higher levels may be complex.

When talking about verbs, it is important to make a difference between these two temporal categories encoded in the verb. Tense has been defined as the “grammaticalisation of location in time” (Comrie, 1985, p. 1), whereas aspect as the “grammaticalisation of expression of internal temporal constituency” (Comrie, 1985, p. 6). In other words, tense refers to the temporality of a situation, whereas aspect refers to duration or completedness of occurrence. Tense is used to tell the hearer whether the situation occurred in the past, in the present or if it will occur in the future. Aspect is used to show the speaker’s point of view. That

is to say, whether the occurrence is bounded or unbounded. The aspectual distinction has been traditionally classified into morphosyntactic categories: simple, progressive, perfect and perfect progressive, but the present research is only concerned with the simple and the perfect, more specifically, with the difference between the simple past and the present/past perfect. When a situation usually occurs, the simple aspect is employed (e.g., *Ronald smokes a lot*). However, when a situation occurs prior to a significant reference time, the perfect aspect is used (e.g., *Ronald has smoked a lot*).

The Aspect Hypothesis (AH), developed by Andersen and Shirai (1994, 1996), Bardovi-Harlig (1992, 2000), Robison (1995), Shirai (1991), and Shirai and Andersen (1995), is a theory used to explain the L2 development of tense-aspect morphology, a predominant concern in L2 language development. The AH seems to conclude that the initial stages of tense-aspect development are restricted by the different classes of lexical aspect: states, activities, accomplishments, and achievements. Therefore, it predicts that language learners use the perfective form first with telic<sup>4</sup> verbs (achievements and accomplishments) as they have inherent endpoints, and later with atelic events and states. As for the imperfective, it is first used with stative verbs and with activities, accomplishments, and achievements. Consequently, learners are expected to use fewer prototypical selections as they become more proficient (Ayoun & Salaberry, 2008).

In the case of Spanish, the distinction between perfective and imperfective is expressed by means of overt tense/aspect morphology on the verb and all the aspectual classes of verbs (states, activities, accomplishments and achievements) combine with preterite and imperfect. The preterite is understood as a bounded/perfective event (e.g., *María leyó el periódico*) whereas the imperfect is interpreted as unbounded/imperfective (e.g., *María leía el periódico*). On the contrary, English lacks a past tense equivalent to imperfect in Spanish. Thus, lexicalising is the way to mark it, instead of morphosyntactically, that is to say, by using alternatives such as *used to* and *would* or the progressive tense.

The English simple past is an absolute tense (Comrie, 1985) as the event time and the reference time are concurrent. That means the verb describes a complete action or a series of

---

4. Telic verbs describe an event with an inherent endpoint (e.g., *Sara walked home*) and atelic verbs describe an event without an inherent endpoint (e.g., *Lisa builds cars*).

complete actions before the speech time. As for the English present perfect, it relates a past situation to the present moment. It can be perfective (for example *Mary lost her keys*), or imperfective (*Mary has looked for her keys*), depending on its completedness (Uno, 2014).

The present perfect is “one of the most semantically complex verb forms” (Kearns, 2011, p. 182) because of its complex nature (Han & Hong, 2015; Karpava & Agouraki, 2013; Thumvichit, 2016) and the meanings it can convey (Greenbaum & Nelson, 2002; Kearns, 2011). Therefore, errors such as overgeneralisation (i.e., when the present perfect is used in contexts that require the simple past) and undergeneralisation (i.e., when the simple past is used in contexts that require the present perfect) occur. If there is no correspondence in the speaker’s L1, overgeneralisation errors decrease as the use of the present perfect is reduced (Collins, 2002; Odlin & Alonso-Vázquez, 2006; Svalberg & Chuchu, 1998). In addition, the use of adverbials such as *yesterday* and *in the past* helps learners to identify the action as a simple past and *since* and *already* as a present/past perfect (Bulut, 2011; Declerck, 2006). Likewise, the use of definite adverbial modifiers with the present perfect is entirely incompatible (Bardovi-Harlig, 2002; McCormick, 2008; Rothstein, 2007; Suh, 1992).

When learning English as a FL, learners acquire first the simple past and then the present perfect (Bardovi-Harlig, 2002; Housen, 2002), because for learners to acquire the present perfect, they must reach a certain English proficiency level (Hawkins, 2009). This means that they have to acquire certain morphosyntactic and semantic features necessary to employ and understand the present perfect. In addition, they tend to use it in writing earlier than in speaking (Bardovi-Harlig, 2002).

Speakers whose L1s are tenseless (e.g., Chinese, Japanese, Korean) may identify time references and employ tenses differently from those whose L1s have a tense system (e.g., Spanish, French). For instance, Hinkel (1992) found that Arabic, Chinese, Korean, Japanese and Vietnamese learners of English found it difficult to identify temporal meanings with English tense markers and seemed to interpret time references according to their L1’s standards, a fact that can produce a low rate of learner comprehension. Similarly, Lim (2007) delved into Malaysian students’ intralingual and interlingual errors on the English present perfect by means of elicitation. The errors “were analysed and explained by (1) identifying the

differences between the subjects' mother tongue and the target language, (2) considering the subjects' tendency to use certain verb forms in various situations, and (3) referring to the rules governing its use in English" (p. 368). The results show that most errors are associated with "crosslinguistic differences in the use of perfective verb phrases and intralingual difficulties in differentiating the temporal references of certain verb forms" (p. 368). For these reasons, learners whose L1 does not have a formal or functional parallel in the L2 (e.g., Arabic, Korean, Russian, Vietnamese) will be more dependable on teacher instruction and textbooks.

On the contrary, learners whose L1 is similar in form and meaning to the L2 tend to perform better (Izquierdo & Collins, 2008; Shirai, 1992; Terauchi, 1994). According to previous research on language transfer in SLA, the initial stage of L2 acquisition is L1, which seems to be the foundation for foreign language development (Butzkamm & Caldwell, 2009; Hall & Cook, 2012; Littlewood & Yu, 2011; Schwartz & Sprouse, 1996). In this regard, there are parallelisms between Spanish and English, as they both use adverbs and morphological inflections to convey temporal meanings (Birdsong & Molis, 2001). The English simple past is represented with the 'pretérito imperfecto' and 'pretérito perfecto simple' in Spanish and the present perfect is represented with the 'pretérito perfecto compuesto'. L1 transfer effects are tangible in tense-aspect similarities between the learners' L1 and L2 verbal systems (Izquierdo & Collins 2008; Shirai, 2004; Terauchi, 1994). Therefore, learners have the possibility to apply the new knowledge into the existing one in their L2. However, these parallelisms are, to some extent, discrepant, which can interfere negatively in L2 acquisition (Spada, Lightbown & White, 2005).

Studies on the acquisition of tenses have emerged over the past years but none focus on the use of the simple past vs the present/past perfect by L1-Spanish learners of English. In order to examine how these tenses are acquired, some studies are going to be reviewed. McCormick (2008), for example, focused on the teachability of the present perfect in the written form with intermediate adult L2 learners of Canadian English. He realised that the learners' L1 strongly affected their success. Izquierdo and Collins (2008) also observed that the greater similarity between L1 and L2, the easier it became to master the formal aspects.

On the contrary, the lower correspondence between the L1 and L2, the more they relied on semantic aspects.

Terán (2014) noticed that Spanish learners of English experienced some developmental stages in the acquisition of the present perfect which were characterized by an important role of proficiency level. That is to say, the more proficient the participants were, the more accurate they were in using the present perfect. In turn, Uno (2014) investigated the effects of lexical aspects on the accuracy of Japanese learners of English to use the present perfect tense. She noticed that adverbial modification and boundedness affected the accuracy of learner production, as learners used unbounded verbs in the present perfect with durative verbs, and if there was no durative modifier, learners were less accurate.

In general, according to previous research (e.g., Ayoun & Salaberry, 2008; Chan et al., 2012; Comajoan, 2006), it seems that the use of modifiers aids learners to be more accurate when using either the simple past or present/past perfect. In addition, it is clear that L1 plays an important role in the acquisition of tenses in English (Cai, 2001; Collins, 2004; Giacalone, 2002), and there is also evidence to suggest that instruction accelerates acquisition (Bardovi-Harlig, 2002; McCormick, 2008).

## 1.6 Chapter Summary

Over the years, there have been many changes in regard to language learning. Accordingly, L2 learning has been studied from different perspectives, (i.e., behaviourism, passing through nativism to interactionism), and different theories have also emerged in response to previous ones, such as the *Universal Grammar Model* by Chomsky (1964), the *Monitor Hypothesis* by Krashen (1982, 1985), the *Output Hypothesis* by Swain (1985), or the *Noticing Hypothesis* by Schmidt (1990). Although not all theories agree on the components that are necessary for language acquisition, the truth is that for SLA to occur a combination of elements have to occur to some extent to aid the language learning process: input, noticing, interaction, output, and corrective feedback.

L1 and L2 acquisition processes have been compared with the aim to ascertain whether or not learning the L2 can be equated with learning the mother tongue. Opposing views see both

processes as different, since not only a child brain and an adult brain are different, but also the context and timing constraints.

As Swan (2009) stated, “learning languages is a notoriously complex business, involving the mastery of several different kinds of knowledge and skill” (p. 118). For this reason, it is beneficial to pay attention to areas such as interlanguage and EA and delve into the cause and type of error to better understand and approach the acquisition process.

In this chapter we have analysed how the process of acquisition takes place with a focus on key issues in language learning. The next chapter will centre more in depth on CF, as it is a crucial element in the present dissertation.

## **Chapter 2. The Role of Feedback in Language Learning**

Many definitions of feedback can be found in the literature, as explained in Section 1.3.1. For the reader's convenience, we present two short definitions: according to Leeman (2007) feedback is defined as "a mechanism which provides the learner with information regarding the success or failure of a given process" (p. 112). For correction of written texts, Calvo and Ellis (2010) define feedback as input that leads the writer when revising a composition. Despite the numerous definitions, research converges on the fact that the evidence provided to the learner can be positive or negative. Feedback is a type of negative evidence when it provides the learner with information about non-grammatical patterns in the target language. Gass (2003) refers to positive evidence as "a set of well-formed sentences to which learners are exposed" (p. 226). That is to say, positive evidence is the information or input that describes what is correct and appropriate in the target language, both in an authentic situation or in the classroom with foreigner learners. On the other hand, negative evidence refers to the information about incorrect production in the target language in the form of correction, an explanation, etc.

Positive evidence is available to every child at any age, from any culture and social environment. Despite controversy, although children learn to speak through positive evidence, it may not be enough to acquire a language since interaction and negative evidence become essential. Different theories differ on the importance of input and negative evidence. For instance, researchers following Chomsky (e.g., Krashen, 1982; Truscott, 1996, 1999), state that there is no need to provide negative evidence, since L2 acquisition occurs as a result of positive evidence. On the contrary, researchers in the cognitive-interactionist domain (e.g., Long, 1996) attribute some roles to negative evidence as a useful tool to raise learners' awareness and bridge the gap between form and meaning (Ellis, 2010).

As explained in the previous chapter, feedback plays a role in SLA since it is a useful tool that encourages students to produce output and provides learners with valuable information during the language acquisition process. Learners notice the gap between their interlanguage and the target language, thus, contributing to and facilitating language acquisition. For this reason, it has been

studied over the years by many researchers such as Li (2010), Lyster and Saito (2010), Long (1996), Mackey and Goo (2007), Russell and Spada (2006), Schmidt (1990, 1995, 2001) and Swain (1985, 1995).

Effective feedback can be provided in different ways. In this chapter we will focus on written feedback, since it is the one used in the study, and we will also analyse different types of feedback: explicit and implicit, focused and unfocused, and immediate and delayed.

## 2.1 Types of Feedback and Uptake

Whereas some researchers (e.g., Krashen, 2003; Truscott, 1999) recommend avoiding error correction since it may hinder the learning process, some others (e.g., Ammar & Spada, 2006; Azar, 2007; Doughty & Varela, 1998; Lyster, 1998) support the idea that making students notice there is a mismatch between the input received and the output produced contributes positively to the students' development of the language. In fact, even though they receive the same language exposure, those students whose attention is drawn to some specific structures show a better command of these structures (Lightbown & Spada, 1990). The drawback is that most of those studies have been conducted in a short-term period. Therefore, proving they lead to acquisition is not an easy task.

As previously stated, making and correcting errors are natural processes in language learning. To be able to correct them, learners need to be aware of the mismatch between their output and the target language form (Doughty, 2001; Ellis, 2008; Gregg, 2001; James, 1998). To accomplish this goal, the teacher may make use of different types of feedback. Various classifications have been proposed (Ellis, 2009; Lyster & Ranta, 1997; Lyster & Saito, 2010; Ranta & Lyster, 2007). One of the most well-known was proposed by Lyster and Ranta (1997), which is explained below with examples taken from Suzuki (2005).

- a. *Explicit correction.* The teacher uses an error indicator to make the learner aware that s/he has made a mistake and provides him/her with the correct form. Therefore, the student does not need to sort out the correct form.

S: So we write pacific [pazifik]  
T: Say [pasifik], not [pazifik]



- b. *Recasts*. The teacher repeats the student's utterance, or part of it, but corrected. It is the teacher who provides the correct form without explicitly indicating that the student has made a mistake. Therefore, sometimes the student may not be aware of the correction and may continue with the conversation.

S: In the middle is good. Neutral [neutral].

T: Oh neutral [nju:tral].

- c. *Clarification requests*. It is an attempt to reveal the intended form by using rising intonation, usually in the form of question. It refers to both linguistic and comprehension problems.

S: result [result] of something

T: What did you say?

- d. *Metalinguistic feedback*. Without giving the correct form the teacher tries to elicit it by giving information or asking questions related to the mistake. It is the learner's task to sort out the correct form and, therefore, s/he needs to analyse the utterance and establish linguistic connections.

S: She without.

T: without... what is the verb?

- e. *Elicitation*. The teacher tries to elicit the correct form from the student by making a pause before the wrong form or by asking questions without providing the correct form.

S: Because I enjoy city life [laip]

T: City ...

- f. *Repetition*. The teacher repeats the student's mistake usually with rising intonation so as to emphasize and make the student aware of the correction.

S: When I don't understand what garden [kuden] is in Japan ...

T: [kuden]?

All of them offer learners the possibility to reflect about their input and improve it, stimulating their cognitive processes and leading them to acquisition. On the contrary, not all of them are equally successful depending on what is being corrected and, for this reason, it is the teacher's task to decide what kind of feedback suits best to both the learners and their proficiency level.

The provision of feedback may generate a response which Lyster and Ranta (1997) called *uptake*, a term previously introduced by Austin (1962) in his speech act theory. Uptake represents the learner's attempt to repair the mismatch according to the feedback received. Figure 4 illustrates the error treatment sequence which includes uptake.

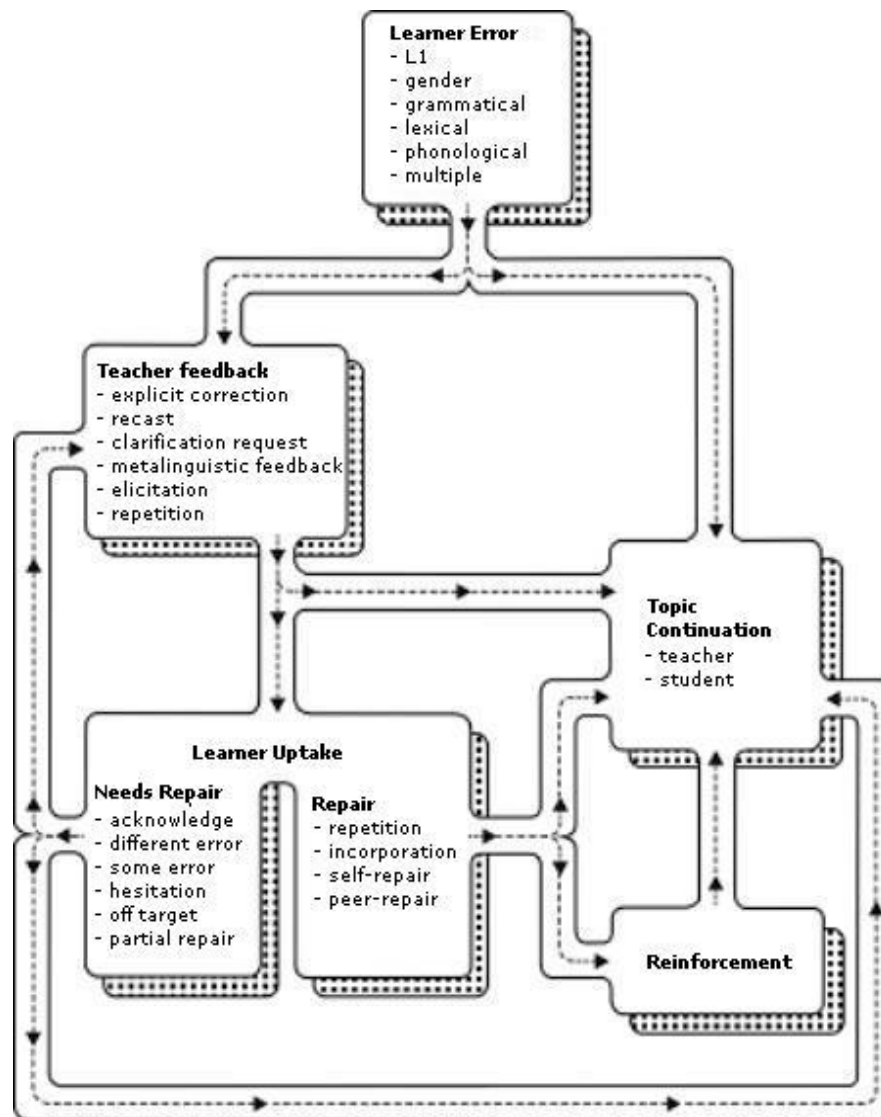


Figure 4. Error treatment sequence

(Lyster and Ranta, 1997, p. 44)

As Figure 4 shows, the error treatment sequence starts with the learner's error, which can be followed either by the teacher's correction by means of any of the types of feedback or topic continuation if there is no feedback provided. If the error is followed by the teacher's correction, this feedback can be followed by the learner's uptake or topic continuation if there is

no correction on the student's part. Lyster and Ranta (1997) distinguish between two types of uptake: first, *repair*, which includes error correction (correct reformulation) and, second, *needs-repair*, where error correction is missing. In the case the teacher's correction is followed by the learner's uptake, there may be two possible outcomes. If it is *needs-repair*, topic continuation or more teacher feedback are two possible follow-ups. On the contrary, if it is *repair*, the possible follow-ups are either topic continuation or teacher's reinforcement, which leads eventually to topic continuation. Regarding the first type of uptake (i.e., *repair*), four different types can be found, as explained below:

- a. *Repetition*. The learner repeats the teacher's correction which includes the correct form.

S: You should go to see doctor.

T: The doctor.

S: The doctor.

- b. *Incorporation*. The learner repeats the teacher's correction which includes the correct form and incorporates it into a longer utterance. It refers to the student's repetition of the correct form provided by the teacher, which is, then, incorporated into a longer utterance produced by the student.

S: When I will be at university I ...

T: No, it isn't future. You should say: when I am at university, I will ...

S: When I am at university I won't live in Castellón.

- c. *Self-repair*. The learner self-corrects after receiving the teacher's feedback.

S: Do the parents time to do so?

T: What?

S: Do the parents ... pare, parents time, do the parents have time to do so?

- d. *Peer-repair*. It is similar to self-repair but in this case it is another student different from the one who made the mistake who provides the correct form in response to the teacher's feedback.

S1: There is poor

T: Sorry?

S2: Pool.

On the contrary, in the second type of uptake (i.e., needs-repair) the student does not repair the original error even though feedback has been delivered. In this case, six different types of uptake can be found:

- a. *Acknowledgment*. The learner accepts the teacher's feedback by saying "yes" or nodding.

S: Two people go out, and pay for one people price ... I don't know  
 T: Exactly. That's exactly what you said. Two people go out and pay for one person.  
 S: Yeah.

- b. *Same error*. Despite receiving the teacher's feedback, the student repeats the initial error.

S: Take one [kuri]  
 T: Take one what?  
 S: [kuri]. [kuri].

- c. *Different error*. In this case the student neither corrects nor repeats the initial error but makes a different one.

S1: Take it from [poket]  
 T: Pocket?  
 S1: Not pocket, uh, [bok]  
 S2: Bottom.  
 S1: Yeah bottom.

- d. *Off target*. The learner responds to the teacher's feedback by using a different correct linguistic form.

S: Many shops are downtown.  
 T: Sorry?  
 S: Downtown, many shops and places everywhere, a lot of people.

- e. *Hesitation*. The student's hesitation is the only response to the teacher's feedback.

T: There is something wrong with the word "business"  
 S: uh... because uh...

- f. *Partial repair*. It refers to partial correction of the initial error but the utterance is not completely repaired.

S: When I don't understand what garden [kuden] is in Japan ...  
 T: [kuden]?  
 S: [guden]?

Both *repair* and *needs-repair* are usually followed by reinforcement in order to strengthen the correct form before proceeding with topic continuation, to encourage students to continue making an effort, and to congratulate them for the good results obtained after correction. Common ways to reinforce this are short statements of approval such as “Good!”, “Better!”, “Yes!” and “Well done!”. At the same time those statements may be followed by more detailed explanations. For example:

S: When I will be at university I ...  
 T: No, it isn't future. You should say: when I am at university, I will ...  
 S: When I am at university I won't live in Castellón.  
 T: Well done! Remember that 'when' used as a conjunction before subjunctive in Spanish, in English we use the present simple.

As explained in the previous section, the choice of feedback may highly depend on the nature of the error and the skill used. For instance, according to much research, although recasts may sometimes go unnoticed (Mackey, Gass & McDonough, 2000, Allwright & Bailey, 1991; Chaudron, 1990; Netten, 1991), they are the predominant response to learner errors as they are immediate and unobtrusive, thus making them appealing to teachers (Roberts, 1995; Doughty, 1994; Lyster & Ranta, 1997).

In Lyster and Ranta's (1997) study, although the authors claimed that recasts seemed to be the most popular feedback technique, it was the least likely to lead to uptake (31%) (see also Mackey et al., 2000; Panova & Lyster, 2002). Explicit correction came second on the list with 50% of uptake with the difference that the repair percentage was much higher than the needs-repair. Clarification requests, metalinguistic feedback, and repetition were similar as they were effective at eliciting uptake from the student (88%, 86% and 78%, respectively), although metalinguistic feedback and elicitation were more successful at eliciting repair (45% and 46% respectively) than either clarification requests (28%) or repetition (31%). The fact of eliciting the correct form forces the student to provide an answer and, therefore, no uptake rarely occurs unless the student keeps silent.

The fact that recasts lead to the lowest rate of uptake may be a consequence of a problem related to ambiguity. Implicit types of corrective feedback, especially recasts, may be too ambiguous for learners to perceive them as corrections. Recasts go so unnoticed in order not to break the flow of communication that students may perceive them as mere confirmation forms of meaning. In

this line, having recasts at the top with the lowest rate of uptake casts doubt on its effectiveness in L2 development. However, it is important to state that the percentage of uptake depends on the choice of feedback that, at the same time, depends on the type of students addressed (level, timing, learning style, motivation, etc.). Havranek and Cesnik (2001) and Kennedy (2010) found that learners with higher levels of proficiency benefit more from corrective feedback than lower proficient learners. In addition, students with a higher level produce a larger amount of uptake following the teacher's feedback than students with lower levels. The reason is because they may be more aware of the fact that they are being corrected, as they comprehend the target language more easily and they also have more linguistic devices to respond to the teacher's feedback. In the case of recasts, Mackey and Philp (1998) reached the conclusion that learners with lower proficiency may not be developmentally ready to move to the next stage. In Ammar and Spada's (2006) study, the results showed that the groups that received prompts and recasts improved but only learners with a high proficiency level were helped by recasts.

As stated before, noticing the gap is crucial to uptake and SLA (Gass, 1997; Schmidt, 1990, 1994; Yousefi, 2011). If uptake takes place, learners have the possibility to reflect about the incorrect form (see Figure 5).

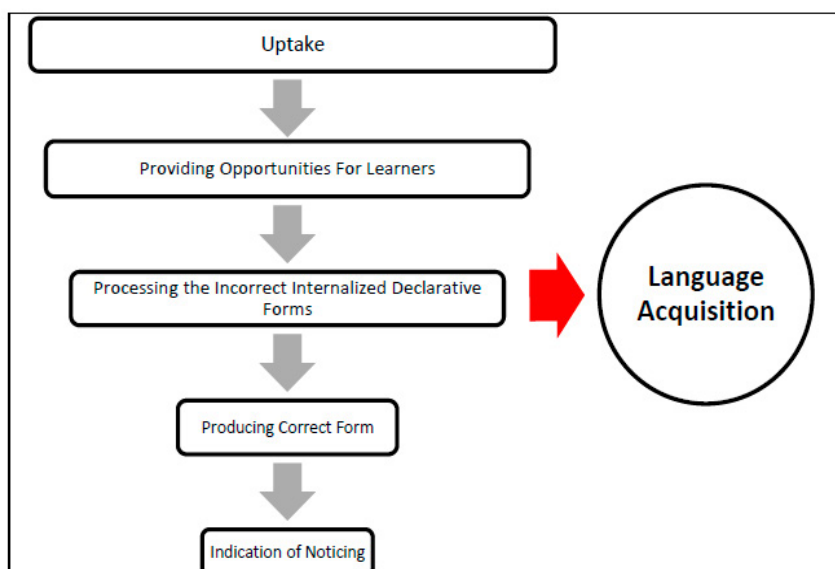


Figure 5. Uptake as a facilitator of language acquisition (Hassanzadeh Nezami, 2012, p. 7)

Some researchers (Ellis et al., 2001; Lyster & Mori, 2006; Nabei & Swain, 2002) agree on the idea that uptake and language acquisition are not equivalent, although it may have a facilitative role (Ellis, Basturkmen & Loewen, 2001; Lyster, 1998; Swain, 1995). If there is no response to the teacher's feedback from the student or if this response fails to repair the error, it does not mean that the learner has not noticed the linguistic form (Loewen, 2004). Mackey and Philp (1998), for example, found that although the feedback students received was not incorporated into the learners' output, they benefited from recasts. On the other hand, even though there is a response to the teacher's feedback from the student, uptake does not necessarily ensure acquisition (Lightbown, 1998). However, there exist theoretical bases to believe that uptake may lead to language acquisition (see Ellis, Basturkmen & Loewen, 2001; Loewen, 2004; Lyster and Ranta, 1997; Schachter, 1983). Nobuyoshi and Ellis (1993) found that learners who responded with self-completed repair following clarification requests improved more than the learners who did not modify their output following feedback. Nevertheless, there is not enough empirical evidence to confirm that uptake guarantees language acquisition.

In conclusion, although CF is thought to be beneficial for the learner and may lead to acquisition (e.g., Bitchener & Knoch, 2008, 2010a, 2010b; Chandler, 2003; Ellis et al., 2008; Ferris, 1999; Sheen, 2007), other researchers like Truscott (1996), Kepner (1991) and Rouhi and Smaiei (2010) hold the opinion that CF is not only ineffective on SLA but also harmful. Feedback plays an important role on SLA but contending that after feedback a form is immediately acquired is simplistic (e.g., Carson, 2001; Ferris, 2003). Sometimes students may feel overloaded with too much feedback, which may be in turn counterproductive. Therefore, employing either focused or unfocused CF in the language classroom is important. This is the topic we address in the next section.

## **2.2 Focused and Unfocused CF**

What to correct or not is an issue in the literature on feedback that depends on many factors. The hotly-debated topic centres on whether all errors or just some specific ones should be corrected, that is to say, focused versus unfocused correction. On the one hand, implementing focused correction means leaving many mistakes uncorrected. On the other hand, implementing

unfocused correction means overloading students with corrections and then probably discouraging and confusing them.

As just stated, the issue of whether focused or unfocused CF is more beneficial for the learner has generated much debate. Researchers such as Bitchener (2008) and Sheen (2007) claim that asking L2 learners to focus on a wide scope of linguistic features at the same time may be out of their processing capacity. Likewise, “the return of papers covered with the inevitable red marks results in looks of disappointment and discouragement on students’ faces” (Semke, 1984, p. 195). In addition, noticing that only some errors have been corrected may confuse the learner by leading them to repeat the uncorrected errors and, therefore, fossilise them (Lalande, 1982; Selinker, 1972; Vandergrift, 1986). Learners’ working memory and attentional capacity is limited, especially when processing oral feedback online more than when handling WCF offline (e.g., Sheen, 2010).

On the one hand, there are studies that show that L2 learners prefer unfocused over focused CF (Anderson, 2010; Ferris, 2006; Leki, 1991). Researchers such as Van Beuningen, De Jong and Kuiken (2012) explained that pupils whose errors were corrected comprehensively made fewer errors in new pieces of writing than learners who did not receive [W]CF” (p. 32). In addition, as Hartshorn and Evans (2012) stated, manageable, meaningful, timely, and constant unfocused CF “can be both practical and effective in improving accuracy” (p. 1).

There are some past studies already mentioned in previous sections that show no improvement of unfocused CF groups over the control group. Semke (1984) conducted a study with 141 first year German FL students who were divided into three WCF groups (direct written CF, error code, and direct correction combined with positive comments, respectively) and a control group who received written comments and questions. The study showed no significant differences between the three WCF groups and the control group, and for this reason they concluded that CF was ineffective in accuracy improvement.

Similarly, Kepner (1991) compared two groups of intermediate Spanish FL learners (direct WCF and control group with content comments in their texts). After 12 weeks, no significant differences in the number of errors were reported. Sheppard (1992), in an ESL context, obtained similar results. Twenty-six upper-intermediate ESL students were divided into two groups



(coded error correction and comments relating to content) and no significant differences were found. On the contrary, Chandler's (2003) study did show significant differences between the experimental and control groups. However, due to the flaws these studies have, the results obtained are questionable for two different reasons: the control groups were no real control groups and in one of the studies there was no pretest to make sure all the students had a similar level. Taking into account that those studies were conducted in the 90s, further research has improved its implementation and methodology.

For example, in Hartshorn's (2008) study forty-seven advanced students of English as a second language were tested on their writing accuracy. Students in the treatment group received manageable, meaningful, and timely comprehensive (i.e., unfocused) WCF on their writings. Results showed improvements in lexical accuracy and in some grammar categories. Constant unfocused CF revealed accuracy improvement, an interpretation shared by Evans, Hartshorn, McCollum, and Wolfersberger (2010), and Evans, Hartshorn, and Strong-Krause (2011).

Another study that avoided the flaws mentioned before is the one by Truscott and Hsu (2008). Forty-seven high-intermediate EFL graduate students, who were divided into two groups (WCF and control group), had to write a guided narrative story based on eight pictures in thirty minutes. One week later, both groups received their writings back so that they could revise them; the WCF group with all the grammatical errors underlined and the control group with no marked errors. One week later, they repeated the process and results showed that although there were no significant differences in error change between both groups, the WCF group outperformed the control group in revision of the errors. Similarly, Van Beuningen et al. (2008) conducted a study on sixty-two secondary Dutch FL students who received WCF in the form of error code, direct correction on different error categories and self-correction. The direct correction group improved in both the revisions and the new texts, whereas the error code group only in the revisions.

Van Beuningen et al. (2012) also investigated the effect of direct and indirect comprehensive (i.e., unfocused) written feedback on second language learners' written accuracy but took it a step further and classified errors into grammatical and ungrammatical. For grammatical errors alone, the direct correction group showed significant improvement, whereas for ungrammatical

errors, error code was more effective and long-lasting in time. Results showed that both direct and indirect comprehensive feedback benefitted from both feedback types, as they improved when compared to the two control groups.

On the other hand, focused CF may seem more beneficial owing to the fact that it targets specific errors and does not entail a cognitive overload for the learner, as Ellis et al. (2008) claimed. If the objective is making the learner aware why he/she has committed the error, focused CF has greater impact for acquisition to take place. Many recent studies applying focused WCF have shown beneficial effects (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010a, 2010b; Bitchener, Young & Cameron, 2005; Ellis et al., 2008; Farrokhi & Sattarpour, 2012; Sheen, 2007, 2010; Sheen, Wright & Moldawa, 2009). However, it is true that “although practitioners certainly want students to use articles and other linguistic features accurately, a heavy emphasis on a few narrowly drawn structures in instruction and feedback would seem too limited a focus for a writing class” (Ferris, 2010, p. 188).

Sheen (2007), for example, examined the effects of focused CF on ninety-one adult ESL learners' accuracy in the use of the articles 'the' and 'a' by means of pretests, posttests and delayed posttests. The learners were divided into three groups: a direct group (errors were indicated and correct forms provided), a direct metalinguistic group (the same as the direct group but with metalinguistic explanations), and a control group. Sheen concluded that both direct CF groups outperformed the control group because the processing of the feedback provided was easy as it was limited to two linguistic forms.

It is true that results from previous studies (Bitchener, 2008; Bitchener & Knoch, 2008, 2009, 2010a, 2010b; Ellis, Sheen, Murakami & Takashima, 2008; Sheen, 2007, 2010) exhibit beneficial and long-term effects when using focused WCF. Despite the fact that most researchers opt for focused CF, it can also be susceptible to avoidance strategies. That is to say, when CF focuses on a limited set of structures, students can choose a different way to express themselves without using that set of structures (Hartshorn & Evans, 2012).

In the field of focused WCF, there are a number of studies that have investigated the use of articles ('the' and 'a') and time retention in subsequent writings, as Table 1 below illustrates.

Table 1. Sample studies on focused WCF

(Qui Guo, 2015, p. 45)

*Focused new text studies*

| <i>Studies</i>                       | <i>Participants</i>                       | <i>Treatment</i>   | <i>Long-term effectiveness</i>             |
|--------------------------------------|---|--|--|
| <b>Bitchener (2008)</b>              | 75 low-intermediate ESL                   | 1. Direct+written&oral metalinguistic explanation<br>2. Direct+written metalinguistic explanation<br>3. Direct correction<br>4. Control  | 2 months                                   |
| <b>Bitchener &amp; Knoch (2008)</b>  | 144 low-intermediate ESL                  | 1. Direct+written&oral metalinguistic explanation<br>2. Direct+written metalinguistic explanation<br>3. Direct correction<br>4. Control  | 2 months                                   |
| <b>Bitchener &amp; Knoch (2009b)</b> | 52 low-intermediate ESL                   | 1. Direct+written&oral metalinguistic explanation;<br>2. Direct+written metalinguistic explanation<br>3. Direct correction<br>4. Control | 10 months                                  |
| <b>Bitchener &amp; Knoch (2010a)</b> | 52 low-intermediate ESL                   | 1. Direct+written&oral metalinguistic explanation<br>2. Direct+written metalinguistic explanation<br>3. Direct correction<br>4. Control  | 10 months                                  |
| <b>Bitchener &amp; Knoch (2010b)</b> | 63 advanced ESL                           | 1. Written metalinguistic explanation<br>2. Written&oral metalinguistic explanation<br>3. Circling<br>4. Control                         | 10 weeks (not circling group)              |
| <b>Sheen (2007a)</b>                 | 91 intermediate ESL                       | 1. Direct correction<br>2. Direct + written metalinguistic explanation<br>3. Control   | 4 weeks                                    |
| <b>Ellis et al. (2008)</b>           | 49 intermediate EFL                       | 1. Direct correction<br>2. Control   | 10 weeks                                   |
| <b>Bitchener et al. (2005)</b>       | 53 post-intermediate ESL migrant students | 1. Written metalinguistic explanation or direct+oral conference<br>2. Written metalinguistic explanation or direct<br>3. Control         | 12 weeks (only on past tense and articles) |

All the studies in Table 1 report outperformance of the treatment groups over the control groups and effectiveness over time. Some in the short-term (2-4 months) and most of them in the long-term (10 months). According to Table 1, it seems that degree of explicitness has an effect on retention over time. This fact may explain why the circling technique in Bitchener and Knoch (2010b) was not equally effective over time than other more explicit forms of WCF. However, as Van Beuningen (2010) remarks, deeper analysis would be necessary to confirm this assumption and to assess the validity of these results on other grammatical or linguistic aspects.

When comparing focused and unfocused CF (e.g., Ellis et al., 2008; Frear & Chiu, 2015), research shows contradictory results. Ellis et al. (2008) found that focused and unfocused groups benefited equally from CF on definite and indefinite articles. Frear and Chiu (2015),

who examined the effectiveness of focused and unfocused WCF on the accuracy of weak verbs and the total accuracy of all structures in new pieces of writing, also found that the focused and unfocused indirect WCF groups outperformed the control groups. On the contrary, Sheen et al. (2009), apart from criticizing the way Ellis et al. (2008) had conducted their study for not distinguishing the focused and the unfocused CF clearly, analysed a group of 80 ESL intermediate students on the use of articles and on some grammatical structures (i.e., articles, copula 'be', regular past tense, irregular past tense and preposition) and found that the focused group outperformed the unfocused and control groups over time. Therefore, in line with Farrokhi and Sattarpour (2011), Sheen et al. (2009) came to the conclusion that unfocused CF may be of limited pedagogical value while focused CF can contribute to grammatical accuracy in L2 writing.

In light of these opposing results, perhaps further research comparing both kinds of CF is necessary. The effectiveness of focused or unfocused CF may also depend on other aspects such as the objective of the lesson, the type of learners, students' preferences, type of errors and CF or time constraints, among others. As an example, if we take into account the level of students, lower proficiency learners may benefit more from focused WCF since they can notice they are being corrected and can concentrate on a specific category without feeling overwhelmed (Bitchener, 2012; Robinson, 1995, 2003; Skehan, 1998, 2003; Skehan & Foster, 2001; Schmidt, 2001; van Patten, 1996, 2004). On the contrary, higher proficiency learners are able to concentrate on a broader set of categories at the same time and would benefit more from unfocused WCF.

In summary, on the basis of previous studies that prove the effectiveness of CF (e.g., Bitchener, 2005; Bitchener & Knoch, 2008; Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russell & Spada, 2006; Sheen, 2006, 2007, 2010; Sheen, Wright & Moldawa, 2009; Van Beuningen, De Jong & Kuiken, 2012), teachers have the option to implement unfocused CF and correct every single mistake or implement focused CF and concentrate on only some of them based on their students' needs and abilities. Although unfocused CF demands more from students because of the range of errors it covers, it seems to be more effective in the long term, and focused CF seems to work better in the short term. In any case, its effectiveness may also depend on other

factors such as the explicitness of feedback or the time span between error and correction. For this reason, the immediate-delayed dichotomy is the focus of the next section.

### **2.3 Immediate vs Delayed CF**

This section will delve into different studies concerning immediate and delayed feedback. Immediate feedback takes place at an early stage of a task performance, whereas delayed feedback occurs sometime after having completed the task. On the one hand, research has shown the benefits of delayed feedback (e.g., Butler et al., 2007; Mousavi & Gorjian, 2018; Mullet et al., 2014). First of all, in the case of the speaking skill, it allows learners to continue with the flow of the conversation avoiding unnecessary interruptions (Long, 1977, cited in Quinn, 2014). Secondly, it gives learners time to retrieve the content and self-correct (Hays et al., 2010) and helps learners achieve greater accuracy in new writings (Bitchener & Ferris, 2012). It is evident that oral CF is much more ephemeral than WCF (Williams, 2012). Consequently, as Bitchener (2012) claims, the fact that the time available to reflect on CF is longer with WCF than oral CF may contribute to deeper processing and, therefore, greater effectiveness.

Butler et al. (2007) conducted a study on the type and timing of feedback. After reading some prose passages participants took a multiple-choice test on information contained in the passages. Feedback either in the form of correct answer (standard feedback) or answer-until-correct was provided immediately for some of the multiple-choice items or after delay for other items. The results revealed that test performance was higher when receiving delayed feedback independent of the feedback provided, something the researchers attribute to the spaced presentation of information.

Mullet et al. (2014) also conducted several experiments in which the timing of feedback was manipulated to check effectiveness. A sample of College Engineering students were asked to hand in some weekly assignments and received feedback immediately after the assignment deadline or one week later. Results showed that the students performed better on subsequent course exams when receiving delayed feedback. Mousavi and Gorjian (2018) also investigated the effectiveness of immediate and delayed CF with two groups of 15 Iranian intermediate EFL learners. They were asked to discuss one of the topics they had covered during the term. The

immediate CF group was corrected straightaway and the delayed CF group 10 minutes later. By means of a pre-test/post-test design, the authors found that delayed CF was significantly more effective in oral fluency, which was not the case for immediate CF.

On the other hand, there are also studies that show beneficial effects on immediate CF. Aubrey and Shintani (2014), for instance, conducted a study which compared the effectiveness of immediate feedback and delayed feedback shortly after having finished writing a text on the use of hypothetical conditions. Results reflect that both groups showed large effect sizes in the immediate posttest when compared with the comparison group. However, the effect sizes for the synchronous corrective feedback group were higher than those for the asynchronous corrective feedback group.

In a similar vein, Li et al. (2016) conducted a study where 120 English students were tested on the passive voice by means of pre-tests, dictogloss tasks, immediate post-tests and two-week delayed post-tests. They were divided into four groups: immediate CF, delayed CF, task only and test only. The analysis of the data gathered illustrates that the four groups improved significantly. However, further analysis shows that on the delayed post-test, immediate CF learners outperformed the test-only but not the task-only group. In addition, in the delayed post-test only the immediate CF group significantly outperformed the task-only group. Therefore, immediate CF seemed to prove some advantage over delayed CF.

Finally, Farmani, Akbari and Ghanizadeh (2017) researched the effect of immediate and delayed CF on students' motivation and performance. Ninety students were divided into three groups (immediate, delayed and control group) who completed a motivation questionnaire and a language test. The results evidence that motivation in the immediate correction group was the highest as well as their language achievement. In light of their findings, the authors claimed that "type of error correction affects performance" (p. 85).

After having discussed some studies favouring immediate feedback and some others in favour of delayed feedback, there are also others that show no difference between both types. Varnosfadrani (2006), for example, compared immediate and delayed oral feedback with a group of 56 intermediate EFL Iranian university students. They had to read a text and retell it with their own words in an interview. They received CF during or following the interview. Results reveal no

significant differences for timing (immediate vs. delayed). Similarly, in another study by Quinn (2014), ninety intermediate adult ESL learners were divided into three groups (immediate, delayed and no CF). After three pretests, instruction on passive voice, three communicative tasks under specific CF conditions and a posttest, results revealed statistically significant improvement for all conditions over time on all measures, but no statistically significant differences between conditions. This means that different timing of CF did not affect L2 development.

In Lavolette et al. (2015), immediate and delayed computer-generated CF were compared. Immediate CF was provided right after the writing was completed, and delayed CF some days later. Again, no statistically significant difference was found. The type of CF (immediate vs delayed) has also been investigated in terms of proficiency level in task-based instruction. It is the case of Li et al.'s (2016) study, where 120 EFL Chinese middle school learners were divided into high and low proficiency levels and further divided into the different groups (immediate feedback, delayed feedback, task-only, and control). Findings showed that low proficient learners benefited only from immediate feedback, whereas high-proficiency learners took advantage of both immediate and delayed feedback. The authors attributed “the disparities between the effects of the two feedback types to the different cognitive demands imposed on learners rather than the timing of feedback” (p. 1).

In spite of ample research on CF, studies have provided no conclusive answer for the issue of whether one type of feedback is more beneficial than the other. Perhaps, teachers should first take into account the skill that is being practised and analyse what type of error has been committed before providing CF, which would help to decide how to treat the error. Therefore, if it is a common mistake, according to Chastain (1971), it should be corrected at the end of the learner's speech or activity. Scrivener (2005), for example, suggested making a list of errors and treating them once the activity has finished. Thus, if fluency is the goal, probably delayed correction is more suitable. However, “if the objective is accuracy, then immediate correction is likely to be useful” (Scrivener, 2005, p. 299).

As important as deciding when to provide CF, how to provide it is also crucial. The next section analyses the dichotomy implicit-explicit CF and aims at examining the effectiveness of both kinds of feedback through a selection of studies.

## 2.4 Implicit vs Explicit CF

CF can be described as implicit or explicit. Explicit feedback takes place by means of direct corrections when learners make an error. It is the preferred way to address corrections when time is limited and learners need to speed up their language development. Instructors make students pay attention to specific linguistic forms by bringing their attention to the errors they have committed. As Ellis (2009) explains, “direct CF has the advantage that it provides learners with explicit guidance about how to correct their errors. This is clearly desirable if learners do not know what the correct form is (i.e., are not capable of self-correcting the error)” (p. 99). In this sense, direct CF can consist of crossing out an unnecessary element, adding one, or providing metalinguistic comments or even the correct form.

On the other hand, implicit feedback offers an indication by means of, for example, a recast, a correction code, underlining or highlighting the errors, or eliciting the correct answer, leaving the student the responsibility to decipher the correct form. In fact, implicit teaching aims “to attract learner’s attention and to avoid metalinguistic discussion, always minimising any interruption to the communication of meaning”, whereas explicit teaching aims “to direct learner’s attention and to exploit pedagogical grammar in this regard” (Doughty & Williams, 1998, p. 232). However, researchers such as Ellis (2009) and Sheen (2007) claim that this dichotomy does not apply to written feedback as it tends to be on the more explicit end. Oral feedback can go unnoticed (e.g., recasts) whereas written feedback is explicit by nature. Even so, there are a myriad of studies which have analysed this difference and effectiveness in both skills.

Explicit correction is useful for the reasons mentioned above. However, in Chandler’s (2003) words, “students feel that they learn more from self-correction”, as “both are viable methods depending on other goals” (p. 267). Authors such as Ferris (1995), Ferris and Roberts (2001) and Lalande (1982) take a stand in favour of implicit corrective feedback as these studies provide evidence on the fact that implicit correction is similar or even superior to explicit correction as it promotes self-discovery when trying to work out the correct form and, therefore, facilitates long-term gains. Similarly, Ellis (2009, p. 100) states that:

*In accordance with the general line of argument by Ferris and Roberts, it might be claimed that indirect feedback where the exact location of errors is not shown might*



*be more effective than direct feedback where the location of the errors is shown [...] as students would have to engage in deeper processing.*

However, although some studies prove evidence on the effectiveness of implicit correction over explicit correction, authors such as Chandler (2003), Doughty and Varela (1998), and Leeman (2003) conclude that CF has to be sufficiently explicit to be noticed by learners if any effect in the classroom is desired. In fact, it is explicit CF the one which gives learners the necessary information to test their hypotheses about correct language (Bitchener & Knoch, 2010b). As it has been proved (Lightbown & Spada, 1990; Tomasello & Herron, 1988; White, 1991), CF is beneficial for learners' accuracy. In any case, the difficult question to answer is to ascertain what contributes more to L2 development, whether explicit or implicit feedback.

As previous studies suggest, explicit CF does not only help to promote noticing of feedback (Ellis, Loewen & Erlam, 2006; Mackey et al., 2007; Nassaji, 2009; Robinson, 2003) but also to facilitate language learning (Bitchener & Ferris, 2012; DeKeyser, 2003, 2007; Ellis, 2009). In fact, it results in higher levels of both uptake and repair (Ellis, 2005). Despite the fact that a lot of research about this topic has been carried out to devise which approach is better, there seems to be no consensus on which one is more effective.

Some are the studies whose results show that explicit feedback (in the form of metalinguistic feedback) results in greater gains than implicit feedback such as recasts (e.g., Carroll & Swain, 1993; Ellis et al., 2006; Havranek & Cesnik, 2003; Lyster, 2004; Nagata, 1993; Zohrani & Ehsani, 2014). Carroll and Swain (1993) conducted a study on the use of dative verbs in Spanish by 100 low intermediate ESL learners who were divided into five groups: A (direct metalinguistic feedback), B (explicit rejection), C (recasts), and D (indirect metalinguistic feedback), and E (control group). After two feedback sessions followed by production tasks, the results showed that all treatment groups outperformed the control group and group A (direct metalinguistic feedback) outperformed the other groups. Similarly, in Lyster's (2004) study, students were divided into the implicit group in the form of recasts, the explicit group and the control group. Both groups outperformed the control group but the explicit group obtained greater gains than the implicit one.

Ellis et al. (2006) also examined the effects of implicit and explicit corrective feedback on the acquisition of past tense -ed. Three groups of ESL low-intermediate learners (two experimental groups and a control group) performed some communicative tasks while receiving either recasts (implicit feedback) or metalinguistic explanation (explicit feedback) after an error in the target structure. Results showed a clear advantage for explicit feedback over implicit feedback. Likewise, Zohrani and Ehsani (2014) investigated the effect of implicit and explicit CF on EFL learners' awareness of and accuracy in English grammar. They concluded that grammar accuracy and awareness in both groups improved. Besides, the explicit group outperformed the implicit one and the authors argued that explicit CF seemed to be more effective. As previously stated, implicit feedback tries to elicit the correct form. Among the different forms of implicit feedback, recasts seem to be the most common type (Ellis et al., 2006) because the flow of the conversation is maintained and the focus is kept on meaning. Their effectiveness is evident if we take into account some previous studies (e.g., DeKeyser, 1993; Doughty, 2001; Leeman, 2003; Long, 1996; Muranoi, 2000). According to Doughty (2001), recasts are an ideal form to facilitate cognitive comparison of interlanguage and target forms while focusing on meaning. In the same line, Leeman (2003) concluded that recasts work for acquisition as they provide positive evidence.

The difference in effectiveness between explicit and implicit is sometimes not too evident. Sanz (2003) analysed the use of pronouns by 28 first-year university learners of Spanish and divided them into two groups. One group received metalinguistic feedback and the other group implicit feedback. By means of sentence completion and written video retelling, results show that both groups considerably improved and there was little difference between them. Similarly, Carroll (2001) conducted a study on word-formation with 100 adult low-intermediate ESL learners. The results show that both kinds of feedback contributed positively to the students' learning and differences between the effectiveness of implicit and explicit corrective feedback were scant. In addition, Kim and Mathes (2001) examined the use of dative verbs by 20 Korean adult ESL learners who were divided into Group A, who received metalinguistic feedback and Group B, who were provided with recasts. Although both groups expressed their preference for explicit feedback, the difference in gains between both groups was not significant.

On the other hand, a different outcome is reached in Salazar-Campillo's (2003) study on accuracy rates in short and long-term learning. One group received more implicit feedback through repetition plus recast and the other group more explicit feedback through metalinguistic information and elicitation. Results reveal that the more implicit feedback (repetition plus recast) led to greater gains. Notwithstanding, the explicit combination resulted in more accuracy rates in the delayed test. In addition, although explicit CF seems to be more effective than implicit CF and taking into account that recasts are highly used because of their effectiveness, Lyster, Saito and Sato (2013) rightly argued that "a variety of CF types is probably more effective than consistent use of only one type", so "it may not be necessary or even possible for researchers to identify the single most effective CF strategy" (p. 21).

The choice of feedback may highly depend on the error that is being corrected (Ellis & Sheen, 2006). For instance, recasts are an immediate and unobtrusive way to provide feedback since they do not break the flow of communication (Lyster & Ranta, 1997). Nevertheless, recasts may be effective when speaking but not so much when writing. However, when returning students their corrected writings, if they are not able to have a personal interview to comment on mistakes, explicit correction "provides explicit signals to the student that there is an error" and then can provide the correct form (Panova & Lyster, 2002, p. 584). In this case, teachers have to make sure that the student is aware there is a mistake and can reflect on it. In fact, in general, students are more likely to attend to the feedback if it is explicit (see, for example, Samuda, 2001).

In a similar vein, Vamosfadrani and Basturkmen (2009) concluded that explicit correction is more productive with phonological mistakes and especially with easy rules. On the contrary, implicit correction is more effective with difficult rules and negotiation of form with lexical errors. Likewise, according to Van Beuningen, De Jong and Kuiken (2012) explicit CF may be better for grammatical errors, whereas implicit CF may be preferable for non-grammatical errors. However, one kind of feedback may work better with some grammar points than others. Ellis (2007), for example, reported that metalinguistic feedback benefited the comparative form in English more than the learning of the past form -ed, whereas recasts showed no learning difference. These results may be related to the prior metalinguistic knowledge students had, which is the next issue to address.

Explicit CF may seem more effective at lower proficiency levels since the learner may receive the necessary information from the selected forms to manage how to correct the error (Li, 2009; 2010). This type of feedback can prevent confusion and undesirable corrections due to lack of learners' understanding and limited linguistic knowledge. In Ammar and Spada's (2006) study, for instance, students with a limited mastery of the possessive determiners 'his' and 'her' benefited more from prompts. On the other hand, indirect CF works well with higher levels due to the students' linguistic ability to deal with the feedback received and relate it to their linguistic knowledge (Bitchener, 2012; Bitchener & Knoch, 2008; Li, 2010). In addition, when used together with metalinguistic feedback, it can be as effective as direct feedback (Baker & Bricker, 2010; Bitchener, 2008; Ellis, Sheen, Murakami & Takashima, 2008; Ellis, 2009; Ene & Upton, 2014; Ferris & Roberts, 2001; Sheen, 2007).

The effectiveness over time may be another differentiating aspect when choosing one type of CF over another. Although explicit CF may be more salient and evident for the learner (Lyster et al., 2013), there are studies that suggest that implicit CF might be more effective in the long term due to its engagement nature (Bitchener & Knoch, 2008; Li, 2010; Mackey & Goo, 2007). In this sense, Tootkaboni and Khatib (2014) investigated the effects of direct feedback, indirect feedback and no feedback on enhancing learners' writing accuracy both in the short and long term. In the short term, findings showed a significant superiority of the direct feedback group over the other groups. In contrast, the learners in the indirect coded feedback group were slightly better than the group that received instruction through direct feedback. This suggests direct CF strategies may be beneficial for short period mastery of linguistic structures. Probably, the best option may be to combine both to obtain improved outcomes (Ferris, 2010).

In conclusion, although both types of CF contribute to acquisition, explicit CF seems to be more effective than implicit (Ellis, Loewen & Erlam, 2006). However, asking what type of CF (explicit or implicit) is more beneficial is too ambitious since multiple factors such as age, level of proficiency, linguistic aspect being corrected, individual abilities, etc. are at stake. In fact, the diversity and even sometimes contradictory results may be due to methodological grounds. For these reasons, teachers should tailor the use of CF to match their students' abilities and characteristics.

In the next section, the explicit and implicit aspects of CF will be further analysed in relation to oral and written CF.

## 2.5 Oral and Written CF

Deciding the way to provide feedback (oral or written) may depend on the kind of activity requested (speaking task, writing exercise, video, etc.) and the situation the teacher and the students are in (e.g., in the classroom, at home, online, etc.). According to Tuzi (2004), oral and written CF share similarities and differences, as depicted in Table 2. What is presented below is a summary, as both oral and written CF will be analysed further in their own sections (Section 2.5.1 and Section 2.5.2, respectively).

Table 2. Main differences & similarities between oral and written CF (Tuzi, 2004, p. 219)

| Oral Feedback                | Written Feedback  |
|------------------------------|---|
| Face-to-face                 | Face-to-face/distant  |
| Oral                         | Written   |
| Time dependent               | Depends whether feedback is given to the writer in class or by the next class |
| Pressure to quickly respond  | Pressure to respond by next class   |
| Place dependent              | Depends whether feedback is given to the writer in class or by the next class |
| Nonverbal components         | No nonverbal components   |
| More personally intrusive    | Depends on whether extended comments are provided or not                      |
| Oral/cultural barriers       | Written/cultural barriers   |
| Greater sense of involvement | Greater sense of involvement  |
| Negotiation of meaning       | Negotiation of meaning  |
| Less delivery effort         | Greater delivery effort   |
| N/A                          | No cut & paste  |

Despite the fact that both written and oral CF are usually directed at individual learners, oral CF usually occurs online whereas WCF occurs offline. Moreover, oral CF is also available to those learners who are listening to the teacher and it demands higher levels of memory since online cognitive comparison is required, whereas WCF does not, as it can be reviewed over time. In addition, oral CF tends to be more focus-on-form oriented as opposed to written correction which may address different aspects in the same writing (e.g., grammar and lexical accuracy, punctuation, coherence, etc.).

Although there are many studies on the effectiveness of CF, few compare oral and written CF. As an exception, Sheen (2010) compared the effectiveness of oral and written CF on the use of English articles with four groups of students who received the following feedback: oral metalinguistic, written direct correction, written direct metalinguistic, oral recast. The findings indicate that the three first CF types outperformed the control group, which suggests that the degree of explicitness of CF impacts on the effectiveness of CF more than the medium in which CF is provided.

Other studies such as the one by Bitchener, Young and Cameron (2005) investigated the effectiveness of different types of CF (direct, explicit written feedback and student-researcher five-minute individual conferences; direct, explicit written feedback only; no corrective feedback) on the use of prepositions, past simple tense, and definite article. The study showed how the combination of written and oral CF in the form of a 5-min individual oral conference resulted in improvement in the use of the past simple tense and the definite article in new pieces of writing but there was no overall effect on accuracy improvement for feedback types when the three error categories were regarded as a single group.

In view of such scarce studies comparing both types of feedback in a single study, in the next two sections oral and written corrective feedback will be analysed separately.

### **2.5.1 Oral CF**

Oral feedback in the form of tutoring and face-to-face conferences is a common practice in language classes. In this way, the student and the teacher can view the text, refer to it, mark it up and/or interact in real time and negotiate at the very moment of providing feedback. Thus, students perceive their own writing as a process.

Taking into account Table 2, oral feedback is face-to-face spoken communication that depends on a time and a place. The participants can respond immediately, interrupt and even overlap each other as well as using nonverbal components typical of oral communication (i.e., body language). Probably, the fact of being able to interact in real time without much effort encourages participants to feel more involved in the task.

Following the classification by Lyster and Ranta (1997), there are six types of oral CF (recast, elicitation, clarification request, metalinguistic feedback, explicit correction and repetition). Panova and Lyster (2002) added a seventh type called *translation*. Other authors have proposed a different classification taking into account explicitness and whether or not the correct form is provided. In 2007 Lyster and Ranta classified them in two broad categories: prompts and reformulations. Reformulations include recasts and explicit correction, and prompts include elicitation, metalinguistic clues, clarification requests and repetition as they constitute a range of signals other than reformulations to push learners to self-repair. Some years later, Sheen and Ellis (2011) suggested a similar taxonomy which also includes the distinction between implicit and explicit CF. These classifications are represented in Table 3:

Table 3. Oral corrective feedback types (Lyster, Saito &amp; Sato, 2012, p. 4)

|                | Implicit   | Explicit   |
|----------------|--|--|
| Reformulations | Conversational recasts <ul style="list-style-type: none"> <li>• a reformulation of a student utterance in an attempt to resolve a communication breakdown</li> <li>• often take the form of confirmation checks</li> </ul>   | Didactic recasts <ul style="list-style-type: none"> <li>• a reformulation of a student utterance in the absence of a communication problem</li> </ul> Explicit correction <ul style="list-style-type: none"> <li>• a reformulation of a student utterance plus a clear indication of an error</li> </ul> Explicit correction with metalinguistic explanation <ul style="list-style-type: none"> <li>• in addition to signalling an error and providing the correct form, there is also a metalinguistic comment</li> </ul> |
| Prompts        | Repetition <ul style="list-style-type: none"> <li>• a verbatim repetition of a student utterance, often with adjusted intonation to highlight the error</li> </ul> Clarification request <ul style="list-style-type: none"> <li>• a phrase such as 'Pardon?' and 'I don't understand' following a student utterance to indirectly signal an error</li> </ul> | Metalinguistic clue <ul style="list-style-type: none"> <li>• a brief metalinguistic statement aimed at eliciting a self-correction from the student</li> </ul> Elicitation <ul style="list-style-type: none"> <li>• directly elicits a self-correction from the student, often in the form of a wh-question</li> </ul> Paralinguistic signal <ul style="list-style-type: none"> <li>• an attempt to non-verbally elicit the correct form from the learner</li> </ul>   |

As previously explained in Section 2.4, explicit CF is an overt and deliberate correction of an erroneous utterance (e.g., metalinguistic explanation, repetition, etc.), whereas implicit CF is a covert correction by restating the utterance with rising/falling intonation (e.g., clarification request, recasts, etc.).

Over the years, recasts have shown to be the most prevalent form of oral feedback (Ellis et al., 2001; Lyster & Ranta, 1997) since they are unobtrusive and immediate. When the teacher reformulates the sentence without the error, learners are supposed to make an immediate comparison between their erroneous utterance and the target language (Doughty & Varela, 1998; Mackey & Philp, 1998). However, whether recasts are salient enough to be noticed by learners has also been a controversial issue (Lightbown, 2001).

Studies analysing their efficacy in L2 learning and in comparison with other explicit types of oral CF are numerous and revealing (e.g., Ammar & Spada, 2006; Dilans, 2010; Ellis, 2007; Ellis, Basturkmen & Loewen, 2001; Ellis, Loewen & Erlam, 2006; Ellis & Sheen, 2006; Erlam & Loewen, 2010; Goo, 2012; Han, 2002; Lochtman 2002; Loewen, 2009; Lyster, Saito & Sato, 2013; Mackey, Gass & McDonough, 2000; McDonough & Mackey, 2006; Nassaji, 2009; Nicholas, Lightbown, & Spada, 2001; Ohta, 2000; Panova & Lyster, 2002; Philp, 2003; Rassaei, 2013; Révész & Han, 2006; Romanova, 2010; Saito & Lyster, 2012; Sato & Lyster, 2007; Sauro, 2009; Sheen, 2004, 2007, 2008, 2010; Shintani & Ellis, 2010; Smith, 2010; Yang & Lyster, 2010).

For instance, Lochtman (2002) studied the impact of recasts, explicit corrections and teacher initiations on learners' uptake with 15 and 16-year old pupils studying German as a foreign language. The results indicated that metalinguistic feedback and elicitations led to uptake at a higher rate than recasts and explicit corrections. However, depending on the different classroom activities, Lochtman reports that "the more analytic and form-focused the activity (grammar exercises), the more initiations to self-correction leading to negotiations of form occur. When the focus shifts to meaning as well (text comprehension), the number of recasts is significantly higher" (p. 281).

Nassaji (2009) studied the effects of recasts and elicitations in interaction of a group of 42 English adults. By means of learner-specific pre-interaction scenario descriptions and immediate and delayed post-interaction error identification/correction tasks, results showed that both recasts and elicitations may be beneficial but those students who received recasts developed a higher degree of immediate post-interaction correction than with elicitation. This may suggest that the effectiveness of CF depends on the degree of explicitness.

In a similar vein, Ellis, Loewen and Erlam (2006) investigated how the use of recasts and metalinguistic explanation with three low-intermediate ESL learners (two experimental groups and a control group) affect the acquisition of the past tense -ed. The posttests (an oral imitation test, an untimed grammatically tests) show that explicit feedback is more appropriate than implicit feedback and how the metalinguistic explanation benefited both implicit and explicit knowledge. Likewise, Rassaei (2013) also investigated how recasts and



explicit correction influenced the acquisition of the definite article *the* and the indefinite *a*. Results revealed that learners who received explicit correction outperformed those who received recasts, since the former noticed the gap or perceived the interlocutors' corrective utterances as corrective feedback more easily.

In addition, Lyster et al. (2013) conducted a study to review different aspects of oral CF (e.g., frequency, preferences, laboratory vs. classroom studies and the targets studied). Results showed that oral CF was significantly more effective than no CF. Furthermore, prompts and explicit corrections resulted in more gains than obtaining feedback in the form of recasts.

Large classes make it difficult for teachers to provide students with individual feedback, which constitutes a big challenge with regard to developing the students' writing skills. At the same time, as a consequence of this situation, even though written work is essential and fulfils different important goals in the learning process (Bean, 2001; Gibbs & Simpson, 2004; Hyland, 2000; Race, 2010), teachers may feel reluctant to assign written tasks to avoid work overload. For this reason, as a variation of oral CF, audio feedback can be a solution to provide individual formative feedback without having to worry much about our students' lack of technological knowledge or devices as in the past. Audio feedback is the recorded response from the teacher which can be listened to more than once and at the student's pace. Its interest started as early as in the 60s (Moore, 1977; Tanner, 1964). The first studies already showed improvement in the students' writings (Coleman, 1972; McGrew, 1969). This way, students can reflect on their own writing process, apply it in further writings, feel motivated to open an audio file, have active participation and take this correction seriously as it is personally addressed to them. In fact, Sipple (2007) notes that audio feedback made students feel more confident in their writings because it "provided more genuine and frequent praise" (p. 24). Similarly, teachers feel unburdened and motivated to continue providing them with quality feedback.

Dagen, Matter, Rinehart and Ice (2008) found that feedback was richer in audio format because it was more detailed and both teachers and students perceived it as something positive. Similarly, Merry and Orsmond (2008) reported the same perception on the students' part but not on the teachers', who did not feel they had provided higher quality feedback. Ice,

Curtis, Phillips and Wells (2007) showed that students detected nuances more effectively and understood content more in detail. Yet, there exist some differences as for the students' perceptions. While Oomen-Early et al.'s (2008) study showed that 52.6% of the students did not like receiving audio feedback as the only means, 84.6% opted for both, audio and written feedback. On the contrary, the participants in Ice et al.'s (2010) study favoured audio feedback for overall quality and written feedback for more specific aspects such as grammar.

In general, students felt that audio feedback helped them to be more involved in the course (e.g., Olesova et al., 2011; Wood, Moskovitz & Valiga, 2011), understand the teacher's comments more effectively (Wood, Moskovitz & Valiga, 2011) and have a more personal relationship with their teacher (Kirschner, van den Brink & Meester, 1991). In this sense, audio feedback can benefit both teachers to save time and provide better quality feedback and students to receive richer and extensive feedback to apply in future written work.

To sum up, taking into account the studies on oral CF analysed in this section, it would be necessary to consider aspects such as the degree of explicitness, classroom activities and students' abilities to implement different types of oral CF correctly. The next section will deal with the counterpart of oral feedback, that is, written corrective feedback.

### **2.5.2 Written CF**

WCF is the information provided by the reader to the writer in the form of comments. In comparison to oral feedback, written feedback is usually provided when the teacher and the student do not share either the space or the time. Therefore, the immediacy and pressure to respond disappear and can be done hours and even days later. Due to the lack of immediacy, the feedback provided is unidirectional (i.e., from teacher to student) and may not lead to further comments or negotiation on the part of the student, which may undermine the students' sense of involvement.

As a subcomponent of CF, effectiveness of WCF is still an arduous debate. While some studies show it has little or no effect for L2 development (e.g., Dulay & Burt, 1977; Kepner, 1991; Semke, 1984; Sheppard, 1992; Truscott, 1996, 1999, 2007; Truscott & Hsu, 2008) some others present the opposite view (e.g., Bitchener & Knoch, 2009; Bitchener & Storch,

2016; Ellis, Erlam & Loewen, 2006; Ellis, Sheen, Murakami & Takashima, 2008; Evans et al., 2010; Ferris, 1999, 2006; Ferris & Roberts, 2001; Gu enette, 2007; Guo, 2015; Hartshorn et al., 2010; Hyland & Hyland, 2001; Sachs & Polio, 2007; Stefanou, 2014). Already in 1996 Truscott argued that CF on L2 was ineffective and even harmful for the learner. Since then, the debate about the role of CF has been of considerable interest.

Over the last 20 years the countless studies on the effectiveness of WCF have revealed a positive influence, both on revision (Ashwell, 2000; Ferris, 1999, 2006; Ferris & Roberts, 2001; van Beuningen, Jong & Kuiken, 2008) and new texts (Bitchener, 2008, 2009; Bitchener & Knoch, 2008, 2010a, 2010b; Bitchener, Young & Cameron 2005, Chandler, 2003; Ellis, Sheen, Takashima & Murakami, 2008; Sheen, 2007; Sheen, Wright & Moldawa, 2009). On the contrary, other studies have been criticised for flaws when conducting the study and, therefore, for unreliable results. For example, some studies lacked a control group (e.g., Fazio, 2001; Kepner, 1991; Lalande, 1982; Robb et al., 1986; Sheppard, 1992) and that made it impossible to compare the effects of correction groups over no correction groups. Some other research was based only on revised texts (e.g., Ashwell, 2000; Fathman & Whalley, 1990; Ferris & Roberts, 2001; Robb et al., 1986), which, according to Gu enette (2007) and Truscott (2004, 2007), may not generate enough evidence to confirm lasting effects over time as compared to new writings. Another shortcoming open to criticism is the freedom provided to students when carrying the writing task at home without being able to control time or not following the rules established by the teacher (e.g., Ashwell, 2000; Chandler, 2003). In Truscott's (2007) words, "A writing task that students do with help from the teacher (the revision) is obviously not comparable to one they do on their own (the original essay), and so a study with this design does not yield any measure of learning, short-term or otherwise" (p. 257).

All in all, WCF has proved to be beneficial, but evident deficiencies in research as the ones mentioned above make it difficult to reach a consensus on the value of WCF. In order to shed some more light on this controversial issue, in the next section we will deal with the different types of WCF.

### 2.5.2.1 Direct and indirect WCF

Although both terms have been previously explained (see Section 2.4), their role in the field of WCF will be explained in this section. Direct (explicit) written corrective feedback (DWCF) is defined as “the provision of the correct linguistic form by the teacher to the student” (Ferris, 2006, p. 83). On the other hand, indirect (implicit) written corrective feedback (IWCF) is understood as the indication the teacher gives “in some way that an error has been made by means of an underline, circle, code, or other mark—but does not provide the correct form, leaving the student to solve the problem” (Ferris, 2006, p. 38). IWCF fosters students’ reflection and noticing (Bitchener & Knoch, 2009) as well as language engagement, problem solving and control of their own learning process, which facilitate their progress and reduction of error (Rahimi, 2009).

Indirect feedback can be provided in different ways. The error can be signalled by means of colours, underlining or circling, which would be an effective way to draw the student’s attention. Another option is indicating there is an error in the margin by simply signalling or using a code, which would be an indirect and more demanding way to identify the error. This kind of feedback requires more attention on the students’ part since they are the ones who have to come up with the correct form after receiving the teacher’s feedback.

Direct feedback can be provided in different ways too. The teacher can provide the learner with the correct form by crossing out unnecessary words, inserting missing words or even writing the correct form, which could be seen as the most direct form of direct feedback. According to Ferris (2002, 2003), this type of feedback is useful in the sense that it provides learners with the correct form when dealing with untreatable errors, especially at lower levels, because it requires little cognitive process. On the contrary, as discussed in Section 2.4, it may be a disadvantage for long-term acquisition.

Both types of feedback have been the focus of WCF research. Regarding DWCF studies, Chandler (2003) proved that the students’ correction of grammatical and lexical errors between assignments reduced errors in subsequent writings. Results showed that underlining and direct correction are more beneficial types of WCF than describing the

kind of error for reducing long-term errors. Likewise, Bitchener et al. (2005) studied the effects of different types of CF (direct CF with and without oral conferencing) on the accuracy of new pieces of writing. Although it was evident in the definite article and past tense but not in the use of prepositions, results showed that direct correction plus oral metalinguistic explanation was more beneficial than direct correction alone in new pieces of writing over a period of 12 weeks. In a similar vein, Sheen (2007) studied the impact of direct-only explicit correction and direct metalinguistic correction (more explicit WCF) on the acquisition of articles by intermediate ESL learners. Results show that although both experimental groups outperformed the control group on the immediate posttest, the direct metalinguistic group outperformed the direct-only correction group in the delayed posttests.

Other combinations were also analysed by Bitchener (2008). The effectiveness of four direct feedback combinations (1: direct error correction with written metalinguistic explanation and oral metalinguistic explanation; 2: direct error correction with written metalinguistic explanation; 3: direct error correction; and 4: control group) on three pieces of writing of 75 low intermediate ESL students was analysed. Groups 1 and 3 outperformed the control group while group 2 failed to do so. The same year, Bitchener and Knoch (2008) analysed the use of the referential indefinite article *a* and the referential definite article *the* by four low-intermediate ESL groups which received different combinations of written CF (1: direct error correction, written, and oral metalinguistic explanation, 2: direct error correction and written metalinguistic explanation, 3: direct error correction, and 4: control group). The results show that those who received WCF outperformed those who did not, improved the accuracy of their writing to a similar extent, and retained their level of accuracy over seven weeks. One year later, Bitchener and Knoch (2009) obtained similar results in all four post-tests.

Similarly, Santos, López-Serrano and Manchón (2010) investigated two types of direct WCF (reformulation, in which the teacher rewrote the incorrect sentence, and error correction, where the teacher underlined the error and provided the correct form) on eight secondary-school EFL learners at the intermediate level. Results show more benefits for

error correction over reformulation probably due to the fact that “reformulations led to many more changes to the students’ original texts than in the error correction condition” (p. 149).

At this point, it is important to consider that the linguistic aspect being treated or the proficiency level of the students have to be taken into account if reliable conclusions are expected. The previous studies suggest that direct WCF alone may be beneficial for low-intermediate students but intermediate ones may benefit from metalinguistic explanations too.

As for IWCF, different types have been analysed (e.g., underlining, correction code, metalinguistic WCF). Lee (1997), for instance, reported that after having compared two indirect WCF conditions (underlining and marginal feedback) with EFL students in Hong Kong, underlining showed to be more effective than marginal feedback and no feedback for self-correction. Ferris and Robert (2001) also analysed two indirect types of feedback (underlining and underlining plus code) of five different error categories. The study showed that both groups outperformed the control group but no significant differences on accuracy were noticeable between underlining and underlining plus codes.

The use of a correction code was also analysed by Greenslade and Félix-Brasdefer (2006) who investigated the effects of two types of feedback: coded vs. underlining. They found that although both types of WCF were effective, the coded feedback technique exhibited more significant gains in the self-correct process. The same happened in Muñoz’s (2011) study, which compared coded WCF with underlining on university sixty-two students of Spanish. The results showed that the coded condition group performed significantly better both in the short and long run than the underlining condition group. In addition, the latter did not outperform the control group, which only received feedback in the form of praise and suggestions. Similarly, the study conducted by Ferdouse (2013) with twenty university students exhibited more benefits from receiving coded feedback than non-coded feedback and showed their preference for it, as it happened in Chandler (2003) and Nakazawa (2006). However, in the last two studies, no significance between coded and uncoded WCF was found. On the contrary, a study conducted by Ferris and Roberts

(2001) with a group of seventy-two ESL college students showed that underlining was as effective as coded WCF from draft to draft. Indeed, underlining an error was salient enough to have a positive impact on the learning process.

Bitchener and Knoch (2010b) discovered that although the three advanced ESL treatment groups (written metalinguistic explanation with examples but no direct correction, circling of errors and written metalinguistic WCF) outperformed the control group in the two functional uses of the English article system, those who received written metalinguistic explanation and those who received both written metalinguistic explanation plus an oral form-focused review demonstrated a superior longitudinal effect (ten weeks) than simply underlining.

Most research carried out on direct and indirect CF so far have shown positive results. However, we find mixed results when comparing both types of CF in the same study. There are some studies which show no difference (e.g., Bitchener and Knoch, 2010a; Mantello, 1997; Nakazawa, 2006; Vyatkina, 2010) and studies that show the advantage of one of them over the other. As for studies exhibiting no differences, Mantello (1997), for example, compared two types of feedback: coded WCF (indirect) and reformulation (direct). Both groups benefited equally from coding and reformulation in the narrative past tense. Nakazawa (2006) provided WCF under four conditions: direct WCF, coded WCF, lists of revising criteria (no correction but a list with frequently occurring errors), and control group. After revising their five writings, results showed that students who received direct WCF benefited over the other groups in the short-term. However, after 15 weeks there was no advantage between the four groups. Similar findings are reported in Vyatkina's (2010) study which reported significant gains in direct feedback over indirect with just underlining or coded correction in the short-term in verb-related, noun-related, lexical, structural, word order, and spelling errors. Nevertheless, after 16 weeks, no significant differences between the two groups were found.

Finally, Bitchener and Knoch (2010a) investigated the effects of WCF of four groups of low-intermediate ESL students (direct corrective feedback and written and oral metalinguistic explanation; direct corrective feedback and written metalinguistic

explanation; direct corrective feedback only, and control group). The treatment groups outperformed the control group but no difference in effectiveness was found between the three treatment groups.

With reference to research presenting differences in effectiveness, there are studies that show the advantage of DWCF over IWCF (e.g., Chandler, 2003; Nakayama, 2002; Shintani, Ellis & Suzuki, 2014; van Beuningen, Jong & Kuiken, 2008). Chandler's (2003) study, as mentioned before, exhibited the advantage of direct WCF over the three forms of indirect WCF (underlining with and without codes) after 10 weeks of treatment. Similarly, van Beuningen, Jong and Kuiken (2008) compared the effectiveness of direct and indirect CF. Results show that although short-term effects were found for both direct and indirect corrective feedback, only direct feedback proved to have a significant long-term effect.

Studies such as the one by Nakayama (2002), which found that 29 students learning to write Japanese appreciated more detailed feedback, may suggest that the more explicit the feedback is, the more problems the learner may have to correct the error, especially if their level of proficiency is low (Ferris, 2006). Some studies (e.g., Thouësny, 2011; Vyatkina, 2010) show how students found difficulties to correct their errors because the underlining technique was not accompanied by a code or explanation or because they did not understand the linguistic terms. For these reasons, the teacher has to provide sufficient information to make sure that students understand the terms used (Lee, 1997; Muñoz, 2011). In addition, although students sometimes prefer direct correction because it is straightforward, they admit they learn more from self-correction (Chandler, 2003). In regard to direct feedback, Ellis (2010) maintains that "focused error correction leads to gains in linguistic accuracy and that the more explicit the feedback is, the better the results" (p. 173).

In turn, Shintani, Ellis and Suzuki (2014) conducted a study comparing written metalinguistic explanation and direct correction in new pieces of writing. They analysed the use of indefinite articles and the hypothetical conditional by a group of Japanese EFL university students based on a dictogloss writing task. Contrary to a very similar study conducted one year before, results report increased accuracy for the hypothetical



conditional but not for the indefinite article and longer lasting effects in direct correction than metalinguistic explanation, which may be due to the complexity of the hypothetical conditional. Differences between both studies may be explained in terms of language exposure (ESL and EFL) and proficiency level (lower-intermediate and pre-intermediate).

The previous studies have shown positive effects of direct feedback. However, there are also studies which show advantage of IWCF over DWCF (e.g., Baleghizadeh & Dadashi, 2011; Ferris, Chaney, Komura, Roberts & McKee, 2000; Lalande, 1982; Sachs & Polio, 2007; Sheen, 2007; Storch & Wigglesworth, 2010; van Beuningen, Jong & Kuiken, 2011). These studies favour indirect CF because it is argued that students have to make the effort to find the right form and correct one's own errors (Corder, 1967; Lyster & Mori, 2006). For example, Lalande's (1982) study compared direct CF and indirect coding CF on grammatical and orthographic errors with two groups of 60 German intermediate FL learners. The group that received indirect CF benefited more than the direct one. Lalande (1982) suggested that making the student engage in "guided learning and problem solving" may foster deeper language processing and, therefore, "long-term acquisition" (p. 415). Likewise, Baleghizadeh and Dadashi (2011) compared the effect of direct and indirect CF on spelling accuracy in English of 44 Iranian male students. By comparing the written work dictated by their teachers, results show that indirect CF was more effective than direct CF in correcting spelling errors.

Sachs and Polio's (2007) research examined the effectiveness of two types of feedback, reformulation and written error corrections. Students in the error correction condition outperformed the reformulation group. Similar results appear in Storch and Wigglesworth (2010), whose study examined the nature of the learners' engagement with the feedback received (reformulation and editing code) to better understand why some feedback is taken up and retained and some is not. Results showed that students engaged more with coded WCF than with reformulations.

Sheen's (2007) study reported that those students who received metalinguistic explanations retained the gains they had made in their immediate post-test, whereas those who received direct WCF alone did not retain their level of performance in the use of

English articles. Ferris, Chaney, Komura, Roberts, and McKee (2000) found that direct correction was more effective in the short term, and indirect feedback in the long term. Van Beuningen, Jong and Kuiken's (2011) study exhibited more benefits from coded WCF than direct CF in non-grammatical categories (lexical and orthographical errors).

Shintani and Ellis' (2013) study compared written metalinguistic explanation and direct correction in revised and new pieces of writing by means of three picture composition tasks done by 49 lower-intermediate ESL students. The results showed improvement in the use of indefinite articles by the metalinguistic explanation group but not by the direct correction group in the immediate post-test. Nevertheless, this improvement did not last long, as reflected in the delayed post-test after two weeks.

It is also important to point out that the effectiveness of feedback may also depend on the kind of error. Yang and Lyster (2010), based on Ellis' (2005) work, classified linguistic error types in rule-based and item-based errors. Similarly, Ferris (1999) distinguished between *treatable* and *untreatable* errors. Rule-governed or treatable errors happen when the wrong rule is applied and can be easily corrected or treated by referring to the rules, e.g., the past of regular verbs. On the contrary, item-based or untreatable errors are those that may not be solved or treated by applying a set of rules, e.g., collocations.

In this sense, the type of CF may depend on a wide range of aspects, such as the kind of error, the learner's existing knowledge or their cognitive processing. Therefore, more explicit types of WCF may be more effective when applied on more complex rule-based errors and less-governed errors, and less explicit forms of WCF may work better with simple rule-based errors. Likewise, when the learner has the existing knowledge, less explicit types of WCF may be more effective when dealing with complex rule-based errors. On the contrary, when the learner does not have the existing knowledge, more explicit types of feedback may be more beneficial for simple and less rule-based errors. In other words, WCF on rule-based errors may be more effective for knowledge consolidation since the processing of the explicit knowledge may be activated over and over again to correct different errors which share the same rule. Nevertheless, with item-based errors, the explicit knowledge may only be activated once due to the nature of the error. Some

research supports these facts. For example, Storch and Wigglesworth (2010, p. 319) showed that proficient learners “clearly noticed the reformulations and were able to address their errors in their subsequent writing” if the errors were superficial. Ellis (2008) also stated that “the effectiveness of direct and indirect feedback is likely to depend on the current state of the learners’ grammatical knowledge” (p. 355).

In view of the above-mentioned research on both direct and indirect written feedback, we may say that there are still many questions with uncertain answers. Sometimes the reason for inconclusive results is due to poor research design and lack of comparability (Ferris, 2004; Guénette, 2007). Other researchers (e.g., Bitchener, 2008; Ellis, Sheen, Murakami, & Takashima, 2008; Sheen, 2007) attribute it to an overwhelming process in which students are given feedback on all their errors, something that confuses the learner. Guénette (2007), for example, points out some studies lacked a control group (e.g., Polio et al., 1998; Robb et al., 1986), in some studies the effectiveness of CF was measured differently (e.g., Fathman & Whalley, 1990; Ferris, 2002), others measured effectiveness over a period of time (e.g., Polio et al., 1998; Robb et al., 1986) while others just did it measuring a pretest and a posttest (Semke, 1984), and in some other cases the group differences were often not statistically significant (e.g., Lalande, 1982; Semke, 1984).

In sum, restricting one type of CF to certain errors or students may be too simplistic, whereas using a combination of both can be more enriching. Ferris’ (2010) words seem to capture this feeling by claiming that “there is some role for written CF [corrective feedback] in L2 instruction, although the nature and extent of this role remains in dispute” (p. 183).

In the previous sections we have focused on a wide range of CF techniques, how they work and their advantages and disadvantages. Taking into account the aim of the present study, the next section will focus on how both parties in the learning process, that is, learners and teachers, perceive WCF.

## **2.6 Learners’ and Teachers’ Perceptions of Written Corrective Feedback**

The study of how learners and teachers value WCF has sparked increasing interest in the literature on feedback (e.g., Amrhein & Nassaji, 2010; Brown, 2009; Diab, 2005; Karim & Nassaji,

2015; Lee, 2008; Montgomery & Baker, 2007; Simard et al., 2015). Choosing the type of WCF should not be a unilateral decision but an agreed one (Evans et al., 2010). In this way, feedback providers and learners may make the most of CF. Hence, to ensure instructional effectiveness it is of great concern to refer to both students' and teacher's WCF perceptions (e.g., Diab, 2005; Ferris & Roberts, 2001; Hyland, 2003; Leki, 1991; Makino, 1993; Montgomery & Baker, 2007; Schulz, 1996, 2001). If students and teachers share similar perceptions of the effectiveness and usefulness of WCF, the learning process may be effective. However, the opposite may lead to students' dissatisfaction and, therefore, the learning process takes the risk to fail (Brown, 2009; Rauber & Gil, 2004; Schulz, 2001; Yoshida, 2008).

First, it is relevant to make a difference between two terms that do not necessarily need to go hand in hand when receiving feedback: *attitudes* and *beliefs*. According to Dörnyei (2005), *attitudes* refer to an influence set in our minds, whereas *beliefs* refer to an empirical conviction. Therefore, a student may think that a teacher's correction is helpful for him/her but may not want to be constantly corrected. That means that what learners want is not always what is best for acquisition. In fact, students tend to prefer grammar and teacher's correction (Leki, 1991), which means leaving content and organisational patterns aside. It seems that the effectiveness of WCF may sometimes hinge upon students' preferences, since they may be more likely to pay attention to the correction for the fact of trusting its results more (McCargar, 1993; Schulz, 2001). For this reason, a more effective environment can be developed if students' WCF expectations and preferences are shared with the teachers'.

Despite the fact that some researchers distrust teacher's written feedback, (e.g., Altena & Pica, 2010; Randolph & Lea, 2010, cited in Elashri, 2013) some others (e.g., Connors & Lunsford, 1993; Listiani, 2019; Sommer, 1982; Zamel, 1985; Zhan, 2016) have shown that students in general find it a useful tool to improve their writings. In the same vein, some studies (e.g., Ferris, 1995; Lee, 2005; Schulz, 2001) have noted that students opt for a grammar-based approach while some others (e.g., Amrhein & Nassaji, 2010; Ashwell, 2000; Lee, 2008) have found that students prefer content-based correction. Another difference was made by Hedgcock and Lefkowitz's (1994) study, whose results showed that FL and ESL students seem to have different perspectives. Whereas FL students prefer receiving feedback on grammar, lexicon and mechanics

of their written texts, ESL students prefer receiving feedback on content and organisation. One possible explanation may refer to the difference in the learning context. Probably L2 learners focus more on their L2 knowledge, while ESL learners are more concerned with the writing skill itself.

In addition, some circumstances (e.g., when students do not understand the meaning of WCF, or they like having all the errors marked, or they prefer being corrected in a different way) may lead to corrections that do not match the teacher's expectations. Some studies (e.g., Cathcart & Olsen, 1976; Ferris, 1995; Lee, 2005; Radecki & Swales, 1988) demonstrate that students prefer receiving a variety of different types of WCF or WCF on content rather than surface errors such as grammar (e.g., Semke, 1984; Woroniecka, 1998; Zamel, 1985). For example, Radecki and Swales (1988) analysed 59 ESL students' perceptions on feedback by means of a questionnaire. Eighty-seven percent of the students expressed they were happy to receive feedback on any linguistic error and tried to correct themselves. On the contrary, only 13% were unwilling to revise and modify their writings and preferred to receive direct correction of the most relevant errors. Similarly, Lee (2004) found out that teachers thought marking all errors was beneficial for students and most students wanted teachers to mark all their errors.

In Olajedo's (1993) and Saito's (1994) studies, students showed their preference to be corrected by their teacher and not by peers, despite the fact their teacher thought it was advisable. Students usually appreciate and value their teachers' feedback for the development of their writing skill. It is the case of Mahfoodh's (2011) study, whose students expected their teachers to focus on all aspects of written texts. However, there are numerous studies which show discrepancies between teachers' and students' preferences, a fact which may sometimes be the cause of ineffectiveness of correction (see Chiang, 2004). For instance, in Diab's (2006) study, teachers drew their attention to grammar, spelling and punctuation while students placed more emphasis on grammar and writing style. In addition, students expected their teachers to correct all their errors by using direct correction. Likewise, in Hamouda's (2011) study, although students favored overall correction, teachers did not. Nevertheless, in this study, teachers and students agreed on the use of the red colour for corrections by means of circling and underlining techniques.

As for studies on the teachers' perceptions, there are some that provide interesting results regarding the combination of teacher's perception and practice (e.g., Ko, 2011; Lee, 2008; McMartin-Miller, 2014; Montgomery & Baker, 2007; Shahrani, 2013). Ko (2011) compared the perceptions of teachers of Korean as a foreign language (KFL) and teachers of English as a second language (ESL) in North America in terms of written feedback. After analysing the 153 KFL/ESL teachers' practices with an online questionnaire, the author reports that

*KFL teachers in relation to ESL teachers were more concerned with language-related local issues over writing-related global issues. The dissimilarity was consistent even with multiple-draft approach usage: ESL teachers changed their focus of concern from global issues on early drafts to local issues on later drafts, while KFL teachers focused on local issues across all drafts (p. 41).*

Montgomery and Baker (2007) analysed a set of questionnaires answered by ESL teachers on the feedback they provided. After comparing the teachers' feedback with the students' writings, results report the teachers' performance did not match their beliefs. Similarly, Lee's (2008) study analysed teachers' perceptions and practice. Based on a set of interviews, this study proved that teachers' feedback implementation did not match their beliefs.

Al Shahrani (2013) carried out research on teachers' beliefs and practices in WCF in one Saudi university and on forty-one students' preferences. By means of a collection of interviews and questionnaires, results presented a general equivalence between the students' preferences and the teacher's practice. In addition, as in the previous two studies, the teachers' beliefs did not match their own practices. One year later, McMartin-Miller (2014) analysed the feedback provided by three SL teachers to nineteen university students in the United States. The interviews of both teachers and students revealed that the feedback provided by the three instructors differed from one another, and that students preferred to obtain comprehensive feedback.

As a conclusion, it can be inferred from the previous studies that learners' and teachers' preferences and perceptions do not always go hand in hand. In fact, teachers do not even match their practices to their perceptions on some occasions. That means that communication between instructors and students in a specific context is necessary so that student learning can be ensured. Forming a negative attitude towards teachers' WCF or thinking the marking students receive is useless can weaken the effectiveness of correction.

After having reviewed some relevant aspects related to WCF perceptions and practices, it is important to discuss who the providers of WCF may be and the differences among them.

## **2.7 Feedback Provider**

One of the main features of feedback to be effective, and therefore, enhance learning should be that it primarily addresses the students' needs (Kulhavy & Stock, 1989), but it should also be objective, goal-oriented, clear, simple and understandable (Shute, 2008). However, as discussed in Section 2.2, the provision of timely and accurate feedback is not an easy task. It is, in fact, one of the challenges teachers have to face. Classrooms are usually packed and teachers find difficulties to fully attend to their students' writing demands. For this reason, teachers try to find shortcuts (e.g., peer feedback, self-feedback, automated feedback) to reduce their workload. This choice cannot be made at random, as it becomes necessary to take into account how a specific kind of feedback may affect students' learning. We are now going to analyse the different agents that can provide feedback.

### **2.7.1 Teacher's feedback**

Teacher's feedback seems to prevail over peer and automated CF (e.g., Bai, 2012; Zhou, 2013), as it has been traditionally considered the teacher's responsibility for the students' benefit (Hyland & Hyland, 2006; Knoch, 2011). For example, it is a common practice in English writing classes as a method to improve the students' writing skill (Black & William, 2009; Wei, 2015). According to Kroll (2003), L2 writers benefit most when teachers employ a variety of CF which is accessible in classroom settings.

Despite being overloaded with corrections, oversized classes and time constraints, the advantages of teacher feedback are numerous and the impact it can have on the students' attitude and learning is of utmost importance. Apart from helping students to improve their writing competence (O'Mahony et al., 2013; Wang & Liu, 2012), teacher's feedback has been claimed to help to learn more and better (Chandler, 2003; Morch et al., 2017). In addition, although it is not too common, providing feedback does not only imply helping to correct

errors but praising students' success (Ferris, 1995). That is to say, apart from the pedagogical function, teachers also have a psychological role in the classroom. In this line, Bruno and Santos (2011) point out the importance of being selective and only providing feedback on the most essential things in order to maintain students' interest and motivation. In addition, Askew and Lodge (2000) and Brookhart (2008) agree on the fact that exorbitant written and/or oral comments should be avoided so as not to overwhelm students. If noticing (and ultimately, learning) is the objective, disproportionate quantities of feedback will only confuse students.

Despite the fact that students value their teacher's written feedback for its effectiveness as a way of improving their written work, both its immediate and long-term effects are still dubious. WCF is usually one-way and the dialogue disappears unless learners need clarification. That is to say, when teachers provide their students with written feedback, they expect them to understand what the teacher means and to react in a specific way. However, this may not always happen because the feedback provided may be excessive, inconsistent, misunderstood, of poor quality, vague and authoritarian, among others (Connors & Lunsford, 1993; Ferris, 2003). Thus, it seems students rely unconditionally on their teachers, but sometimes they misunderstand their comments or do not even know how to correct the error (Conrad & Goldstein, 1999), a fact that may even make them feel overwhelmed and less enthusiastic up to the point of weakening their initiatives (Zhan, 2016). Some studies, in fact, (e.g., Brown, 2009; Diab, 2005; Montgomery & Baker, 2007; Simard et al., 2015) have revealed that the lack of understanding between the teacher's feedback and the students' perceptions can lead to erroneous hypotheses and corrections. Therefore, teachers must make the effort to provide tactful, clear and concise feedback as well as in the right quantity.

There are numerous studies, already mentioned in previous sections, which have proved the effectiveness of teacher's feedback (e.g., Chandler, 2003; Ferris, 2006; Ferris & Roberts, 2002; Lalande, 1982; Robb et al., 1986; Sheen, 2007). Ferris and Roberts (2001) divided 72 university ESL students into three groups (1-correction code, 2-errors underlined, 3-control group) to check their abilities to self-edit after receiving teacher's feedback. Both treatment groups considerably outperformed the control group and no significant differences were



noticed between Group 1 and Group 2. Thus, they concluded that although it is still debatable how explicit error feedback should be, teacher's feedback helped to improve students' accuracy over time. Likewise, Chandler (2003) conducted a study on various kinds of error feedback (direct correction, underlining, describing) provided by the teacher. Teacher's feedback resulted in students' improvement on grammatical and lexical errors between assignments. In fact, students felt that their teacher's feedback helped them learn better and not repeat the errors.

Sheen (2007) examined the effect of different types of WCF (direct-only correction group, direct metalinguistic correction group, and control group) on the acquisition of articles by adult intermediate ESL learners of various L1 backgrounds. Both treatment groups outperformed the control group and, specifically, the direct metalinguistic group performed better than the direct-only correction group in the delayed posttests.

In conclusion, despite having been blamed for not offering immediate feedback on some occasions (e.g., Bai, 2012), teacher's feedback helps improve the students' writing competence (Wang & Liu, 2012; O'Mahony et al., 2013). Moreover, it has a significant and influential role on the students' learning process, as it not only helps to learn more and better (Chandler, 2003; Morch et al., 2017), but it also contributes to greater grammatical accuracy and pragmatic appropriateness (e.g., Li & Ye, 2016; Sarie, 2017).

### **2.7.2 Peer feedback**

Peers are another possible figure who may provide corrective feedback in language classes. Peer assessment is the mutual evaluation of the work of classmates who have an equal status. It is defined as "a set of activities through which individuals make judgments about the work of others" (Reinholz, 2016, p. 1). This way of providing feedback is particularly useful in large classes in which the teacher struggles to deal with all their writings. It is not a substitute for teacher's feedback, but a cooperative and collaborative process between learners. Nicole (2010) proposes that "feedback should be conceptualised as a dialogical and contingent two-way process that involves coordinated teacher-student and peer-to-peer interaction as well as active learner engagement" (p. 503). This means that students write a piece of work, read and

assess their peers' pieces, receive their writing corrected and determine which corrections to modify. This sort of feedback is more cognitively demanding for the students because they do not merely rely on and accept their teachers' feedback, but they form an opinion on what is acceptable or not "giving students a better understanding of assessment criteria and leading to deeper learning" (Logan, 2009, p. 30).

Taking into account that receiving feedback is an important step in the learning process, learners may feel reluctant to receive it from peers (Tsui & Ng, 2000), because "some students may feel that their classmates are not qualified to provide insightful feedback" (Liu & Carless, 2006, p. 7). For this reason, Sato and Lyster (2012) reported that

*learners often avoid negotiation and solely focus on task completion, [...] their feedback is usually made up of simple segmentations of their partner's erroneous utterances. This is not quality feedback because it lacks a corrective force. [...] learner's perceptions of one another may hinder the effectiveness of peer interaction" (p. 597).*

On the contrary, the use of peer and self-assessment has been proved to have many advantages. Students are responsible for the whole process and are motivated to understand the assessment criteria and to write for an authentic audience moving from the private domain to a more public one. In this way, the benefits are twofold: first, students are encouraged to reflect on their classmates' work to provide them with feedback, and second, students check the validity of the suggestions on their own writings.

Some studies show how different the reaction can be when interacting with a peer or with a teacher. In the writing skill, Yu and Lee (2014) found that students tried to write more clearly and avoid mistakes with peer assessment to make their writings easily understandable. In the speaking skill, for example, Tulung (2008) discovered that students felt more relaxed and self-confident in oral interaction with other students than with the teacher. Similarly, Sato (2013) investigated learner's beliefs regarding peer feedback and concluded that although anxiety is lowered when they interact with peers, their preference for peer interaction depends on whom they interact with. Other students may experience the same difficulties when writing, and peer feedback may help them to develop self-confidence by reducing their anxiety (Cotterall & Cohen, 2003; Curtis, 2001).

A number of studies exhibit advantages of peer feedback on structure and content, grammatical knowledge internalization, critical thinking competence, active participation, and overall improvement (e.g., Cai, 2011; Forrer et al., 2015; Ion et al., 2016; O'Mahony et al., 2013; Tsui & Ng, 2000; Wang, 2016; Wei, 2015; Yu, 2013; Zhan, 2016; Zhou, 2013). Moreover, De Grez et al. (2012) also reported positive effects of peer feedback on oral presentations. Planas-Lladó et al. (2013) explain that “students perceive it as both a motivating and recommended methodology that facilitates the acquisition of learning at different levels” (p. 2) but that does not change the fact that students see it as a responsibility and they sometimes distrust their fellow students’ abilities to peer-assess. Likewise, the studies by Topping (1998), Villamil and De Guerrero (1997), and Yang et al. (2006) also reported positive effects of peer feedback on achievement and students’ attitudes towards learning and quality of writing, and learner autonomy increased.

Comparing teacher and peer feedback, one of the differences between the teacher’s feedback and peer feedback is that the latter is not authoritative. Yang, Badger and Yu (2006) compared the responses of one group receiving teacher’s feedback and another one receiving peer feedback. Surprisingly, although peer feedback was considered to be less trustworthy, those students were more actively involved in checking the feedback they received. On the contrary, students receiving the teacher’s feedback made more surface changes. Going even further, Cho et al. (2008) proposed multiple peer reviewers instead of a single expert or peer assessor to improve the quality of the piece of work. In addition, they also discovered that peer reviewers tend to praise and use non-directive corrections leading, thus, to improved writings. Cho and MacArthur (2011) also concluded that students who read and reviewed their peers’ writings, wrote higher quality papers on a new topic.

Sato and Lyster (2012) also proved the effectiveness of peer interaction and CF on language development. Four groups of Japanese students received different treatments: CF with prompts, CF with recasts, participation in peer interaction activities and the control group. All the treatment groups outperformed the control group, and particularly, accuracy and fluency were improved by peer feedback. Miao, Badger and Zhen (2006) also compared peer and teacher feedback in written production and stated that although peer feedback can

contribute to learning development, “the research broadly indicates that teacher feedback has a much greater impact than peer feedback, though with considerable variation” (p. 182) because “students said that the teacher was more professional, experienced and trustworthy than their peers. Often peer feedback was not accepted by the writer for the reason that it seemed incorrect to them” (p. 189). The same happened in Connor and Asenavage’s (1994) and Paulus’ (1999) studies. In the first one, Connor and Asenavage researched the impact of peer and teacher feedback on 8 ESL students from different countries. Results showed that teacher’s feedback had much more significance for students than peer feedback. In the latter, Paulus also explored the effect of peer and teacher feedback on 11 ESL students. Results indicate the effectiveness of teacher feedback over peer feedback since 87% of teacher comments produced a change in their production and only 51% of peer feedback.

Nevertheless, although teacher’s feedback seems to be more authoritative and trustworthy, Zhao (2010) explained that whereas teacher feedback triggered more revisions than peer feedback, “only 58% of teacher feedback instances were found to be used with a real understanding of their necessity” (p. 14).

In summary, peer interaction can contribute to L2 learning as it provides students with opportunities to engage in contextualized production practice. However, lack of training to provide CF or students’ L2 proficiency level can hinder the process. In addition, if learners do not trust their classmates’ linguistic ability, peer feedback may be ineffective (Yoshida, 2008). Students may also find themselves uncomfortable grading their classmates and exerting power over them and vice versa (Falchikov, 2001; Isaacs, 2001). All in all, peer feedback is a beneficial tool in SLA, as stated in previous literature (Byrd, 1994; DeGuerrero & Villamil, 1994; Lockhart & Ng, 1995; Min, 2005; Tsui & Ng, 2000; Villamil & DeGuerrero, 1996).

Another possibility of feedback together with peer and teacher feedback is self-correction, which is addressed in the next section.

### **2.7.3 Self-feedback**

Nowadays the tendency seems to be towards educating autonomous learners who can critically modify, correct, assess and/or evaluate their own written production. The challenge is

to be able to detach from one's own text and read it critically. Self-assessment is an important part of the learning process where students develop their critical thinking and metacognitive behaviour (Cooper, 2006). It is defined as a process where "students are directed to assess their performance against predetermined standard criteria ... [and] involves the students in goal setting and more informal, dynamic self-regulation and self-reflection" (Bourke & Mentis, 2011, p. 859). Self-assessment can develop a sense of internal responsibility for their own learning (Yorke & Longden, 2004), since the teacher is sharing his role of assessor/evaluator (Usher & Pajares, 2005). Moreover, learners control their own learning by identifying their own strengths and weaknesses (Cubukcu, 2008).

Self-assessment can be performed in different ways. For example, the use of checklists or self-rating prompts helps students to reflect on their written work by making sure they have followed the task characteristics, such as format, register, vocabulary, structures, etc. (e.g., Clarke, 2005; Clarke, Timperley & Hattie, 2003). Other techniques are the use of rubrics or sample answers for self-marking or grading (Todd, 2002) or for detecting the quality characteristics of the student's composition. These three methods of self-assessment entail self-reflection on one's work. Nevertheless, estimating one's own level of performance by means of rubrics or a score may not always be reliable, as Brooks (2002) claims. Sometimes students overestimate their abilities or do not realise they are not good enough in that field (Dunning, Heath & Suls, 2004; Harris & Brown, 2010). For these reasons, self-assessment is not usually given a high percentage in the final grade, but it serves as part of the students' self-learning and progression.

According to Ross (2006), self-assessment comprises three processes, as illustrated in Figure 6:

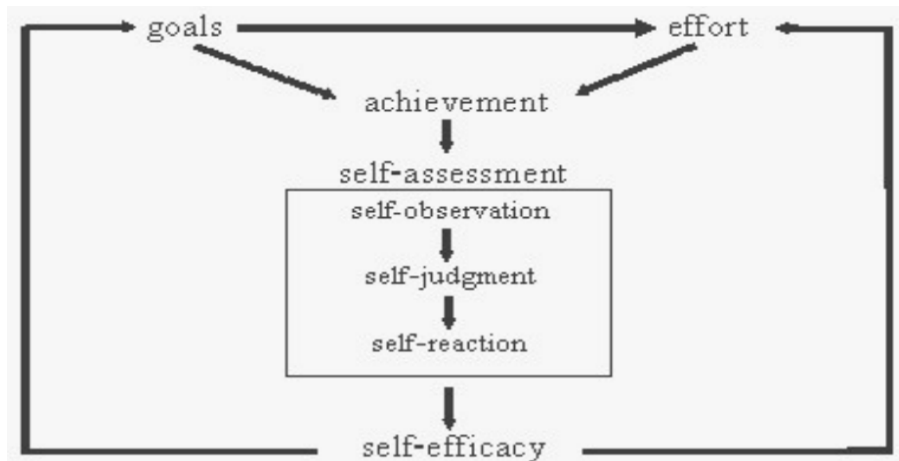


Figure 6. Contribution of self-assessment to learning (Ross, 2006, p. 6)

Students first need to observe their written work based on the established criteria. Then, they judge their own work to determine if they met those standards. Finally, once this process has finished, students react by feeling satisfied or disappointed. Self-assessment plays a role in the students' persistence and willingness to continue progressing by making them aware they need to make a greater effort or they are able to meet the standards and be successful in subsequent writings (self-efficacy). On the contrary, a feeling of frustration may also emerge in this process.

To conduct self-assessment correctly, the students need clear instructions (e.g., approach, marking criteria, schedules, etc.) in order not to feel frustrated during the correction/revision process. Ross (2006, pp. 8-9), for instance, suggests the following steps:

1. defining the criteria by which students assess their work
2. teaching students how to apply the criteria
3. giving students feedback on their self-assessments
4. helping students to use self-assessment data to improve performance

There are many studies that have proved the advantages of self-assessment (e.g., Andrade & Boulay, 2003; Andrade & Valtcheva, 2009; Dochy et al., 1999; McDonald & Boud, 2003; Mentis, 2011; Rolheiser & Ross, 2000; Topping, 2003) to identify the students' strengths and weaknesses, to develop their critical thinking (Cooper, 2006), to encourage them to be

responsible in their own learning process (Cyboran, 2006) and to improve future work. However, measuring how effective self-assessment is results in a complex task. While it may create more reflective and autonomous students, its contribution to L2 learning is an intricate debate. What is clear is that although peer and self-assessment are hardly used in the classroom, the former nourishes the latter. When receiving peer feedback, students are the ones who make the ultimate examination in the form of individual reflection or self-assessment.

Finally, the next section will deal with a type of feedback in which human interaction does not play a role. We are referring to automated or computer-generated feedback, the type used in the research carried out in this dissertation.

#### **2.7.4 Automated feedback**

Aspects such as lack of enough practice in the classroom, high ratio of students or self-learning methodologies have given rise to the development of sophisticated software (e.g., *Criterion*, *Grammarly*, *Grammar Checker*, etc.) which is based on extensive corpora and generates immediate feedback and even a score. Ellis Paga is regarded as the pioneer of *Automated Essay Scoring* as he designed the so-called *Project Essay Grader* in 1966, a computer program to grade writings. Broadly speaking, these computer programs provide students with a tool to receive immediate individualised and extensive feedback in order to improve their writing skill, relieve the teacher from the amount of work and promote autonomy among students. In fact, they can be used as extra tools to receive CF (Bull & McKenna, 2004).

Despite the fact that some studies (e.g., Attali, 2004; Coniam, 2009; Hutchison, 2007; Shi, 2012; Yang, 2004; Yang & Dai, 2015; Wei, 2015; Zhou, 2013) show how automated feedback helps learners in their learning process, others report the opposite or mixed results (Lai, 2010; Lee et al., 2009; Tuzi, 2004). Although automated feedback has been increasingly used in higher education, there is still a grinding reluctance to automated feedback. The common problems of automated feedback are lack of accuracy and surface correction since they do not usually notice content or cohesion mistakes and decrease students' active role.

In Rezaee and Ahmadzadeh's (2012) words, "computers have become an inseparable part of everybody's life. By far, their roles in education, especially in language learning and

teaching, have expanded so drastically that no language instruction can ignore them in its curriculum” (p. 346). Using computers to deliver feedback has spread and is still in full development. Online courses and distance learning are commonplace but certain aspects are still in need of improvement (e.g., quality of Internet connection, quality programs, lack of resources, etc.). According to AbuSeileek (2012), computer-mediated CF may support students in the development of their writing performance when receiving CF.

Our aim in this last section of Chapter 2 was to present an overview of different agents who can provide feedback. Being automated feedback a key element in the present dissertation, a deeper review of computer-generated feedback, along with an analysis of the main computer programs available nowadays, will be discussed in Chapter 3.

## 2.8 Chapter Summary

After having dealt with issues such as the different types of CF that students can receive, the different agents that can provide that feedback and their effectiveness, conclusions are not clear-cut and further research is still necessary to solve the feedback conundrum. The choice and implementation of CF are not arbitrary decisions, since they are influenced and determined by the circumstances (e.g., students’ abilities, preferences, attitude, timing, level, methodology, etc.). By understanding how students learn and use CF and taking into account the characteristics of the group, the effectiveness of CF may be maximised.

As explained in this chapter, students can be assisted through feedback and how to employ it in their learning should be studied in depth. Writing is seen as a process rather than an end-product (Kamal & Faraj, 2015), where CF has a crucial role to enhance student learning. Thus, teachers must be aware of the different corrective strategies if they want to aid their students (Livingstone, 2015). However, although sometimes they are, they may not know how to optimise them in the classroom (Caranza, 2007).

Once the term CF and what it involves have been analysed, the next chapter will focus on Computer-Assisted Language Learning (CALL, henceforth) and, particularly, on the most relevant computer-generated feedback programs, a crucial issue in the present dissertation.



## **Chapter 3. Computer Technology Applications in Language Learning**

Although it evolved in the 1970s, the origin of the term Information Technology (IT) can be traced to World War II when the military and the industry worked together in the development of electronics and information. IT represents any technology that helps to create, manipulate, collect and distribute information. As time went by and with the advent of electronic communication, the term IT developed into ICT (Information and Communications Technology).

Over time, the use of new technologies has increased due to different factors, for example, its possibilities in different fields (e.g., economic, social, environmental and educational), and access to the Internet and computers. Without forgetting all their advantages, e.g., interactivity, immediacy, automaticity, adaptability, innovation, lifelong learning, etc., one of their main applications is in the language learning field. In fact, nowadays most teachers use them in their classes and little by little they are better trained and feel self-confident to use new technologies as an aid in the teaching/learning process.

According to Taylor (1980), computers can have different roles: computers as tutors, computers as tools and computers as tutees for students. In the 'computers as tutors' option, the computer presents the material (e.g., grammar, listening, pronunciation and reading exercises) and evaluates the answers to determine the next step. In 'computers as tools', the computer assists students to do the task (e.g., grammar checkers). In 'computers as tutees', it is the student who teaches the computer by programming it. Therefore, taking into account Taylor's (1980) classification, in the present dissertation, the computer will be used as a tool for language learning.

The benefits of using ICTs in language learning are countless and may contribute as a type of feedback by helping students to regulate their own learning (e.g., Cabero, 2001; Mooij, 2009; Nicol, 2006). Learners become an active entity who uses the language in a meaningful context by cooperating and interacting autonomously. All this increases learners' motivation, concentration, cognitive development and critical ability. Therefore, taking into account the importance of ICTs

in language learning, the next section will specifically focus on the use of the computer in the language learning process.

### 3.1 Computer-Assisted Language Learning

The use of new technologies has contributed and is still contributing positively to many fields such as medicine or science, and also education. It is not surprising they have revolutionised the education sector as it represents a way to improve the students' language competencies by adapting to their learning needs and styles, increasing their autonomy, creativity and productivity. Interactive materials, presentations and online assessment are some of the possibilities they offer. Interaction -either with the teacher, peers or a computer- is "a necessary and fundamental mechanism for knowledge acquisition and the development of both cognitive and physical skills" (Barker, 1994, p.1). With all the technological advances computers are ready to deal with this interactive aspect of the learning process.

The use of computers and the Internet in education offers a different perspective of the classroom environment, the role of teachers and students, the way of working and interacting, methodologies, materials, etc. CALL is a term that has evolved over time and been defined as:

*the search for and study of applications of the computer in language teaching and learning (Levy, 1997, p. 1)*

*any process in which a learner uses a computer and, as a result, improves his or her language (Beatty, 2003, p. 7)*

*learners learning language in any context with, through, and around computer technologies (Egbert, 2005, p. 4)*

The first definition focuses on the computer as the tool to teach and learn a language. However, in the second and third definitions not only is the idea of learners learning a language emphasised over the use of the computer but the relation between it and language learning also becomes relevant. As Warschauer (1999, n. p.) states, "the use of computers should be an integral aspect of the language learning process, as we live in times of rapid change and technology is constantly evolving."

Computers can function as a tutor to guide students, as a stimulus to develop critical thinking, as a tool to develop other tasks such as writing, and as a medium of communication. However, as pointed out by Garrett (1991), “the use of the computer does not constitute a method” but a “medium in which a variety of methods, approaches, and pedagogical philosophies may be implemented” (p. 75). In other words, computers offer opportunities to enrich language learning but are ineffective if misused.

CALL started in the 1960s and has developed gradually in three phases: behaviouristic, communicative and integrative (Warschauer, 1996). Behaviouristic CALL, based on the main behaviourist theories of learning, was implemented in the 1960s and 1970s when audiolingualism was influential and drilling commonplace. The computer had the role of a tutor working as a deliverer of resources to students. An example of this CALL tutoring system was *PLATO*, which was designed by the University of Illinois and included chats, forums, online testing, vocabulary and grammar drills and multiplayer video games, among others. In the 1970s, when drilling started to be seen as insufficient and not authentic, communicative CALL flourished. Communication in an environment where the use of the target language, forms, original utterances, and implicit grammar teaching were prominent, ruled this phase. The computer keeps being the know-it-all but it is the students’ responsibility to find the answer through critical thinking and reflection. Some of those programs are *SimCity* or *Sleuth*, along with other spelling and grammar checkers.

Along with this use of computers, authors such as Kenning and Kenning (1990), Pusack and Otto (1990) and Rüschoff (1993) have tried to use computers in a more integrative way (e.g., projects). The integrative phase is based on two main technological developments: computers and the Internet. Regarding the use of computers, language learners can easily communicate asynchronously by email or by means of synchronous programs or applications such as Skype or WhatsApp with peers or even native speakers from anywhere at any time (i.e., computer-mediated communication).

As for the Internet, all the multimedia resources students can find are easily accessible. This is referred to as hypermedia and helps to create an authentic learning environment where students learn at their own pace and where the different skills are all integrated. An example

of this technique is *Dustin*, a program designed by Northwestern University. It simulates the different situations which may be experienced at an airport. The student has to role play different scenarios to continue with the next one. The main drawback is that those programs are not 100% interactive because they still have some limitations; moreover, they are expensive and not all teachers have the required knowledge to use them.

The presence of technology in language teaching and learning is considerably notorious, which has impacted on the amount of literature in the field of SLA. Figure 7 represents how the learner interacts through the computer with the teacher, peers, materials and others to fulfil the learning objectives. That is to say, computers act as mediators or facilitators to achieve learning goals.

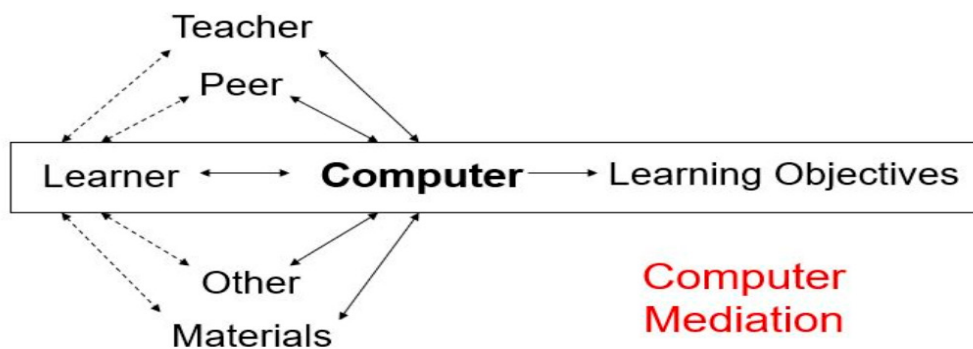


Figure 7. A simple conceptualisation of the CALL perspective. (Levy & Hubbard, 2005, p. 146)

Within the field of CALL, ICALL (Intelligent Computer Assisted Language Learning) is the focus of much attention for its role on error detection (Heift & Schulze, 2007). The difference between more traditional CALL tools and ICALL is that the latter applies Artificial Intelligence (AI) based on algorithms and technologies (Gamper & Knapp, 2002; Heift & Schulze, 2007; Nerbonne, 2003; Schulze, 2008). Obviously, the use of AI applied to the educational field is a revolutionary transformation. In regard to the main focus of this study, i.e., written expression, automated feedback systems are a form of ICALL that can identify errors, suggest solutions and even provide corrections based on parsers, algorithms and corpora. While CALL systems offer feedback on superficial errors (e.g., spelling and grammar), ICALL goes beyond surface errors and addresses more complex ones such as meaning, content, register, etc. In the past decade, a

number of automated feedback systems were designed, for instance, *E-Tutor* (Heift, 2003) and *The Sentence Fairy* (Harbusch, Itsova, Koch & Kuhner, 2008) for German, *Grammar Checker* (Lawley & Chacón, 2003) and *WERTi* (Meurers et al., 2010) for English, *ROBO-SENSEI* for Japanese (Nagata, 2002), and *TAGARELA* for Portuguese (Amaral, Meurers & Ziai, 2011).

AI has also made it possible to design automated writing evaluation tools that not only provide feedback but also a score. They combine natural language processing (NLP), statistics and AI to evaluate written discourse. As an illustration, we can mention *BETSY* (Rudne & Liang, 2002), *Criterion* (Burstein et al., 1998), *Intelligent Essay Assessor* (Landauer, Foltz & Laham, 1998), *TechWriter* (Napolitano & Stent, 2009), and *Writing-Pal* (Dai, Raine, Roscoe, Cai & McNamara, 2011).

In conclusion, with the continuous growth of the use of multimedia for language learning, applications for mobile phones and social media, among others, the options to implement CALL and benefit from both in and out of the classroom are unlimited. However, thanks to AI techniques a new dimension was added, ICALL. With ICALL the learning process is learner-centred and the computer adopts the role of the teacher. Focusing on the role of feedback and working on the principle that it is beneficial to identify errors and provide feedback (Heift, 2002; Nagata, 1993), ICALL can be seen as the solution to numerous writings waiting to be corrected and even assessed and returned to students. However, although ICALL is a dynamic field of interest that provides innovative language-learning tools for the educational community, reliance on automated feedback systems is still a controversial issue. For this reason, a deeper analysis of the main automated writing systems will be explained in Section 3.2.

The next two subsections will examine the ways in which a computer can be part of the process of feedback provision.

### **3.1.1 Electronic feedback**

As one of the cutting-edge inventions of the 21st century, computers have altered the way life is understood. Especially in the education domain, the importance of technology and the benefits they seem to provide to the learning process show how it is taking over classrooms at all levels. As Sadler (2010) points out, “feedback should not only be of an appropriate

type but also be provided within the available resources” (p. 536). The computer is one of these resources. Thus, its use has widely spread in schools and universities and submitting papers electronically, through forums, chats or learning platforms (e.g., *Moodle*, *Aules*) is commonplace. These papers can be corrected online, what has been termed as “electronic feedback” or “e-feedback”, that is to say, “feedback in digital, written form and transmitted via the web that transfers the concepts of oral response into the electronic arena” (Tuzi, 2004, p. 217). The provision of this kind of feedback is no longer unusual (Elola & Oskoz, 2017) and should not be confused with computer-generated feedback or AWE, which will be explained in Section 3.1.2.

As shown in Table 4, e-feedback differs from traditional oral and written feedback in a number of ways.

Table 4. General differences between oral, written and e-feedback (Tuzi, 2004, p. 219)

| Oral feedback                | Written feedback                  | E-feedback                     |
|------------------------------|-----------------------------------|--------------------------------|
| Face-to-face                 | Face-to-face/distant              | More distant                   |
| Oral                         | Written                           | Written                        |
| Time dependent               | Depends                           | Time independent               |
| Pressure to quickly respond  | Pressure to respond by next class | No pressure to quickly respond |
| Place dependent              | Depends                           | Place independent              |
| Nonverbal components         | No nonverbal components           | No nonverbal components        |
| More personally intrusive    | Depends                           | More personally distant        |
| Oral/cultural barriers       | Written/cultural barriers         | Written/cultural barriers      |
| Greater sense of involvement | Greater sense of involvement      | Greater sense of anonymity     |
| Negotiation of meaning       | Negotiation of meaning            | Less negotiation of meaning    |
| Less delivery effort         | Greater delivery effort           | Less delivery effort           |
| N/A                          | No cut & paste                    | Cut & paste                    |

Taking into account aspects such as time, with e-feedback students can take part in synchronous communication through chats and in asynchronous communication by email, forums or Word comments. The difference is that in asynchronous communication students cannot revise their writings until they have received their corrections electronically. In addition, nonverbal elements in e-feedback are missing, which may make deciphering and effective correction more complex. On the contrary, in traditional oral feedback students and teachers or peers communicate verbally and nonverbally synchronously. As for written feedback, teachers or peers read the text and provide their written response on the same paper either synchronously or asynchronously.

E-feedback through computer conversations allows students “to respond spontaneously [and] to reflect on their ideas, rehearse their responses, and respond at their own pace”

(DiGiovanni & Nagaswami, 2001, p. 269). On the contrary, when time is not shared, spontaneity disappears and the pressure to answer decreases. Lastly, with e-feedback paperwork problems and written-work retrieval or loss is no longer an obstacle. In this way, the teacher makes sure all the students receive their corrections.

Research on the impact of e-feedback on revision is scarce and sheds no conclusive findings. For instance, e-feedback may have no impact simply because teachers are reluctant to use it. In this sense, Lu and Powell (2004) proposed a methodology to discover the barriers that influence EFL teachers in the use of computers in their classroom: (1) technology skills, (2) funding, and (3) the acceptance of technology. Some other authors have also referred to the quality of the instructional software and the high cost (Bani Hani, 2009), technophobia (Lam, 2000; Thelmadatter, 2007), time pressure both outside and during class (Lam, 2000; Levy, 1997; Smerdon et al., 2000; Strudler et al., 1999), lack of resources and materials (Loehr, 1996), insufficient or inflexible guidelines, standards, and curricula (Langone et al., 1998), lack of support or recognition for integrating computers (Grau, 1996; Strudler et al., 1999), a clash between new technologies at universities and older ones in schools and lack of leadership (Smerdon et al., 2000) and inadequate training and technical support (Lam, 2000; Langone et al., 1998; Levy, 1997; Smerdon et al., 2000). Those barriers may make teachers choose alternatives to e-feedback. Regarding students, Tuzi (2005) reported that although they preferred face-to-face communication, online written feedback from peers through chat rooms was more beneficial than oral feedback.

There are studies which show positive effects of electronic feedback on uptake (e.g., Ene & Upton, 2014; Schultz, 2000; Tuzi, 2004). Some of the conclusions reached by Ene and Upton (2014) in their research about the relationship between teacher electronic feedback and uptake showed that the overall rate of successful uptake was high (62.3%), the highest rate was in response to focused grammar (75%) and there was no difference between explicit and direct teacher electronic feedback and uptake.

Schultz (2000) compared face-to-face peer feedback and computer-mediated peer feedback provided by intermediate and upper-intermediate French students. After analysing all the essays and conducting 106 surveys, results show that students in the face-to-face group

suggested more general modifications to facilitate interaction and reflection on the writer's intentions. On the contrary, students in the computer-mediated group made more specific corrections as they had the possibility to save and pay attention to the suggestions. Similarly, Tuzi (2004) analysed the impact of e-feedback on the revisions made by first-year university language students. They made more revisions in response to e-feedback than oral feedback, especially at the clause, sentence, and paragraph levels.

In conclusion, despite the scarcity of studies on e-feedback and specially teacher e-feedback on L2 writing, it has been reported to assist students, (e.g., Ducate & Arnold, 2012; Elola & Oskoz, 2016; Ene & Upton, 2014; Honeycutt, 2001; Sauro, 2009; Schultz, 2000). However, there are also voices that claim it can have negative effects (Blake & Zyzik, 2003; Guardado & Shi, 2007; Ho & Savignon, 2007; Schultz, 2000) for being slow, focusing on form more than discourse and not encouraging interaction or deep revision.

E-feedback can make the correction process easier and faster but what has really revolutionised the writing correction realm is the use of computer-generated feedback, which is the focus of the next section.

### 3.1.2 Computer-generated feedback

As part of ICALL, computer-generated feedback, in contrast to e-feedback, provides automated feedback obtained from algorithms drawn from a database or corpus previously created (Heift, 2001, 2004, 2010; Warden & Chen, 1995; Ware, 2011). Although there is some debate about the terminology (Automated Essay Scoring, Automated Essay Evaluation, and Automated Essay Grading) we will use Automated Writing Evaluation throughout the study, as it is the most commonly used term in the literature. The first AWE programs were developed in the early 1960s when Ellis Page developed the Project Essay Grade. Since then, more advanced scoring engines have been launched into the market, such as the *E-rater*, developed by ETS, *IntelliMetric*, developed by Vantage Learning, and *Intelligent Essay Assessor* by Pearson Education, and incorporated later on into tools such as *MyAccess!* and *Criterion*.

AWE programs aim at saving time, money and increasing reliability in the assessment of writing (Hamp-Lyons, 2001; Rudner & Liang, 2002; Shermis & Burstein, 2003; Warschauer



& Ware, 2006). In this line, workload in schools is a hindrance that affects both teachers and students and although this pressure may be experienced differently by each individual, it is clear that the lack of time is the most relevant one when dealing with teachers' work lives (Apple & Jungck, 1996; Bartlett, 2002; Levin & Riffel, 2000; Vandenberghe & Huberman, 1999). In some educational contexts, the time available is scarce when teachers have to prepare their lessons by using updated materials and realia as much as possible, evaluate their students' essays providing accurate feedback, attend to students' demands and needs, devote time to take control over longer-term planning, reflect on the teaching practice, evaluate students' individually and thoroughly, and cooperate among workmates. It does not only occur during the workday but when arriving home too where the workload seems to continue. Lindqvist and Nordänger (2002, p. 2) emphasised the idea of the "uncompleted character of teachers' work".

Teachers regard their students as a responsibility which depends on them. For this reason, they usually experience feelings of vulnerability, powerlessness, frustration, anger and fear (Kelchtermans, 2005; Nias, 1996). Teachers feel the need to meet their students' demands, and that intensifies the feeling of pressure (Campbell & Neill, 1994; Nias, 1999).

In Gitlin's (2001) words, "For too long now, the working conditions of teachers have been overlooked because it is assumed that teaching is a calling, a profession where one would work and overcome school-related obstacles regardless of their nature" (pp. 254–255). This is the reason why teachers need to rely on external resources such as textbooks, web pages and computer programs, among others, to satisfy their students' needs. Thus, the idea of integrating technology via computer-generated feedback tools can contribute positively to the revision process by providing constructive feedback and reducing the teacher's pressure (Sundeen, 2015). It is true that many teachers have never implemented automated feedback generators (Reiners et al., 2011) and may be reluctant to use them, because the way humans and computers assess aspects such as organisation or coherence may be different. However, these kinds of tools are not necessarily thought to substitute the teacher's feedback but to help them to relieve their workload and to support their students. In fact, its use in the classroom has provoked heated debate as to its effectiveness.

These programs can provide individualised feedback on specific aspects such as grammar or spelling for ESL students (Bolt, 1992; Liou, 1994; Warden & Chen, 1995), as well as on more holistic elements such as content and organisation for L1 and L2 students (Brock, 1993; Burston, 2001; Ferris, 1993; Leacock, 2004). In addition, according to AWE developers, their programs are able to assess and respond to student writing as human readers do (e.g., Attali & Burstein, 2006; Vantage Learning, 2007). However, research into the implementation shows inconclusive results since some studies report favourable results (Coniam, 2009; Hutchison, 2007) while some others do the opposite (Lai, 2010; Lee et al., 2009; Tuzi, 2004). Anson (2006) and Herrington and Moran (2001), for example, distrust how those programs actually read the students' writings or how they redress aspects such as organisation, content, and meaning. Most AWE programs fail to provide specific feedback on the social and communicative dimension, since they are not sensitive enough to identify features such as irony, humour or sarcasm. In Ericsson and Haswell's (2006) words, machine-scoring programs:

*'see' student essays as code. They take students' words, sentences, and paragraphs out of their social/cultural contexts, process them as meaningless 'bits' or tiny fragments of the mosaic of meaning, and claim to have 'read' these essays for meaning. They claim to be able to 'mimic' the way a human reader would read them. And they base these claims on uninformed, possibly fraudulent, understandings of meaning (p. 36).*

The Conference on College Composition and Communication in the U.S., in fact, disagrees with the use of AWE programs for assessment purposes and claims that:

*While they [AWE programs] may promise consistency, they distort the very nature of writing as a complex and context-rich interaction between people. They simplify writing in ways that can mislead writers to focus more on structure and grammar than on what they are saying by using a given structure and style (CCCC, 2006).*

As a result, computers may not fully assess the students' writing ability but merely surface aspects (Condon, 2013; Neal, 2011). That means students can cheat those programs and obtain high scores by giving more attention to those elements that are more easily detected by AWE systems (Herrington & Moran, 2001; Powers et al., 2002; Ware, 2005). For instance, a student can create a nonsensical text based on essay prompts that can obtain top scores. Thus,

AWE programs are useful only for the revision of formal aspects (Cheville, 2004; Grimes & Warschauer, 2006; Yang, 2004; Yeh & Yu, 2004) and for the provision of non-specific and general comments that may be of limited help to students. That is, when the focus is on meaning and content over form, the effectiveness of AWE scoring becomes dubious. What is more, higher scores can also be achieved through avoidance strategies (e.g., El Ebyary & Windatt, 2010). Warschauer and Grimes (2008) found that students resorted to making minor changes such as correcting spelling or grammar to improve their scores instead of addressing more complex issues like content.

Cheville (2004) also pointed out that, as computers cannot notice the sociolinguistic context where some vocabulary or expressions are used, they may take those choices as errors when, in fact, they are appropriate and vice versa. Furthermore, some automated systems focus on the sort of errors made by NS (Burstein et al., 2003), and due to the fact that NNS may make other kinds of errors, the features used to score the writings may not be effective for NNS.

In addition, features such as computer literacy or proficiency can also influence the students' perception students have of these programs. Yang (2004), for example, noted a less positive response to automated feedback from more advanced students than less proficient ones. As automated systems are generally form-focused, more advanced students may not need so much feedback on form but on meaning and content, whereas less proficient students may need more help with formal aspects of the language such as grammar, spelling or punctuation.

Undoubtedly, the main disadvantage of automated feedback in comparison to teacher's feedback is that the social and communicative dimensions are secondary (Ericsson, 2006). On the contrary, some other studies on AWE programs show a more positive view. Grimes and Warschauer (2006), for instance, found two main benefits of using *MyAccess!* and *Criterion* in U.S. high school writing classes. The teachers' management of the classroom was less arduous and students became motivated to practise writing due to the speed of responses. Even though both teachers and students had a positive view on those programs, they devoted little time to using them due to time constraints.

In line with the above said, AWE programs may offer some advantages. They provide immediate feedback (Dikli, 2006), which aids students in the following ways: firstly, they do not have to wait for feedback for too long to make any necessary modifications. Secondly, AWE programs offer multiple drafting opportunities for students (Warschauer & Ware, 2006), which may motivate them to keep trying and improving their written work. Thirdly, these programs adapt to the learning of the students letting them control their own pace by eliminating any pressure or anxiety. Fourthly, they allow for more students' autonomy and critical thinking (Cheng & Cheng, 2008). In other words, based on the feedback received by the computer, students take the decisions and look for solutions to polish their writings. In sum, as Becker (1990) stated, "in the 1980s, no single medium of instruction or object of instructional attention produced as much excitement in the conduct of elementary and secondary education as did the computer" (p. 385).

In view of the growing interest in automated feedback, some research has been carried out in an attempt to compare the feedback and/or scores provided by humans and by computers. For instance, Coniam (2009) investigated the correlation between a computer essay scoring program (*Betsy*) and human raters' scores. Results show that "while computer essay-scoring programs may appear to rate inside a 'black box' with concomitant lack of transparency, they do have potential to act as a third rater, timesaving assessment tool" (p. 259). So far, despite high correspondence with human raters (e.g., Coniam, 2009; Dikli, 2006; Keith, 2003; Phillips, 2007), computer programs still lack transparency regarding the way they score the writers' work. Similarly, Lee et al. (2009) described "a web-based essay critiquing system developed by the authors [...] to provide students with immediate feedback on content and organisation for revision" (p. 57). Their twenty-seven participants were divided into two groups: an experimental group receiving feedback from the web-based system and a control group who typed and revised their essays on the computer in a traditional way. Results show there was no statistically significant difference in length and final scores between the two groups.

Despite the positive aspects mentioned before, students do not make the most of AWE feedback. Attali (2004) found that only a small percentage of students revised their writings

after receiving AWE feedback. Similar results were found by El Ebyary and Windeatt (2010), who realised that few trainees made use of the feedback provided by the computer and less than half of the respondents engaged in a second draft. In the case of Link et al.'s (2020) study with two SL writing classes (AWE plus teacher feedback, and teacher-only feedback), findings revealed that students tended to revise the teacher's lower-level feedback more often than the one provided by the computer. However, the students' long-term retention was higher when they had access to AWE in comparison to only-teacher feedback.

Based on the studies reviewed here, it can be stated that although there are studies that prove the effectiveness and usefulness of AWE programs (e.g., Chodorow, Gamon & Tetreault, 2010; Cotos, 2011; Grimes & Warschauer, 2006) by helping and motivating students to revise and improve their texts, further research on its lasting effects and improvement at the content and discourse level are necessary. Therefore, taking into account the benefits they offer, they can work as a supplement and not necessarily as a substitute for teachers (Burstein, Chodorow & Leacock, 2003; Burstein & Marcu, 2003; Shermis & Burstein, 2003; Ware, 2005; Warschauer & Ware, 2006). Relying too much on them may, on the one hand, force students to make all the corrections suggested by the program and even ignore other more relevant aspects in the text; on the other hand, these programs make students distrust any corrections taking into account that AWE software sometimes fails to provide accurate feedback. On a positive note, AWE programs present the possibility to alleviate teachers from hours of written comments and students benefit from immediate comprehensive feedback (Hyland & Hyland, 2006). However, whether they lead to much improvement in students' revising skills remains a debate (Stevenson & Phakiti, 2014).

As illustrated above, AWE systems have their own characteristics and may assist learners differently. For this reason, after having mentioned their main advantages and disadvantages, some AWE programs will be further analysed in the next section.

### **3.2 AWE Programs**

AWE programs, as it has been explained in the previous section, compare written texts to a large database which analyses quantifiable features such as grammar, spelling, length, and punctua-

tion, among others. These programs can also score the texts through algorithms that take into account the analysis of the previous features mentioned. Some programs provide the correct answer by the click of a mouse and some others offer suggestions for improvements.

AWE programs can be used for formative and summative assessment and writing instruction as a complement for the classroom but not necessarily substituting the teacher. For example, with the introduction of the Common Core Standards in the USA, Pearson's *Intelligent Essay Assessor* graded around 34 million student essays for state and national tests in the USA in 2017 (Smith, 2018). By having some common standards, the use of computer-based testing solutions becomes feasible. Thus, it seems that computer-generated feedback has come to stay. With its advantages and caveats, before making use of an AWE system, its validity in the classroom and for our students needs to be discussed (e.g., used as a substitute or complement, for NS or NNS, what the learning objectives are, etc.).

In the next sections, some AWE systems (e.g., *Microsoft Word*, *MyAccess!*, *GrammarCheck*, *Ginger Grammar Checker*, *Grammarly*, *Criterion*, *Grammar Checker*, and *IADE*), are going to be analysed in order to justify why one of them (i.e., *Grammar Checker*) has been chosen for the present study.

### 3.2.1 *Microsoft Word*

*Microsoft Word* is a word processing software which was released in 1983 and developed by Microsoft. It is a component of Microsoft Office, although it can also be purchased as a standalone software. *Microsoft Word* 2019 is the latest version and together with Google Docs it is the most popular word processing in the world. It allows text manipulation by editing, copying, pasting, embedding screenshots, etc. Moreover, it has some built-in tools in different languages such as a spell checker, a thesaurus, and a dictionary. Proofreading one's own text becomes fast and easy with *Microsoft Word* because it underlines spelling errors in red and grammar errors in blue, and by clicking the right button of the mouse it automatically proposes the right version. Editing and correcting errors becomes a motivating task, but it also restricts the student's scope overlooking other sorts of errors such as style or cohesion or. The disadvantage for EFL students is that *Microsoft Word* has been created for native speak-

ers, that is to say, some of the corrections suggested may not be appropriate for nonnative speakers of English (Lawley, 2003). Other disadvantages refer to the omission of some spelling corrections of words that exist but are not correct in a specific context, or the split-second correction which may lead students to commit the same error in the future as they may have not even noticed they have been corrected. Figure 8 is an example of *Microsoft Word* 2019:

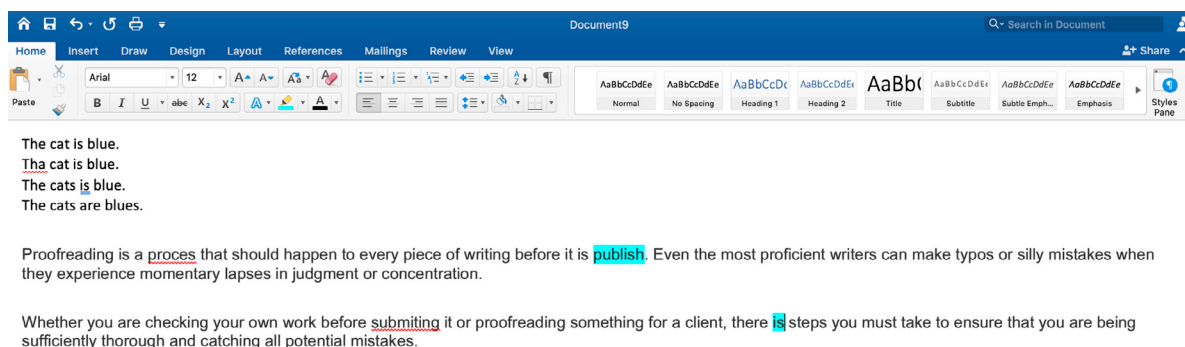


Figure 8. *Microsoft Word* (<https://www.microsoft.com/es-es/microsoft-365/word>)

In the above figure, this software has underlined spelling errors in red and a grammar error in blue, but it has overlooked the other two grammar errors marked in blue. *Microsoft Word* seems to be rather reliable at the orthographic level, but not so much at the syntactic level, as claimed by Fiddienika (2020). In turn, Galleta et al. (2005) found that when *Microsoft Word* identified errors correctly, both high and low ability students equally benefited. When the software flagged errors incorrectly, both types of students made wrong corrections, thus showing the high dependence students have on automated feedback.

### 3.2.2 *MyAccess!*

*MyAccess!* is a web-based AWE program which uses the IntelliMetric® automated essay scoring system developed by Vantage Learning. The scoring system has been designed based on a considerable collection of pre-scored essays scored by human raters. It provides users with immediate scores on a 4-point or 6-point scale that the teacher/administrator can choose. The teacher or administrator can also choose a holistic or domain score. The holistic score

measures the writing as a total work and is based on previously graded essays. The domain or analytic score separates the writing into 5 different traits and analyses them individually (focus and meaning, content and development, organisation, language use, voice, and style and mechanics and conventions). It also includes writing assistance tools such as *My Editor*, *Thesaurus*, *Word Bank*, *My Portfolio*, *Writer's Checklist*, *Writer's Guide*, *Graphic Organizers*, and *Scoring Rubrics*, which help the user to identify spelling and grammar mistakes, provide suggestions to correct those mistakes, offer synonyms and store previous writings to view their progress.

*MyAccess!* can be used for formative and summative assessment. For formative assessment, students can revise and edit their writings as many times as they desire. However, for summative evaluation, they can only submit it once for which they will receive a score.

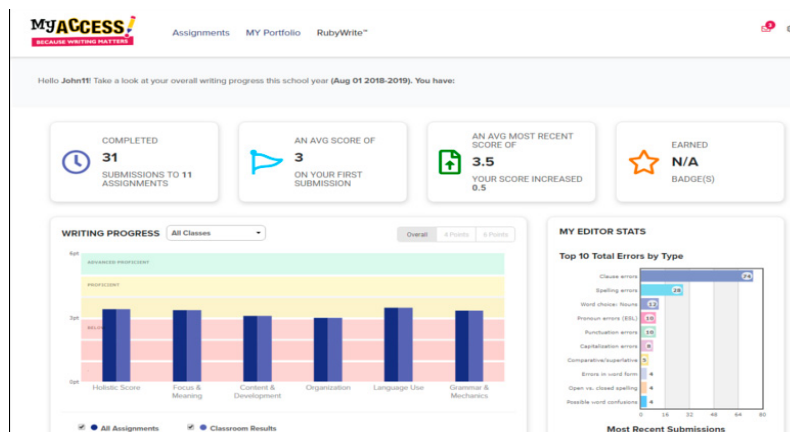


Figure 9. *MyAccess!* (<https://www.myaccess.com/myaccess/assets/newui/images/ma.promo-dashboard.png>)

Studies on *MyAccess!* show contrasting results. On the one hand, Elliot and Mikulas (2004) show it is an effective tool to write better revisions. Similarly, Yeh et al. (2007) found that *MyAccess!* showed improvement of textual quality in focus and meaning. On the other hand, studies such as Yu and Yeh (2003) and Yang (2004) have found some limitations in the feedback provided for being insufficient and vague. What seemed to be useful in the first revisions became not so useful after receiving the same comments several times.

In Chen and Cheng's (2008) study, learners were partly satisfied with the program because they received instant feedback but did not agree so much with the score provided and the



accuracy of feedback (e.g., only formulaic and generic information). In fact, they reported that this program reduced creativity and would be better used by lower levels of proficiency. In addition, *MyAccess!* did not recognise aspects such as humour and irony and it could easily be fooled by students because it praised length over content (Herrington, 2001). Grimes and Warschauer (2010) found *MyAccess!* encouraged students to write but they made little revision on content and organisation, probably because they wrote for a computer instead of a real audience. Although they perceived automated feedback as useful to identify mistakes, the support of the teacher was necessary for interpreting part of the feedback and prioritising it. For this reason, they regarded *MyAccess!* as a tool for writing drafts followed by the teacher's assessment rather than an AWE instrument.

Chou, Moslehpour and Yang (2016) also analysed the effectiveness of *MyAccess!* for self-correction of pre-intermediate EFL students' writings. The findings reveal students improved "in various writing traits in revisions [...], were more capable of self-correcting in usage type of lexical and syntactic errors [but] relatively incapable of handling independently mechanics and style types of errors" (p. 145). Despite the positive effects of this program for independent revision and correction, they also stressed the teacher's role as far as writing instruction is concerned.

### 3.2.3 *GrammarCheck*

This program is a free basic proofreading tool that needs no software installation so texts are never cached in their server. It offers immediate suggestions for spelling, grammar, punctuation and style errors so that students can make corrections. When they enter their text and click the Free Check button, they receive colourful prompts (red for spelling errors, green for grammar suggestions and blue for style suggestions) that help them to correct the errors at the click of a mouse. In addition, it checks the text for plagiarism. If students press the Deep Check button, the system redirects users to *Grammarly* to detect more difficult-to-spot mistakes such as run-on sentences (see Section 3.2.5).

This automated system suggests corrections only at surface level, so it does not provide any feedback on, for example, content or style nor a score. However, it can be very useful to

proofread the students' writings before delivering them to the teacher. Figure 10 shows the interface of *GrammarCheck*.



Figure 10. *GrammarCheck* (<https://www.grammarcheck.net/editor/>)

### 3.2.4 *Ginger Grammar Checker*

*Ginger Grammar Checker* is a free proofreading tool that uses patent-pending technology. Once installed on the computer, students only have to enter their texts and grammar and spelling errors and misused words are corrected. It intends to improve the text in the same way a human reviewer would do. However, it only accepts short texts and with a single click mistakes are corrected without offering any suggestions. Figure 11 shows the interface of this program.

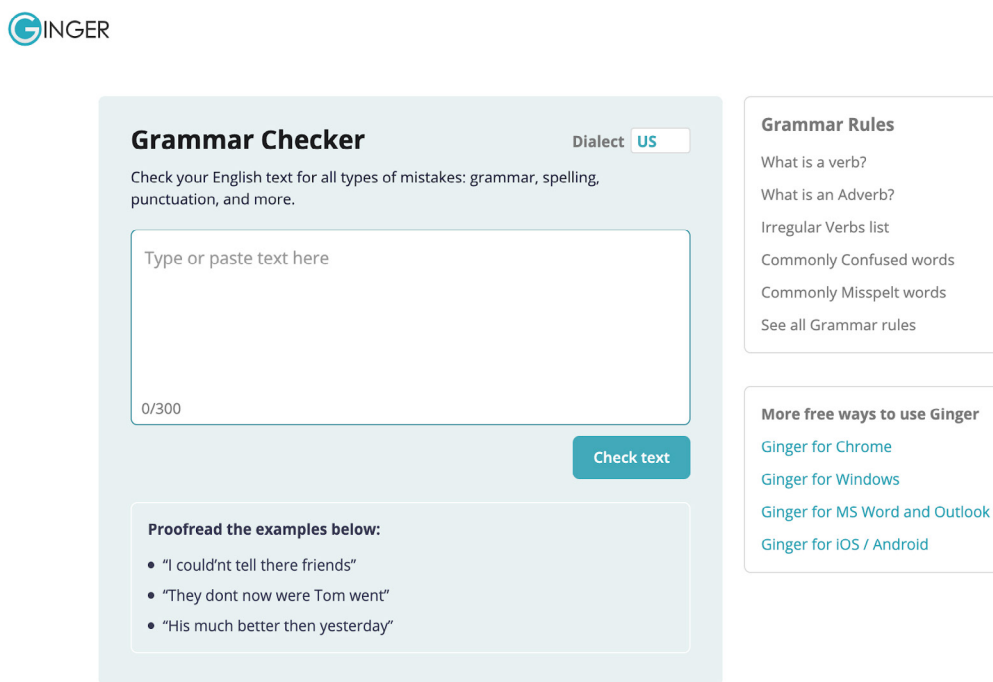


Figure 11. *Ginger Grammar Checker* (<https://www.gingersoftware.com/grammarcheck>)

As can be seen, this simple automated system can be very useful to proofread a writing before handing it in to the teacher. In this way, basic errors can be avoided and the teacher can focus on more relevant ones such as adequacy, content, style or organisation.

### 3.2.5 Grammarly

*Grammarly* is an AWE program founded by Max Lytvyn, Alex Shevchenko, and Dmytro Lider in 2009. It started as a way to support students' grammar and spelling through a subscription and it soon improved its capabilities. It can be added to Google Chrome freely and corrects grammar, spelling and punctuation while writing on Gmail, Twitter, Outlook, etc. For a fee, there are other plans (premium, business, and educational institutions) that also include the revision of sentence variety, compelling vocabulary, tone, politeness, formality and plagiarism, among others. It is a tool that helps the teacher to focus on higher-order mistakes such as complex structures, layout, coherence and cohesion. As for how it works, users upload their writing and receive two scores: one based on percentage accuracy and another

one based on the total number of errors identified. Those errors are organised into six categories: contextual spelling, grammar, punctuation, sentence structure, style and vocabulary enhancement. What is different in this program is that it provides explicit and implicit feedback. Mistakes are underlined, but if the student moves the mouse over the word or phrase, it offers them suggestions.

*Grammarly* has over 20,000,000 active users and although there are few studies on this program, some research (e.g., Cavaleri & Dianati, 2016; Gain et al., 2019; Japos, 2013; Karyautry, 2018; Qassemzadeh & Soleimani, 2016; Reis & Huijser, 2016) has concluded that it is successful to improve students' writing quality. For example, Japos (2013) analysed *Grammarly* as a tool to identify and correct mistakes. The findings showed "significant reductions in the occurrence of the grammar errors indicating that the coaching interventions were effective" (p. 97). In addition, Qassemzadeh and Soleimani (2016) found that the effect of *Grammarly* on learning the passive voice rules was higher than the traditional teacher.

Cavaleri and Dianati's (2016) study showed that Australian higher university students were happy using *Grammarly* because they said it was useful, easy to use and improved their written work and understanding of grammar rules. Similarly, in the study by Gain et al. (2019), 88.6% of the respondents used *Grammarly* to write research papers, reports (53%) and/or emails (45.5%). According to the participants' opinions, the authors claimed that *Grammarly* "improves the writing skill, catches contextual spelling and grammar mistakes, integrates with Microsoft Office, shows definitions and synonyms via double-clicks, has many editorial features [punctuation, syntax, style, plagiarism, etc.]" (pp. 11-12). In addition to these advantages, more than half of the users expressed they did not need any training and did not face any problem while using *Grammarly*. Still, 28% of users expressed lack of awareness as for features and functionality of this software.

In sum, *Grammarly* is not a mere grammar checker but a sophisticated tool that provides both implicit and explicit feedback on different categories. Similar to other AWE systems, *Grammarly* allows teachers to focus on higher-order mistakes instead of surface ones.

### 3.2.6 *Criterion*

Developed by the Educational Testing Services (ETS), this is probably the most well-known AWE system. It addresses mainly high school and college students and uses a scoring engine called E-rater that focuses on many different aspects such as lexis, syntax, content, organisation, style, and grammar in order to provide holistic and specific feedback. It was designed to be an aid for classroom instruction. In this way, students can have extra practice without overloading their teacher. Students can choose from a variety of essay topics organised by grade and different planning templates to outline their essay before writing.

*Criterion* contains two complementary applications that are based on Natural Language Processing<sup>5</sup> (NLP) methods. The scoring application, E-rater, extracts linguistically-based features from an essay and uses a statistical model to determine how these features are related to overall writing quality, so that a holistic score may be assigned to the essay. The second application, Critique, is a suite of programs that evaluates and provides feedback for errors in grammar, usage, and mechanics, identifies the essay's discourse structure, and recognizes undesirable stylistic features (Burstein et al. 2003).

E-rater analyses syntax, discourse, topical content, and lexical complexity by means of a parser and a content vector analysis. In turn, Critique is a corpus-based and statistical application that extracts and counts sequences of adjacent word and part-of-speech pairs called bigrams. Then, the system compares the frequency of the bigrams from the students' essays and the ones in the corpus. Table 5 shows the types of error comments that *Criterion* provides.

---

5. "NLP is a branch of artificial intelligence that deals with the interaction between computers and humans using the natural language. The ultimate objective of NLP is to read, decipher, understand, and make sense of human languages in a manner that is valuable" (Garbade, 2018, n.p.).

Table 5. Types of error comments provided by *Criterion*

|                              |  |
|------------------------------|--|
| Grammar                      | Fragment or missing comma, Run-on sentences, Garbled sentences, Subject-verb agreement, Ill-formed verbs, Pronoun errors, Possessive errors, Wrong or missing word, Proofread this   |
| Usage                        | Wrong article, Missing or extra article, Confused words, Wrong from of word, Faulty comparisons, Preposition error, Nonstandard word form, Negation error  |
| Mechanics                    | Spelling, Capitalize proper nouns, Missing initial capital letter in a sentence, Missing question mark, Missing final punctuation, Missing apostrophe, Missing comma, Hyphen error, Fused words, Compound words, Duplicate |
| Style                        | Repetition of words, Inappropriate words or phrases, Sentences beginning with coordinating conjunctions, Too many short sentences, Too many long sentences, Passive voice  |
| Organization and development | Introduction, Thesis, Main ideas, Supporting ideas, Transitional words and phrases, Conclusion   |

*Criterion* is a more sophisticated program than the ones explained above. On the one hand, according to Burstein, Chodorow and Leacock (2003), there is a high agreement rate on holistic scores between the *Criterion* E-rater and human expert graders (97%). Likewise, Hutchison (2007) reported positive results in a study with 11-year-olds since there was high correspondence between human raters and E-rater in the way more mechanical factors (e.g., paragraphing) were scored. In addition, there are studies that prove the effectiveness of *Criterion* to improve the students' writings and increase their reflection on their own production (e.g., Attali, 2004; Chodorow, Gamon & Tetreault, 2010; El Ebyary & Windeatt, 2010). Chodorow, Gamon & Tetreault (2010), for example, found that *Criterion* helped students to reduce article and preposition errors. In Attali's (2004) study, data from 9,275 essays submitted to *Criterion* were analysed to check whether there were any differences from the first and last (of three) drafts. Results showed that the rate of most of the 30 specific error types was reduced, and thanks to feedback, discourse elements improved (i.e., background and conclusion elements) as did the length of the students' essays. In light of these findings, Attali (2004) argues that the students understand, attend and react to the feedback provided by *Criterion*.

In a study by El Ebyary and Windeatt (2010) on the effectiveness of *Criterion* to amend errors, results showed beneficial effects on the use of articles and tenses suggesting that computer-generated feedback was more effective in revising drafts than the traditional approach of teaching writing.

On the other hand, *Criterion* also presents some weaknesses, as it has been proved that higher scores were obtained by means of avoidance strategies and not for quality writing

(El Ebyary & Windeatt, 2010). In addition, Warschauer and Grimes (2008) conducted a study on the in-class use of *Criterion* and *MyAccess!* in four secondary schools. By means of interviews, surveys, and classroom observation they found both programs were beneficial in different ways: students were encouraged to make more revisions, thus saving teachers' time and none of them interfered in the way teachers instructed their students in the writing skill. In contrast, revisions were made at the surface level (i.e., word or sentence level), and no noticing on content and organisation was significantly fostered. Moreover, taking into account that *Criterion* only recognises standardised American English and the difficulty it has to analyse writing by NNS students, Herrington and Moran (2012) demonstrated that the system may sometimes flag errors that are not truly errors, and, therefore, score accuracy fails. Furthermore, the results of a study conducted by Chapelle, Cotos and Lee (2015) report that nearly 50% of the feedback provided by the software was disregarded by students.

Despite its weaknesses, *Criterion* has been reported to be a precise AWE program in detecting errors accurately. For example, Lavolette, Polio and Kahng (2015) found that *Criterion* identified 75% of errors accurately, a percentage very close to the one Ranalli, Link, and Chukharev-Hudilainen (2017) reported (71-77%), but it missed over half of language errors. Precision in identifying errors may depend on the type of error. Han, Chodorow and Leacock (2006) claimed that *Criterion* was 90% precise in the detection of article errors. This percentage decreased to 80% in the detection of preposition errors in Tetreault and Chodorow's (2008) study. Hence, according to some studies (Burstein et al., 2003; Leacock et al., 2010) it is more desirable to fail to identify some errors than to flag well-formed text as ill-formed. In sum, it seems that *Criterion* is an effective tool "given that the proportion of successful revision is over 70%" (Chapelle et al., 2015, p. 391).

### **3.2.7. Grammar Checker**

In 2001 James Lawley, Rubén Chacón and Sergio Martín from the *Universidad Nacional de Educación a Distancia* in Madrid started to work on *Grammar Checker* (GC), a software nominated for the ELTons<sup>6</sup> in 2016. This software is addressed to Spanish learners of English

---

6. The ELTons (English Language Teaching Innovation Awards) are international awards given annually by the Brit-

as a foreign language and it is based on the Lexical Approach theory. That is to say, it is not based on tagging or parsing, but on a corpus that analyses chunks and suggests revision when it spots strange combinations. Thus, the emphasis is on language as grammaticalised lexis instead of lexicalised grammar (Lewis, 1993).

This computer program provides written feedback on grammar, spelling, words used incorrectly and pairs of words based on a corpus of eighty million words of English as a foreign language. As explained by Lawley (2015), the program divides the text into segments that are compared to that corpus and highlighted in red if they do not appear in it or have a threshold number lower than 0 and 0.1, in orange if they occur in the corpus fewer than 500 times and their threshold numbers lie between 0.1 and 0.5, or yellow if they occur fewer than 75 times and their threshold numbers lie between 0.5 and 0.9. Even though a bigram may be correct, sometimes false alarms occur. When colours appear, students are offered some suggestions on the right margin, which may not be necessarily the solution to the error but some food for thought to amend it. Therefore, this program requires cognitive process from students, as it does not offer the possibility to receive corrections at the click of a mouse. Students are responsible for whether to change the segment or not, so they must think for themselves. In this way, it offers the opportunity to learn from mistakes. GC does not evaluate or score a text, but alerts users to those combinations that are rare or do not occur.

The positive aspects of GC are mainly related to the correction process. GC accelerates the self-correction process since it makes users focus on the coloured bigrams. It also accelerates the teacher's correction process since the text does not contain basic grammar and spelling errors. However, the effectiveness of the program depends highly on the teacher's and students' attitudes toward computer-based feedback and their technological skill in working with computer-based programs.

The program stands out for its simplicity of use. Figure 12 illustrates the interface of GC.





Figure 12. *Grammar Checker* (Chacón, 2016)

After creating an account, the student only has to write the text and press “Enter your text” and then “Start” to check if there are any mistakes. First, spelling mistakes are highlighted in yellow (also purple if it is a very rare word but not necessarily an error, e.g., proper names) and by clicking on the words highlighted useful feedback is provided. By clicking on “Modify” the previous spelling mistakes can be corrected and checked again by pressing “Check again”. Then, the same procedure is followed for the “Incorrect sequences filter” that highlights grammar error, e.g., *these table*, and for the “Problem words filter”, which refers to correct English and does not highlight any word but suggests words that are usually misused by students, e.g., *insano–unhealthy/insane*. Therefore, if after reading the suggestions the students think they have committed an error, they can modify it. Figure 13 is an example of the suggestions GC provides when detecting a possible error.

^ actually

---

**actually**

"Actually" means "in fact"; we use it to indicate a surprising fact:

"I think Mary is a nurse" - "Actually, she's a doctor".

| Common Error  | Correct Version                               |
|---|---|
| He was in Oslo yesterday; <del>actually</del> he's in London. | He was in Oslo yesterday; now he's in London. |

Other ways of saying "now" are: "at present"; "currently", and; "at the moment":

At present, we're quite happy, thank you.

The exhibition is currently closed to visitors.

At the moment, everything's OK.

Find the occurrences of this problem word in the text

Figure 13. Problem Words in *Grammar Checker* (Chacón, 2016).

In the above figure, GC has detected the word 'actually', which is a false friend in Spanish. The correct use of this word is presented, followed by some examples, and then it is the student's task to decide if the word 'actually' has been used correctly in his/her piece of writing. The last step is the "Pairs filters", which highlights phrases that do not usually occur, e.g., *bad do*. In order to know the frequency with which those phrases occur and decide whether it is a mistake or not, the student can use the search engine at the top of the screen. In the version used for this study there was only available a list with frequent combination as the ones that can be found on a corpus; however, in the updated version one can choose the number of words this combination should contain (2-grams, 3-grams, 4-grams or 5-grams) and look for them in a corpus of 80 million words. Figure 14 depicts how frequency is shown by the program.

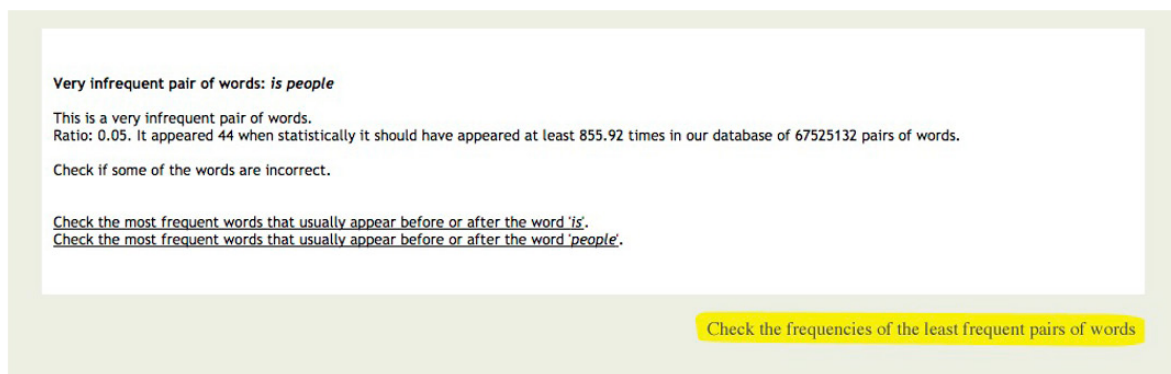


Figure 14. Pairs Filters in *Grammar Checker* (Chacón, 2016)

If the ratio is lower than 1 it means the combination is unusual; on the contrary, when the ratio is higher than 1 it means that the combination has a tendency to occur. GC shows the frequency some combinations occur; however, it does not say they are incorrect if the ratio is lower than 1, as it only indicates it is a rare combination that may contain errors.

In sum, GC does not make corrections at the click of a mouse and although it provides suggestions, it is the student who decides whether modifications are needed or not, and if so, how. In this sense, students need to think for themselves. Nevertheless, the main drawback is that GC only examines one- and two-word sequences and does not assess longer phrases.

### 3.2.8 *Intelligent Academic Discourse Evaluator*

*Intelligent Academic Discourse Evaluator* or *IADE* was developed at Iowa State University to complement writing instruction and provide feedback to ESL graduate students who need to write research articles. More specifically, it was developed as a learner-centred activity and discipline-specific tool to provide feedback on the introduction sections of research articles (Cotos, 2011) but it does not detect or provide feedback on writing errors. From students' written output, *IADE* generates colour-coded feedback (blue, red and green) intended to serve as enhanced input in order to encourage noticing on the learners' part. The different colours (blue, red, and green) are indicative of three moves: Move 1, which establishes a territory (disciplinary knowledge on the topic), Move 2, which establishes or identifies a niche in

the existing knowledge territory, and Move 3, which demonstrates how it addresses the niche.

Figure 15 is an example of an introduction research representing these three moves.

With the development of international communication and corporation, international relations turn up and deserve more attention. Thus, national policies especially to foreign countries would influence a country's images to a large extent. As the former research said that "Most national governments conduct international public relations programs, with varying strategies and effects. These public relations programs are closely connected to the mediation of their images and foreign policies." (Zhang and Cameron, 2003, p 13). However, except for the foreign policies' establishment, excellent images could be set up through media coverage. As we know, mass media is a magnifier for transmitting any issues, opinions etc and it could engender a huge impact in publics and thus public opinion will be generated. Information in this era of globalization, the good or bad image of a country will have a major impact on its political, diplomatic, business and so on. Thus, national images have been gradually thought much of but this is not to say that public relations practice can just select good news and report positive opinions. All the images should be based on truth. One of the most interesting tendencies in political image-making in recent time has been the increasing use of professional public relations consultants by national government. This study predicts the international public relations campaign effectiveness could be improved or measured by news coverage of a given country in several major mass media of other foreign country.

Figure 15. *IADE* colour-coded feedback (Chapelle, Cotos & Lee, 2015, p. 11)

The potential benefits of *IADE* have been analysed, for example, by Cotos (2011). This author claimed that when *IADE*'s feedback was noticed, students "became more cognitively engaged" (p. 436), a fact that, in turn, resulted in new output modifications in the form of a new and more accurate draft. In subsequent research, Cotos (2012) aimed at exploring the impact of *IADE*. Mixed results were obtained: on the one hand, negative impact occurred when learners only relied on the numerical feedback provided, since their cognitive involvement was low. On the other hand, positive impact took place when learners relied on the colour-coded feedback, which helped them make successful changes and, therefore, more effective revision. In addition, it seems that the participants' motivation increased by means of comments on aspects for revision or improvement and not necessarily by a score.

Another study conducted by Chappelle, Cotos and Lee (2015) supported the effectiveness of *IADE* to help students focus on meaning. 92% of the participants stated they had addressed meaning when revising due to the colour-coded feedback. Lexical modifications were also made when realising that certain words could help them build certain moves. In conclusion, unlike other types of software (e.g., *MyAccess!*, *GrammarCheck*, *Ginger Grammar Checker*, etc.), *IADE* is a potential tool to help students focus on how meaning is conveyed and interpreted in research articles.

To sum up, and in light of the above research, Chappelle and Chung (2010) state that AWE systems are commercial products that need to be analysed critically by teachers and students. Thus, it would be sensible to think about AWE systems as useful tools that “can help motivate students to write and revise, increase writing practise, and allow teachers to focus on higher level concerns instead of writing mechanics” (Grimes & Warschauer, 2010, p. 34).

Once some relevant AWE programs have been discussed, a related topic centres on how both students and teachers may perceive the feedback they provide in the EFL classroom. For this reason, the next section addresses this topic and the potential benefits if implemented in this context.

### **3.3 Learners’ and Teachers’ Perceptions on Computer-Generated Feedback**

Engagement at behavioural, emotional and cognitive levels is necessary for students to feel motivated and make the most of computer-generated feedback (Fredricks, Blumenfeld & Paris, 2004). A positive attitude, involvement in the tasks, and cognitive effort go hand in hand if learners want to be successful when writing an essay or when receiving feedback on their writings in order to make the necessary corrections. This engagement, in fact, facilitates language acquisition and writing development, and according to some scholars (e.g., Fredricks, 2013; Skinner & Pitzer, 2012), it may be associated with positive academic results. Unfortunately, despite its relevance in L2 writing, it is an under-researched area (Handley, Price & Millar, 2011) because it is sometimes difficult to determine if the positive or negative opinions the students have of AWE software are due to the system’s characteristics (e.g., easiness of use, efficiency) or

the students' own perception. In the case of e-feedback, although students regard it as useful (Elola & Oskoz, 2016; Liu & Sadler, 2003; Lu & Bol, 2007) and it leads to deeper revision, it seems students prefer face-to-face feedback (Guardado & Shi, 2007; Liu & Sadler, 2003; Schultz, 2000; Tuzi, 2004).

In the case of computer-generated feedback, the use of ICALL has changed both the teacher's and the learner's roles. The teacher, who acts as a guide, stops being the centre of attention and his/her aim is to help learners to solve any difficulties students may encounter and focus on higher-level errors. Regarding the learners' role, they are not in a student-centred environment any more, which means they need to take an active role to interact and negotiate meaning. In this way, they feel they are in charge of their own learning and especially less able students can overcome any self-esteem issues or other feelings they may have (e.g., shyness, fear of mistakes, etc.).

The students' preferences for writing their work by hand or using a computer may depend on several factors, such as mother tongue or proficiency level. According to Wolfe and Manalo (2004, p. 45), "Examinees who speak a language that is not based on a Roman/Cyrillic alphabet are more likely to choose handwriting than are examinees who do speak a language that is based on a Roman or Cyrillic alphabet [and] examinees with poorer English language skills are more likely to choose handwriting than are examinees with better English skills". In addition, students who exhibit "lower levels of computer experience and higher levels of computer anxiety are less likely to choose the word processor as their composition medium" (Wolfe & Manalo, 2004, p. 46).

Some studies have been conducted to measure learners' and teachers' reactions to traditional and computer-generated feedback. After comparing the reactions of learners to both handwritten and computer-assisted feedback, Denton et al. (2008) found that learners preferred computer-assisted feedback for "markscheme clarity, feedback legibility, information on deficient aspects, and identification of those parts of the work where the student did well" (p. 486), and teachers also opted for it because it saved them time when marking. On the contrary, Lai (2010) detected that EFL learners in Taiwan preferred peer feedback over computer-based feedback using *MyAccess!* even though they felt both of

them were effective. Computer literacy may also influence their preferences and although in Matsumara and Hann's (2004) research students were allowed to choose whether or not to use the computer, both high- and low-anxiety students improved. Tuzi (2004) also analysed this relationship between students' preferences and effectiveness of feedback. Although students preferred oral feedback, computer-based feedback was more effective when revising their work and fostered macro-level revision.

El Ebyary and Windeatt (2010) conducted a study which included a pre-treatment stage where students completed a questionnaire on their attitudes to feedback received on their writings, a treatment stage where 24 students had to use *Criterion* (see Section 3.2.6) and a post-treatment stage where the students completed a questionnaire to analyse their attitudes again. In the pre-treatment group, on a four-point scale (positive, neutral, negative, not sure), 25% had a negative attitude, 45% were neutral, 17% held positive attitudes and 13% were undecided. It means that there is a general dissatisfaction with feedback practices which may be due to the system, the instructor, or average of students per class. However, in the post-treatment stage and after having received feedback from *Criterion*, the learners' attitudes had completely changed with 88% of positive opinions compared to the initial 16%. They felt more optimistic about writing and were more conscious of the assessment process. Similarly, Rich (2012) conducted a study to analyse the effects of *Writing Roadmap 2.0* (AWE software developed by CTB/McGraw-Hill) with a group of Chinese students learning English. The experimental group exhibited higher scores than the control group and 94% of the participants liked the software. In addition, 75% of the participants explained they became more confident in English writing and 61% revised their written work 3 or more times. The same happened with Wang, Shang and Briody's (2013) study which investigated the use of *CorrectEnglish* with EFL Taiwanese students. They found that the experimental group had a positive attitude towards the software and even noticed its benefits as for error reduction. Finally, Link et al. (2014) conducted a study on the effect of *Criterion* on writing instruction and performance. Results suggest positive effects since this program led to increased revisions, engaged students in writing practice and helped improve accuracy from a rough to a final draft, something the instructors agreed on.

Students' perceptions on the usefulness of automated feedback may vary depending on the software used and the teacher's approach. In Chen and Cheng's (2008) study, *MyAccess!* was used in three classes of Taiwanese third-year college students majoring in English. In the authors' words, "the students questioned the value of AWE since they did not want to be confined by machine-governed rules nor did they want to be assessed by a machine that did not understand what they wrote" (p. 108). In particular, the use of *MyAccess!* "was not perceived very positively due to limitations inherent in the program's assessment and assistance functions" (p. 107). More specifically, students in class A did not trust the program as they found discrepancies between the automated assessment results and their instructor's assessment. The teacher in class B was against the automated scores and participants in class C did not have anything to compare the scores the software provided. These findings suggest that different writing teachers' practices and attitude toward the program may have influenced the students' perceptions on its effectiveness and usefulness. Research suggests, therefore, that the extent to which learners trust AWE is still undetermined since its effectiveness is somehow doubtful (Lai, 2010).

To sum up, feedback can improve L2 students' writing ability and contribute to SLA but it depends on the students' engagement to make the most of it. Either by receiving electronic or handwritten feedback, teacher or computer-generated feedback, further research on the perceptions of students and teachers is needed. The effectiveness of automated feedback cannot be conferred uniquely to the software itself. There are many other features that contribute to its success or failure such as proficiency level of the students, their motivation and attitude towards new technologies, and familiarity with the software, among others.

### **3.4 Chapter Summary**

Current automated feedback programs, i.e., software which uses AI to generate CF and score writings, are not a panacea but a tool that can assist learners in their written performance and learning process without negative interferences in the writing instruction process. Although AWE systems keep improving and students feel more and more comfortable with them, they seem not to be perfect, as they still present some weakness. Students can manipulate them by



employing certain strategies such as lengthiness and the use of specific vocabulary or grammar features (Cheville, 2004; Herrington & Moran, 2001; Powers et al., 2002; Ware, 2005; Yang, 2004). In addition, AWE systems overemphasise the use of connectors, they ignore coherence and content and they favour formulaic conventions, which restricts students' expression, creativity and originality (Chen & Cheng, 2008).

These programs can only detect surface errors and cannot provide post-grading consultation in the form of, for example, clarifications, extra comments or negotiation of meaning, as it occurs with teacher feedback (Oladejo, 2005). In spite of these drawbacks, automated feedback is a good tool at the teacher's disposal (Warschauer & Grimes, 2008), which assists students in lower-level errors and allows teachers to focus on higher-level errors (e.g., content, meaning, style, organisation, etc.). As some scholars maintain, it should not substitute the teacher but function as a supplement (Burstein & Marcu, 2003; Burstein, Chodorow & Leacock, 2003).

Broadly speaking, teachers appraise the use of computer-based feedback for being quick and effective on formal aspects such as grammar, spelling and punctuation. Integrating computer-generated and teacher feedback in the language classroom may encourage both teachers and students to use automated feedback with more pleasure and confidence. The combination of both assures students that their writing has an acceptable level before the teacher reads it, enhances autonomy and awareness of writing mechanics, gives students confidence on their written production, and finally, allows teachers to focus more on meaning and content.

In conclusion, although the research into the use of AWE applications is inconclusive as some studies report favourable results (e.g., Attali, 2004; Coniam, 2009; Dmytrenko-Ahrabian, 2008; Ellison, 2007; Hutchison, 2007; Ussery, 2007) while others report negative or mixed results (e.g., Chen & Cheng, 2006; Lai, 2010; Lee et al., 2009; Tuzi, 2004), *Grammar Checker* has been used in the present study to contribute to the research on the effectiveness of AWE feedback (see Section 3.2.7). In the end, as Lee et al. (2009) claim, it is necessary to evaluate the validity and effectiveness of the AWE program in the context where it is going to be implemented (for example, classroom dynamics and students' willingness to

use it, learning and teaching styles, writing ability, etc.). It is also important to note that the teacher's attitude towards the AWE program is going to determine its effectiveness. In this sense, it is important that teachers are familiar, believe in and know how to use the program before introducing it to students. As for the students, it is relevant to know what relationship they have with technologies (Grimes & Warschauer, 2006; Kozna & Johnston, 1991).

Following the theoretical framework on computer-generated feedback and the analysis of different AWE systems, Part II will focus on the study itself where the motivation, method, results, discussion and conclusion will be explained.

## **Part 11: The Study**



## **Chapter 4. Motivation for the Study**

A few decades ago, technology in the language classroom was only applied by some teachers and it was limited to the use of television, videos, or tapes (Cunningham, 1998). However, at present, it seems language education is hardly understood without the use of computers and the Internet. Language students widely use these tools for classroom activities, search for information, teacher-student communication by email, online platforms, etc. For some years now, computer-generated feedback has been employed to help language students monitor their learning, being this the main focus of the present research. Therefore, after having dealt with key issues in language learning, corrective feedback and recent computer software used in the educational context in the previous chapters, the present chapter is devoted to explaining the motivation and research questions (RQs) of the study in order to examine the use of computer-generated feedback as a tool in the language learning process.

Writing is one of the most complex challenges for language learners as they do not write very often and not much time is devoted to practising it in the classroom. It is a skill that reveals how autonomous, understandable, fluent and creative a person is, since putting one's thoughts in a meaningful and organised form is not an easy task. However, most of the time students write the same way they speak or even translate from their mother tongue, which leads to transfer errors. Thus, taking into consideration this situation and all the topics discussed in the theoretical framework, the main reasons that motivated the present study are presented in what follows.

After completing my degree in English Studies, I became a teacher of English at the Official School of Languages in Castelló in 2010. Thanks to my teaching practice, I realised students had difficulties when producing texts. Many of my students' essays contained errors because they did not revise their compositions, did not notice there was an error or did not even know how to correct it. A second motivation centres on the increasing number of students EFL teachers have to face. Large classes make it difficult to cope with loads of corrections and with voluntary compositions. Writing is a demanding skill that requires constant practice and comprises many elements (grammar, vocabulary, spelling, punctuation, coherence, cohesion, fluency, content, accuracy, etc.), which are

hard to practise under no supervision. Finally, the last motivation refers to the students' demand to practise their writing skill more often. This issue may seem somehow contradictory, because although students feel they need more practice, when they are asked to write more often, they do not fulfil the expectations. Their personal circumstances, job and studies make it difficult for them to devote the necessary time to the writing skill.

In this respect, the search for new tools that can offer students individualised practice to overcome their weaknesses has been gaining weight. As a consequence, new technologies have triggered important changes in the fields of instruction and writing practice. From an educational point of view, the interest in new technologies is to maximise their versatility in order to offer students and teachers useful resources and tools that can help them improve their work. Thus, in the field of second/foreign language learning, the combination of language, in this case English, and technology increases motivation and provides students with real-life practice (e.g., Bromely, 2005; Halsey, 2007; Keeler, 1996; Kinzer & Verhoeven, 2008; Lee & O'Rourke, 2006; Martin, 2008; Puckett, Judge & Brozo, 2009; Van Leeuwen & Gabriel, 2007; Warschauer, 1996).

For the above reasons, it seems appropriate to use new technologies to motivate and improve the students' writing ability. Students can accomplish this by means of computer technologies that can be used anytime and anywhere, saving both students' and teachers' time, and avoiding students' frustrations due to basic or unnoticed errors. In the next section, the purpose of the study will be explained in detail.

#### **4.1 Purpose**

After having realised that students have problems writing correctly and without minor mistakes on, for example, basic grammar, spelling or punctuation, I researched on the different ways to boost their writing practice outside the classroom and to improve their written work. After having read the relevant literature on the topic, I came up with the so-called computer-generated feedback and AWE programs where I noticed a gap, since the research of AWE systems is not too abundant and there are no conclusive results. Thus, the aims of the present study are, first, to contribute to a better understanding of computer-generated feedback. As explained in Chapter 3, computer-generated feedback is a computer program based on algorithms and on a

corpus that provides feedback by analysing students' production. The urge to rely on this kind of system emerges from the need to cope with the students' needs, although many teachers have never implemented automated feedback generators (Reiners et al., 2011) and, as a consequence, they may be reluctant to use them. For this reason, the present study aims at explaining how the software selected (i.e., *Grammar Checker*) works, and whether it can help teachers and students in the learning and teaching process, for example, by providing immediate feedback (Bai & Hu, 2017; Dikli, 2006), multiple drafting opportunities (Warschauer & Ware, 2006), or acquiring learning adaptability and autonomy (Chen & Cheng, 2008).

The second aim tries to shed light on the effectiveness of AWE tools to improve the students' written work, by means of the software *Grammar Checker*. Although AWE programs can provide individualised feedback on specific aspects (Bolt, 1992; Cotos, 2015; Liou, 1994; Warden & Chen, 1995) as well as on more holistic elements (Attali, 2004; Burston, 2001; Cotos, 2015; Leacock, 2004; Li et al., 2014), research into their implementation shows inconclusive results, since some studies report favourable results (e.g., Coniam, 2009; Hutchison, 2007), while some others show the opposite (e.g., Lai, 2010; Lee et al., 2009; Tuzi, 2004). However, most AWE programs fail to provide specific feedback on elements such as style, content, irony, humor or sarcasm. In addition, students can also cheat the program to obtain higher scores just by using avoidance strategies (Herrington & Moran, 2001; Powers et al., 2002; Ware, 2005). Thus, we aim at checking the effectiveness of GC as a supplement for the revision of formal aspects (i.e., grammar and spelling), since a focus on meaning and content still needs some software improvement and it is out of the scope of the present research.

Finally, our third aim centres on the effectiveness of computer-generated feedback versus teacher's feedback. As explained in Chapter 3, AWE feedback poses some advantages over teacher's feedback. For instance, it can alleviate teachers in terms of time constraints, and students can benefit from immediate feedback and multiple drafting at their own pace while boosting their autonomy and engagement (Chen & Cheng, 2008; Dikli, 2006; Hyland & Hyland, 2006). On the contrary, teacher's feedback also has its benefits. Teacher's feedback tends to address more error categories than AWE feedback, since it is usually more individualised and it is not

based on algorithms. In addition, it is also more valued by students due to personalised and comprehensive corrections (Chen & Cheng, 2006; Fang, 2010; Reiners et al., 2011).

In sum, this study represents a novelty, since, to the best of our knowledge, there is no research on the effectiveness of GC in our specific context. It aims to contribute to the still scant evidence with regard to whether computer-generated feedback results in accuracy development and learning over time, as claimed by Heift and Hegelheimer (2017), who also pointed to the need for long-term research of computer-generated feedback. The validity of GC will be assessed in an EFL setting by including a tailor-made test to analyse learning in the long term. In addition, this software will be used for formative assessment and not for summative assessment, since it does not provide a score and it is not the objective of our research.

Considering the motivation and purpose of the study, four research questions are suggested together with their hypotheses, which will be explained in the following section.

## 4.2 Research Questions and Hypotheses

Despite the fact that GC was developed two decades ago, at present there are few studies on it. In view of the scarce research on the effectiveness of AWE tools, especially GC, and the inconclusive results obtained so far, the research questions proposed are the following:

Research Question 1 (RQ1): *Will GC provide proper CF so as to improve students' accuracy on the use of prepositions, articles and past simple-present/past perfect difference?*

Hypothesis 1 (H1): Previous research on AWE programs report beneficial results related to provision of individualised feedback on specific aspects such as grammar (Coniam, 2009; Hutchison, 2007). Drawing on Potter and Fuller's (2008) claims that grammar checkers are very effective in form-focused writing exercises, GC will help students improve the use of the three grammar aspects as a resource for writing development.

Based on previous studies, L1 transfer due to the similarities between the mother tongue and the target language may benefit students (Ekiert, 2004; Hawkins et al., 2006; Ionin & Montrul, 2010; Romaine, 2003; Snape et al., 2013) but, at the same time, differences between the mother tongue and the target language may also hinder acquisition. The use of individual computer-



generated feedback is expected to improve the students' performance (Attali, 2004; Shi, 2012; Yang, 2004; Yang & Dai, 2015; Wei, 2015; Zhou, 2013).

Research Question 2 (RQ2): *Will students using GC outperform other students receiving no feedback or the teacher's feedback in terms of accuracy?*

Hypothesis 2 (H2): Computer-generated feedback via GC will be more effective than no feedback (Attali, 2004; El Ebyary & Windeatt, 2010; Potter & Fuller, 2008), and will be at least as effective as the teacher's feedback (Dikli, 2006; Landauer et al., 2003; Nadasdi & Sinclair, 2007; Nichols, 2004; Page, 2003).

Research Question 3 (RQ3): *Will students' gains be maintained in the long term irrespective of the feedback provided?*

Hypothesis 3 (H3): Students' gains after receiving WCF (either computer- or teacher-generated) will be maintained in the long term. Providing indirect CF will indicate that an error has been committed, thus promoting reflection and problem solving and, most likely, leading to long-term acquisition (Bitchener & Knoch, 2008; Link et al., 2020).

Research Question 4 (RQ4): *Will students' attitude about the kind of feedback received (teacher and computer) influence the results?*

Hypothesis 4 (H4): Computer-generated feedback by means of GC will engage the students in the revision process (Chapelle et al., 2015; Li et al., 2014). As for the teacher's feedback, although sometimes students may feel frustrated because of the gap there is between the teacher's and student's understanding of the error (Adcroft, 2011; Orsmond & Merry, 2011; Xue-Mei, 2007), students will also have a positive attitude towards the teacher's feedback (Septiana, 2016).

After having presented the research questions and the suggested hypotheses, the next chapter focuses on the study itself (i.e., pilot study, setting and participants, targeted grammatical features, instruments, data collection procedure and data analysis) in order to contextualise the study.



## Chapter 5. Method

### 5.1 Pilot Study

Chapter 5 aims at explaining the procedure followed to implement the research. Before conducting the main study, a pilot study was completed. Identifying potential deficiencies and testing the feasibility of the study itself by examining data collection, instruments and subjects were necessary in order to increase the likelihood of success and guarantee a good study design. After having solid groundwork, the main study was implemented, where some changes applied. The purpose of the pilot study was twofold: firstly, to test the validity of GC for the study, and secondly, to determine important details such as the estimated time necessary for the study, the students' reactions to computer-generated feedback and the type of teacher's feedback.

The pilot study was conducted with students from the Official School of Languages, which is a state institution where students from 16 onwards can study Arabic, Basque, Chinese, English, French, German, Italian, Portuguese, Romanian, Russian, Spanish and Valencian on-site, and blended learning and online learning in some languages such as English and Valencian.

The participants of the pilot study were two 2B2<sup>7</sup> groups of students of English (n=40) who had already passed the 1B2 level, so, in theory, the groups were linguistically homogeneous. They were 39 Spanish students and 1 Argentine whose ages ranged from 20 to 55. They were classified into Group 1 (computer-generated feedback, n=21) and Group 2 (teacher feedback, n=19).

The pilot study did not target any specific features. Despite the fact that both groups received feedback (either by the teacher or the computer) on all their grammar and spelling mistakes, for other corrections on content, structure, style, etc., they received the teacher's feedback, since GC did not provide them with this kind of feedback.

As for the instruments, James Lawley, one of the designers of GC, was contacted to have access to the program and ensure it was appropriate for the study. After a fortnight trial period,

---

7. Levels at the Official School of Languages: A1, A2, B1 (1B1, 2B1), B2 (1B2, 2B2), C1 (1C1, 2C1), C2 (1C2, 2C2). According to the CEFR, the 2B2 level would correspond to the second half of the B2 level.

the teacher/researcher purchased a license so that the students could use it for the pilot study. The version used (version 2013) only analysed Spelling and Pairs Filter. GC was chosen as the best option for a number of reasons. First, it was especially designed for Spanish students of English as a foreign language. Any other software created for non-Spanish speakers may not consider these cases. Second, it works well with students with a B2 level or above as it fosters self-learning by suggesting possible causes for the error and making them critically think about their own production. In addition, they can check the frequency and real use of the word or phrase which contains an error in the corpus available. However, it may not work so well with beginners as the corpus or some of the suggestions could be too challenging. In the case of *Ginger Grammar Checker*, *GrammarCheck* and *Microsoft Word*, they provide superficial and automatic corrections, making students have a passive role and not letting them learn to learn and be autonomous. Third, GC is based on a corpus, which gives solid evidence to corrections since the students' written work is compared to a large database that contains real texts. In this sense, *Criterion* also seems to work quite well (Burstein, Chodorow & Leacock, 2003; Grimes & Warschauer, 2006; Lavolette, Polio & Kahng, 2015). On the contrary, other programs such as *MyAccess!* do not show such positive results, since the feedback provided has been reported to be repetitive and not accurate (Chen & Cheng, 2008).

A fourth reason to use GC points to the fact that it can be used as a complement or first screening rather than a substitute of the teacher, as it does not provide feedback on aspects such as register, structure, meaning, etc., nor a score. Therefore, with GC, students can maintain personal contact with their teacher and do not lose this social and interactive nature of writing. In addition, teachers can focus on more complex errors such as content, organisation, structure, etc., which are sometimes checked superficially because of lack of time or work overload. *MyAccess!* and *Criterion*, on the contrary, provide feedback and a score that can be higher by means of avoidance strategies. A score is not relevant in this study, as the focus is on students' noticing their errors and being able to amend them. In fact, as mentioned in Section 3.2.2, students were not happy with the score provided by *MyAccess!* and the accuracy of feedback (Chen & Cheng, 2008) and the software can be easily fooled by students because it praises length

over content (Herrington, 2001). All those negative aspects could lead to demotivation and too much reliance on a number more than on how much students can improve.

Finally, a fifth reason refers to cost. GC is a low-cost national product. Although it is not free, the use of GC as an instrument to conduct research on feedback and language learning deserved a try, to our mind. It is very gratifying to have the possibility to contribute to the analysis of a tool created in our country which is within anybody's budget and to be able to contact its designer easily.

In sum, the pilot study gave the teacher/researcher the possibility to try GC with students before implementing the actual study. This program was selected for its flexibility (to be used as a complement whenever and wherever), realism (corpus), self-learning nature (autonomy) and low-cost price, as mentioned in Section 3.2.7 of the theoretical framework.

The pilot study started in November 2014 and was developed over a period of three months until January 2015. The first week of the pilot study, Group 1 (computer feedback) was explained how to use the computer program, how to hand in the writings, what kind of feedback they would receive and how to improve their writings after it. Students in Group 2 were explained the same steps excluding the use of the computer program, since the feedback they were going to receive was the teacher's feedback. In the following months, the students wrote three compositions: an application letter, an article and a review, as Figure 16 illustrates:

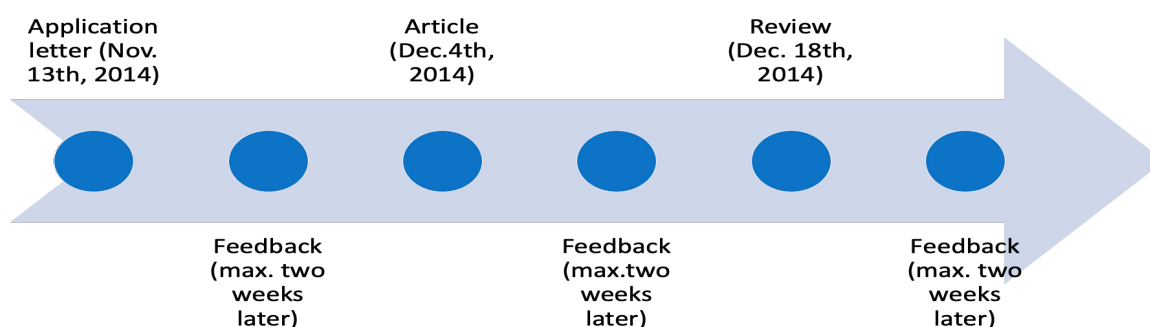


Figure 16. Chronological order of the pilot study

Group 1 wrote their texts using GC and Group 2 did it by hand or on their computers. Group 1 sent the teacher 3 screenshots: one of the Spelling Filter, another one of Pairs Filter and the final draft. After having handed in their writings and within the following two weeks, Group 1 and Group 2 received the teacher's feedback: Group 1 only on higher-level aspects (e.g., structure, style, register, content, meaning, etc.) that are not corrected by GC and on errors not marked by GC or wrongly corrected by the students, whereas Group 2 obtained feedback on all the aspects (e.g., grammar, spelling, style, structure, etc.) at once. Group 2 received their writings back to check their mistakes, which were marked in the same way GC did, and modified them. Then, the second writing was requested and the process was repeated. At the end of the study, all the errors and corrections were collected and analysed to examine the effectiveness of both kinds of WCF.

Thanks to the implementation of the pilot study, some benefits and flaws were identified. For example, students did not have major problems using GC, interpreting its feedback and handing in their compositions by email in the case of Group 1. In addition, the integration of GC in the course was quite smooth as they did the same writing tasks they are usually required at that level but using GC. Despite these advantages, some inconsistencies were observed. For example, checking the students' improvement was difficult, as the compositions they had handed in were not comparable. They wrote three different texts that required different structures and vocabulary. Furthermore, the computer program had some weaknesses too, as working at the sentence level was difficult. Some of the phrases the program had highlighted in yellow and even orange contained no mistakes. Therefore, if the structure the students used was not part of the corpus, it was highlighted in yellow or orange as unusual and the students thought it was incorrect. In addition, the initial idea of checking any modification between the first and the second draft was extremely hard for the teacher, since some of the modifications were incorrect, some others were partially correct, some others were completely changed, some others were not changed, some of them were initially correct but as they were highlighted in orange the students changed them in the wrong way, etc. All these obstacles resulted in improvements for the final study.

As stated earlier, the main aim of the pilot study was to test whether GC could serve as a tool for providing feedback to students, no specific research questions or hypotheses were posed. The expected outcomes of the pilot study were that Group 1 (i.e., students receiving feedback with the computer program) would at least be on a par with Group 2 (i.e., students who obtained teacher's feedback) or even outperform it in the following aspects: writing quality, length and teacher support on higher-order aspects such as content or structure, since GC provides students with immediate feedback on surface-level errors (grammar and spelling), thus letting teachers have more time to interact with students, as attested in previous studies that employed computer-generated feedback (Cochran-Smith, 1991; Greenwald et al., 1999; Tiene & Luft, 2001). The results of the pilot study are presented in Figure 17:

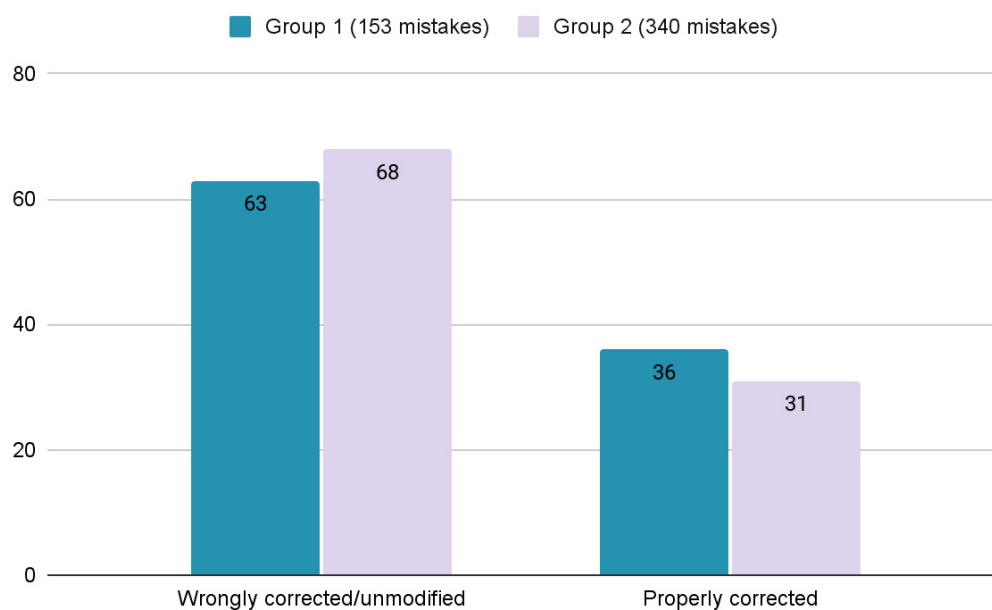


Figure 17. Number of errors and their correction in the pilot study

Group 1 made a total of 153 mistakes, from which 63% were either wrongly corrected or left unmodified, and 36% were properly corrected. In turn, Group 2 made a total of 340 mistakes including grammar and spelling. Out of those mistakes, 68% were either wrongly corrected or left unmodified, and 31% were properly corrected. The above percentages led us to believe that the number of mistakes between Group 1 and Group 2, although very similar, revealed a slight

advantage of computer-based corrective methods over traditional ones. With this basic information, we decided to implement the present study to compare computer and teacher's feedback.

After having checked the validity of GC in the pilot study and having observed some of its limitations, a number of modifications in the final study were made. First, instead of asking for independent writings (i.e., one application letter, one article and one review), the idea of asking for the same type of writing twice gained weight. In this way, comparison between both pieces of writing could reveal more reliable data. Second, the structures analysed narrowed down to common errors at B2 level: *prepositions, articles* and the *past simple-present/past perfect difference*. Therefore, by focusing only on some specific linguistic items, the benefits were twofold. On the one hand, the teacher's work was more manageable, as she would focus her attention on those parts of the essay that contained any of those three sorts of errors and the goal of the study could be accomplished more efficiently. On the other hand, and more importantly, a control group also took part in the study. The fact of focusing on three types of errors made it possible to provide students with feedback on other grammar errors and not leave them unattended (i.e., no feedback on grammar) for an extended period.

Thanks to the pilot study, the teacher could also anticipate students about the use of colours and warn them that when something was marked in yellow, it did not necessarily mean there was a mistake. Similarly, when marked in orange, they had to make sure they checked the corpus to obtain information about the frequency of the words used and the nature of the mistake. In addition, to conduct more rigorous research, the inclusion of a tailor-made test to check the long-term effects and a questionnaire to obtain further information on the relevance of the students' perceptions of both kinds of feedback were considered for the final study.

In conclusion, by limiting the number of targeted structures, adding a control group, warning students about the colour code, and including a tailor-made test and a final questionnaire, the basis to conduct the final study became more solid and reliable. In the next section, the instruments, setting, participants, data collection procedure and data analysis of the study are discussed.



## 5.2 The Study

The present study aims to examine the effectiveness of two types of CF on students' writings. In order to do so, we present the setting and participants, targeted grammatical features, instruments, and the data collection procedure and analysis.

### 5.2.1 Setting and participants

Initially, the participants of the study were two groups of adult students (n=46) studying English as a foreign language at the 2B2 level at the Official School of Languages in Castelló for different reasons (professional, pleasure, job promotion, etc.). Although nowadays there are a lot of different ways to learn a language (e.g., TV, the Internet, chats, language exchanges, private lessons, etc.), the Official School of Languages is an important source that offers students quality teaching, individual monitoring, assessment and continuous practice.

However, from the 46 students only 27 (6 males, 21 females) completed all the writing tasks and the tailor-made test necessary to collect useful information for analysis. As for their nationalities, 26 were Spanish and 1 Romanian. Seven subjects were Catalan speakers, eight Spanish speakers, eleven bilingual speakers (Catalan and Spanish) and one Romanian. As for the participants' age, they ranged from 21 to 56 (mean=39.37 years old) and all of them had studied English for over 6 years either at primary/secondary school, high school and university or the Official School of Languages.

To conduct the study, the 27 students were divided into three groups: Group 1 (computer-generated feedback, n=8), Group 2 (teacher's feedback, n=11), and Group 3 (control group, n=8). Group 1 received feedback from the computer program GC, Group 2 received feedback from the teacher and the Control Group did not receive any feedback on the targeted grammatical features but, for ethical reasons, this group was offered feedback on other language aspects. A proficiency level test was not necessary before conducting the study since they had either 1B2 qualifications or were taking the 2B2 level again as they had failed it the year before.

The researcher in the present study was the teacher of all three groups. She was a female in her late 20s who held an English Studies Degree and a Master's Degree in English Language Teaching and Acquisition. At the time of the study, she had a 4-year-experience as a teacher of English in the Official School of Languages.

### **5.2.2 Targeted grammatical features**

As mentioned in Section 5.1, after having conducted the pilot study, the scope of correction was limited, because depending on the required type of text and the content of the composition, the errors they made varied enormously. That is to say, if students are asked to write a story, they will probably use past tenses, temporal expressions, adverbs, etc. However, if they are asked to write a review about a book or film they will probably use present tenses, adjectives, include a recommendation, etc. That makes the comparison between compositions very difficult, since the sort of errors the students may make are very different. For this reason, the range of errors to take into account was reduced to three grammatical features: prepositions, articles, and past simple-present/past perfect difference.

The compositions in the pilot study differed in content and structure: an application letter, an article and a review. However, the compositions used in the final study were stories based on two comic strips and students had to use similar structures and content (e.g., past tenses, connectors, temporal expressions, similar vocabulary about family, etc). For these reasons, comparison among compositions was feasible, since both storylines were related and had similar characters (Bitchener, 2008).

The selection of targeted grammatical features (prepositions, articles and past simple-present/past perfect difference) was based on the fact that they are recurrent errors in the B2 level that could be easily reproduced in the comic strips due to their characteristics.

## 5.2.3 Instruments

### 5.2.3.1 Grammar Checker

The main instrument of this study is the computer program used by Group 1, *Grammar Checker*. After having conducted the pilot study, the feasibility of using it as a feedback provider tool was confirmed. As explained in Section 3.2.7, it is a corpus-based program that marks errors in different colours according to their frequency rate in the corpus and gives suggestions to correct them. However, it may sometimes mark correct utterances as incorrect due to their low frequency rate, that is to say, for not being included in the corpus. In the present study, and unlike the pilot study, an improved version was used (version 2014), which analysed not only spelling and pairs filters but also incorrect sequences and problem words. In this sense, students sent the teacher five captures instead of three, as was the case in the pilot study: one for spelling (yellow: spelling mistakes with pedagogic feedback; purple: words that are not amongst the 90,000 most frequent English words), one for pairs filters (analysis of the frequency of each combination of two words: red=very suspicious pairs, orange=suspicious pairs, yellow=slightly suspicious pairs), one for incorrect sequences (comparison with one million of the most frequent incorrect sequences in English), one for problem words (words associated with error: words that many students of English use incorrectly but it does not mean they are incorrect) and the final version ‘without mistakes’.

### 5.2.3.2 Comic strips

Taking into account the findings of the pilot study, the repertoire of the selected writings for the final study consisted of two combinations of comic strips (*Dear Abby in Love + Abby’s reply In love* and *Dear Abby Part-time job + Abby’s reply Part-time job*<sup>8</sup>) so as to have enough writings to be analysed. Students had to interpret the comic strips and write

---

8. [https://sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/picture\\_cues/](https://sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/picture_cues/) (Dear Abby In love.JPG+Abby’s reply In love.JPG and Dear Abby Part-time job.JPG+Abby’s reply Part-time job.JPG)

a story. In order to distinguish between the first story (i.e., *Dear Abby in Love + Abby's reply In love*) and the second one (*Dear Abby Part-time job + Abby's reply Part-time job*) in the coding of data for statistical analysis, they were renamed as ABBY and PAM, respectively. The decision of changing the original name of ABBY to PAM in the second story was taken based on a motivational aspect, so that students did not have the impression they were writing about the same four times. Each comic strip was repeated twice (i.e., ABBY\_pre, ABBY\_post and PAM\_pre, PAM\_post). That is, each student wrote four compositions in total. The comic strips are included in Appendix 1. The first comic strip (ABBY) depicts Abby writing a letter to grandma about her new boyfriend, but he does not seem to make a good impression on her. As a consequence, Abby becomes sad and writes back to grandma, who gives her some good advice to amend the situation. As for the second comic strip (PAM), Abby, a high-school student, writes to grandma to express her disappointment for not having any money. Grandma, then, writes back to ABBY to give her a piece of advice that seems to please her.

### 5.2.3.3 Tailor-made test

Six weeks after the last writing (i.e., second draft of *PAM*), each student was asked to complete a tailor-made test in order to see their improvement in the long-term. The tailor-made test consisted of an individualised set of sentences that contained errors from the targeted grammatical features they had made throughout the four writings and some distractors that contained no errors (see Appendix 4). The aim of this test was twofold. On the one hand, it was used to provide the control group with feedback on the errors they had made on the three grammatical aspects during the study, since they had not received any feedback on them. On the other hand, it could reveal information about the long-term effect of the effectiveness of the CF methods implemented (computer-generated and teacher's feedback)

#### 5.2.3.4 Questionnaire

Once the study had finished, the three groups were asked to complete a semi-open questionnaire to know their opinions and experience with GC (see Appendix 2). This questionnaire was based on Cavanaugh and Song (2014) and was delivered to gain insights into the participants' environment and perception of the instructional treatment. It was divided into two parts, which contained some Likert-scale items and some open-ended items. The first set of questions (1-10) aimed at knowing more about the students' background (personal and professional information such as age, mother tongue, education, years studying English, etc.) and the second set of questions (11-20) were about their written language skills and their opinion on how to receive feedback, especially through GC for Group 1 (questions 16-20). For this reason, Group 1 answered all the questions, but Group 2 and Group 3 only answered the first 15 questions since they did not use GC.

#### 5.2.4 Data collection procedure

We may claim that the study approximates ecological validity as it was carried out in a real classroom taking into account the complexities of real-life teaching situations. In this sense, we adhere to Ashcraft and Radvansky's (2009) words that "research must resemble the situations and task demands that are characteristic of the real-world rather than rely on artificial laboratory settings and tasks" (p. 511). Ecological validity also poses some drawbacks, since the teacher cannot always make sure the students hand in all the requested compositions, control if students follow the steps required, or guarantee they do not copy from external sources when writing. All in all, the advantages of data collection in real classrooms far outweigh the probable disadvantages the teacher/researcher may face.

Data were collected over a five-month period, from October 2015 to February 2016 by means of the software *Grammar Checker*. Students in Group 1 used GC to write their stories (*ABBY* and *PAM*) and to receive feedback on the three targeted grammatical features. According to the colour and suggestions of the program, the students had to modify, if necessary, the highlighted errors. To keep record of the suggestions made by the computer

program and the students' modifications, they had to send the teacher the screenshots corresponding to each window by email. That is to say, one shot with the suggestions/colours for Spelling, Incorrect Sequences, Problem Words, Pairs Filter and the final version, amounting to a total of five shots on a Word document.

Students in Group 2 received feedback from the teacher, that is to say, they delivered their written work either by email or by hand and two weeks later at the most the teacher returned their writing with the different errors marked in yellow, orange or red, as did GC. In the same class students were given their essay back, they had to correct their errors and hand it in back again to their teacher as a final draft. If the essay still contained some errors, the teacher provided the correct answer. Finally, students in Group 3 received the teacher's feedback using the same colours as GC except on prepositions, articles and past simple-present/past perfect difference, but in order not to leave important errors on any of the three elements uncorrected, the teacher amended this lack of feedback after the completion of the tailor-made test at the end of the study (see Appendix 3 for some examples).

All the students, regardless of the group, wrote the first draft of *ABBY*. Two weeks later they received their corrections. Then, they wrote a second version of *ABBY* and two weeks later at the most they received the teacher's corrections again: Group 1 received computer-generated feedback immediately and the teacher's feedback on errors different from the three targeted features for the study within two weeks; Group 2 received the teacher's feedback on any type of error, and the feedback provided to Group 3 was similar to Group 1, as no feedback was offered by the teacher on the three targeted features. After *ABBY*, they repeated the same process with *PAM*. All the drafts had no time constraints, as they were written at home with a length limit of 180-200 words. By making students write two drafts of the same comic strip, along with the feedback provided, we aimed at increasing the possibility of more focused and effective correction of problematic grammatical aspects (Nguyen et al., 2015).

As Sheen (2007) argued, "to claim that WF results in learning, one must examine whether the improvement in revisions carries over to a new piece of writing or if the improvement is manifested on posttest or delayed posttest measures" (p. 258), six weeks after having written the last story (*PAM*), students from the three groups were asked to complete a tailor-made

test to check the effectiveness of the two kinds of feedback (computer-generated feedback and teacher's feedback) in the long-term and to provide students from the Control Group with feedback on those mistakes they had not been corrected. All the steps for data collection are presented in Figure 18:

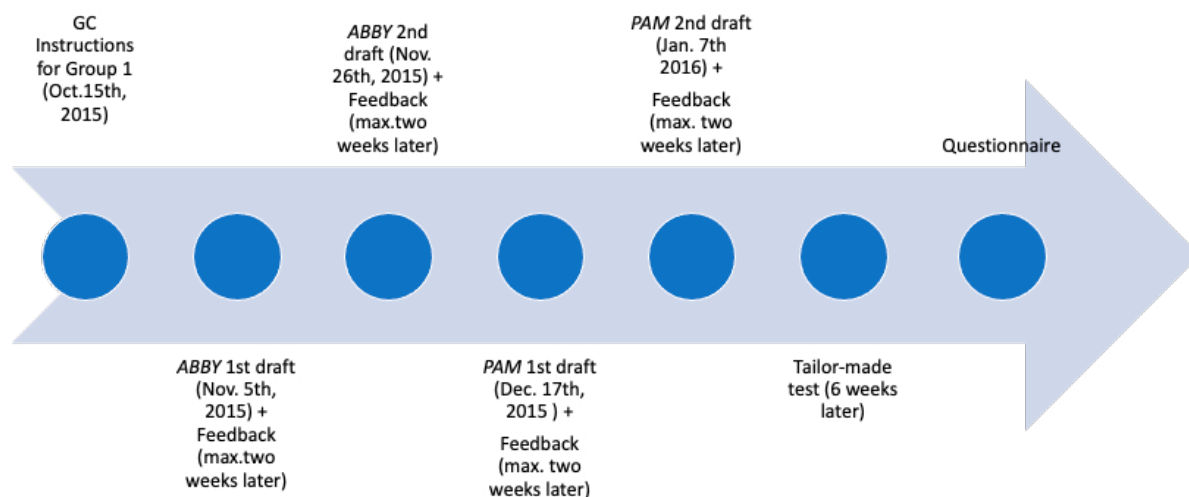


Figure 18. Chronological order of the study

All the students wrote their compositions at home (Group 1 on their computers, Groups 2 and 3 by hand or on their computers); therefore, it was impossible to ascertain whether they wrote their compositions following the teacher's recommendations. That is to say, the three groups were allowed to use any tool that might have proved useful for them, e.g., a dictionary, the textbook, a thesaurus, etc., and Group 1 should have followed the suggestions and the frequency rates provided by the computer. After each writing was handed in, approximately within the following two weeks, the students' revision process took place in the classroom. Group 1 and Group 3 received feedback on the non-targeted grammatical errors and Group 2 on any sort of error. Students could use their phones, textbooks or even their classmates' help to correct the teacher's feedback.

Six weeks after having received feedback on the second draft of *PAM*, students completed a tailor-made test, as already explained in Section 5.2.3.3. For this test, the participants were allowed 15 minutes to complete it and did not receive any help. The participants were expected to correct the errors they had made in their drafts (see Appendix 4 for examples).

The four writings, the tailor-made tests and the questionnaires were analysed by an expert on statistics who helped the teacher/researcher to interpret the results.

After having explained how the data were collected, the next section will deal with how the data analysis was performed.

### 5.2.5 Data analysis

Quantitative data analysis was applied to the drafts together with the tailor-made test and the questionnaire. Due to the small sample, non-parametric tests were used. The statistical analysis, by means of the SPSS version 22, included, first, a within-subjects analysis between the first and the second draft of each story and for each group and each targeted grammatical feature; second, a between-groups for the first and the second drafts; third, a within-subjects analysis between *ABBY* and *PAM* for each group and each targeted grammatical feature, and finally, the analysis of the tailor-made test. A detailed explanation of each type of statistical analysis follows:

The within-subjects analysis between the first draft and the second draft was conducted by means of a Wilcoxon test<sup>9</sup> between the first and the second draft of *ABBY* for each targeted grammatical feature and the sum of the three types of errors contained in each draft. The same statistical test was applied to compare the first and second draft of *PAM* for each targeted grammatical feature and the sum of the errors contained in each draft. The aim of this analysis was to examine whether there existed any difference in errors between the first and the second drafts of both comic strips.

The between-groups comparison analysed the three groups for each targeted grammatical feature and the sum of all the errors of the first draft of *ABBY* and *PAM*, as well as the second draft of both writings using a Kruskal-Wallis one-way analysis of variance<sup>10</sup>. The purpose of

---

9. The Wilcoxon signed-rank test is a non-parametric statistical hypothesis test used to compare two related samples, matched samples, or repeated measurements on a single sample to assess whether their population mean ranks differ.

10. The Kruskal-Wallis one-way analysis of variance is a non-parametric method for testing whether samples originate from the same distribution. It is used for comparing two or more independent samples of equal or different sample sizes.



this analysis is to check whether there is any distinction in each type of error according to the feedback received.

The within-subjects analysis between *ABBY* and *PAM* was carried out by means of a Wilcoxon test of the three targeted grammatical features per group and each targeted grammatical feature between the first draft of *ABBY* and the first draft of *PAM* and the sum of the errors, and also with the second draft and the sum of the errors. The goal of this analysis was to verify whether there was any meaningful variation of errors between the first drafts of *ABBY* and *PAM* and the second drafts of the same comic strips.

Finally, a Wilcoxon test of all errors made in the four written productions in comparison to the tailor-made test was run. By doing so, it could be checked if there were any variations in the long-term between the students' performance at the beginning of the study and nearly five months later.

As mentioned earlier, due to our small number of subjects, nonparametric analyses were carried out. Therefore, a Wilcoxon test was applied (as an alternative to students' t-test<sup>11</sup> for paired samples) for a within-subjects comparison between two levels or moments. In turn, when the comparison was between groups and only two independent samples were involved, a Mann-Whitney U test was run (alternative to students' t-test for independent groups). Finally, when the number of groups was superior to two, a Kruskal-Wallis test was applied (alternative to ANOVA or analysis of variance, a test applicable to more than two independent groups).

In order to code data more efficiently, different shortened terms were used, as Table 6 below depicts:

Table 6. Full and shortened names for the data used in statistical analysis

| Name in full                | Shortened name |
|-----------------------------|----------------|
| First draft of <i>ABBY</i>  | ABBY_pre       |
| Second draft of <i>ABBY</i> | ABBY_post      |
| First draft of <i>PAM</i>   | PAM_pre        |

11. Student's t-test, in statistics, is a method of testing hypotheses about the mean of a small sample drawn from a normally distributed population when the population standard deviation is unknown.

|   |             |
|---|-------------|
| Second draft of <i>PAM</i>                        | PAM_post    |
| Sum of errors of the first draft                  | SUM_pre     |
| Sum of errors of the second draft                 | SUM_post    |
| Sum of the errors of the four drafts              | TAILOR_pre  |
| Tailor-made test                                  | TAILOR_post |
| Prepositions                                      | PREP        |
| Articles  | ART         |
| Past simple-present/past perfect tense difference | TEN         |

### 5.3 Chapter Summary

Chapter 5 featured a detailed account of the pilot study conducted to test the workability of GC. Thanks to the pilot study, some modifications were applied to improve the study design and feasibility of the final study. Then, the setting, participants, targeted grammatical features, instruments, data collection procedure and the different quantitative data analyses conducted for the final study were described.

After the discussion of all the components of the study, Chapter 6 will centre on the analysis of the results and their discussion.

## Chapter 6. Results and Discussion

As stated in the previous chapter, the aim of this study was to ascertain whether the use of computer-generated feedback had an effect on the students' language development in the course of five months. In addition, a comparison between computer-generated feedback and teacher's feedback was also intended. This way, by interpreting the data obtained by means of the students' compositions, the tailor-made test and the questionnaire, valuable information about the usefulness and effectiveness of both types of feedback could be obtained.

Before turning our attention to the results, it is worth noting that the subjects were all enrolled in the same course (2B2) after having passed the 1B2 level and their initial level of proficiency was similar a priori. To confirm this fact, a Kruskal-Wallis test for the three groups and each grammatical feature for ABBY\_pre was conducted, as Table 7 shows.

Table 7. Kruskal-Wallis H Test among groups for ABBY pre and effect size (partial eta square)

|               | Group | Mean | SD    | $X^2_2$ | P    | $\eta^2_p$ |
|---------------|-------|------|-------|---------|------|------------|
| ABBY_PREP_pre | 1     | 0.88 | 1.126 | 0.622   | .733 | 0.018      |
|               | 2     | 1.18 | 1.250 |         |      |            |
|               | 3     | 0.88 | 1.126 |         |      |            |
| ABBY_ART_pre  | 1     | 1.13 | 1.356 | 0.924   | .630 | 0.026      |
|               | 2     | 0.91 | 2.071 |         |      |            |
|               | 3     | 0.50 | 0.756 |         |      |            |
| ABBY_TEN_pre  | 1     | 0.25 | 0.463 | 2.300   | .317 | 0.105      |
|               | 2     | 1.09 | 1.758 |         |      |            |
|               | 3     | 0.38 | 0.518 |         |      |            |
| ABBY_SUM_pre  | 1     | 2.25 | 2.375 | 0.499   | .779 | 0.058      |
|               | 2     | 3.18 | 3.430 |         |      |            |
|               | 3     | 1.75 | 1.035 |         |      |            |

Note: Group 1: "Computer"; Group 2: "Teacher"; Group 3: "Control";  $X^2_2$ : Chi-square test with two degrees of freedom

As can be seen in Table 7, no meaningful differences between the three groups either in each type of error or in the total number of errors were observed. This absence confirms that, regardless of

the students' background (e.g., mother tongue, years studying English, etc.) or slight differences in level of proficiency, the participants' command on prepositions, articles and past simple-present/past perfect difference seems to be similar, since the number of errors committed is statistically nonsignificant.

After ascertaining that the three groups had a similar command of the English language before the study began, the RQs and hypotheses proposed will be addressed in the following sections.

## 6.1 Results and Discussion Related to Hypothesis 1

Based on previous studies (e.g., Coniam, 2009; El Ebyary & Windeatt, 2010; Heift & Schulze, 2007; Hutchison, 2007; Potter & Fuller, 2008), H1 suggested that the provision of individualised feedback on prepositions, articles and tenses by GC would report beneficial results. Despite the fact that the three targeted grammatical features chosen for the study share similarities between the students' L1 (Spanish, Valencian, Romanian) and FL (English), and thus they may help in their acquisition, (Cai, 2001; Collins, 2004; Ekiert, 2004; Giacalone, 2002; Ionin, Zubizarreta & Maldonado, 2008; Romaine, 2003; Snape et al., 2013), they can also be complex to learn because of the L1 influence or the differences between languages (Alonso et al., 2016; Celce-Murcia & Larsen-Freeman, 1999; Evans & Tyler, 2005; James, 2007; Jie, 2008; Master, 1988). In order to confirm or refute H1, some tests were conducted.

A Wilcoxon test was employed to compare each kind of error between both drafts of *ABBY*, both drafts of *PAM* and the sum of all the errors between both drafts of *ABBY* and both drafts of *PAM* before receiving feedback. In addition, some comparative within-subjects analysis with the number of errors committed on each targeted grammatical feature before and after receiving feedback is also included in the following lines. The aim of the Wilcoxon test is to identify whether there are any variations in errors committed between drafts and the aim of the comparative analysis is to ascertain if the feedback received is beneficial and effective.

Each targeted grammatical feature was analysed individually for each group so that H1 (*Grammar Checker will help students improve the use of prepositions, articles and past simple-present/past perfect difference*) can be tested. The first type of error we address is the use of prepositions.

### 6.1.1 Prepositions

As for prepositions, the participants in the three groups made different types of errors (wrong use of preposition, unnecessary use or missing preposition), which may be the consequence of lack of knowledge or L1 transfer, among others. Table 8 shows the results of the comparison between ABBY\_pre and ABBY\_post and PAM\_pre and PAM\_post for errors with prepositions in Group 1:

Table 8. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 1 and for prepositions

|                | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|----------------|------|-------|-----------|------|-------|
| ABBY_PREP_pre  | 0.88 | 1.126 | 1.134     | .257 | 0.411 |
| ABBY_PREP_post | 1.38 | 1.302 |           |      |       |
| PAM_PREP_pre   | 1.13 | 1.126 | 1.190     | .234 | 0.457 |
| PAM_PREP_post  | 0.63 | 1.061 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Table 8 shows the lack of statistical differences between the first and second draft of each comic strip ( $Z_{(w)}=1.134$ ,  $p>0.5$ ;  $Z_{(w)}=1.190$ ,  $p>0.5$ ). In addition, the effect size (Cohen's d) was also calculated to examine the magnitude of the difference in errors after having written the second draft. Therefore, taking into account the fact that for Cohen's  $d=0.20$  is a small effect size,  $d=0.50$  a medium effect size and  $d=0.80$  a large effect size, Table 8 shows that the effect sizes for prepositions in Group 1 range from small to medium, which confirms the lack of statistical significance between the first and the second drafts of each comic strip.

In view of such lack of effectiveness between drafts before receiving the computer feedback, a further analysis after having received it was conducted. Figures 19 and 20 show the distribution of errors students made on prepositions in each draft for Group 1 before and after receiving feedback ( $n=32$ ,  $n=31$ , respectively). As indicated before, three different kinds of prepositional errors were identified: unnecessary, missing and wrong preposition.

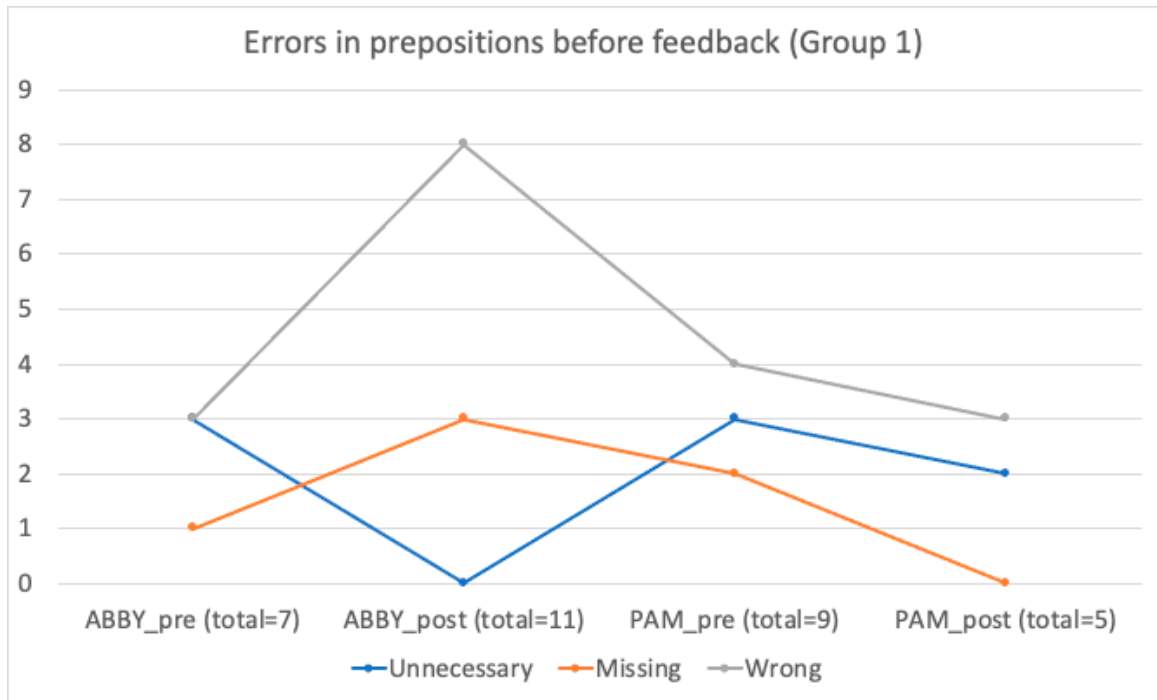


Figure 19. Errors in prepositions before feedback (Group 1)

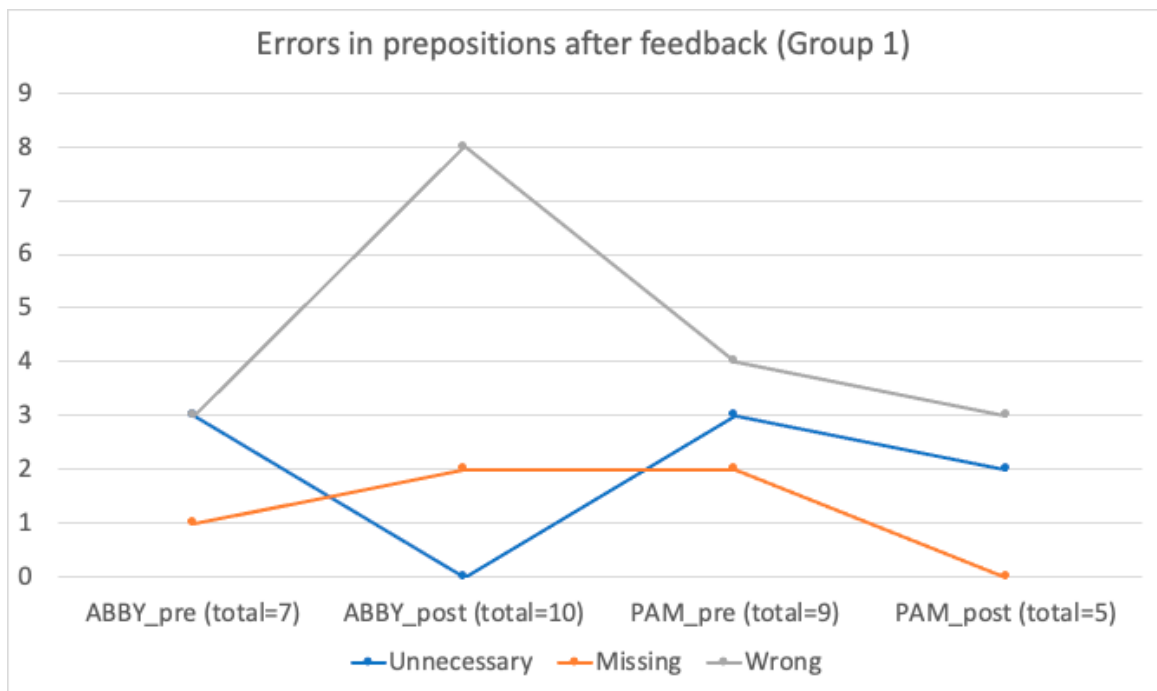


Figure 20. Errors in prepositions after feedback (Group 1)

As can be seen from the above figures, errors committed before and after receiving feedback by GC amount to almost the same number, which can be regarded as positive because students' performance did not decline. Only a decrease of one mistake can be noticed in ABBY\_post in the case of missing prepositions. Thus, after analysing error by error we concluded that GC did not mark those errors and consequently students did not correct them. In order to understand why they were not highlighted at least in yellow<sup>12</sup>, all prepositional errors were clear inconsistencies except for "At the end", which is a combination that is grammatically correct but not as it was used by two of the students as they should have used "In the end". GC did help one of the students to modify "by the other hand" for "on the other hand" but, in general, it failed to help students notice evident errors because they were not highlighted.

As for the progression between drafts, the incidence of errors slightly increased between ABBY\_pre and ABBY\_post and slightly decreased between PAM\_pre and PAM\_post. In view of these findings, and in order to have a better perspective on the effectiveness of computer-generated feedback, the same analysis of errors was carried out with Group 2 and Group 3. Table 9 presents the errors in prepositions for Group 2.

Table 9. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 2 and for prepositions

|                | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|----------------|------|-------|-----------|------|-------|
| ABBY_PREP_pre  | 1.18 | 1.250 | 0.962     | .336 | 0.300 |
| ABBY_PREP_post | 1.55 | 1.214 |           |      |       |
| PAM_PREP_pre   | 1.45 | 1.572 | 0.730     | .465 | 0.195 |
| PAM_PREP_post  | 1.18 | 1.168 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

The results of Table 9 show again a lack of statistical significance between ABBY\_pre and ABBY\_post, and PAM\_pre and PAM\_post ( $Z_{(w)}=0.962$ ,  $p>.05$ ;  $Z_{(w)}=0.730$ ,  $p>0.5$ ). That is to say, no significant improvement was noticed between the first draft and the second draft. In order to delve into the lack of variation between drafts, Figures 21 and 22 illustrate the number of errors for Group 2 before and after receiving the teacher's feedback (n=59, n=28, respectively):

12. Colours used by GC were: red (very suspicious), orange (suspicious), yellow (slightly suspicious).

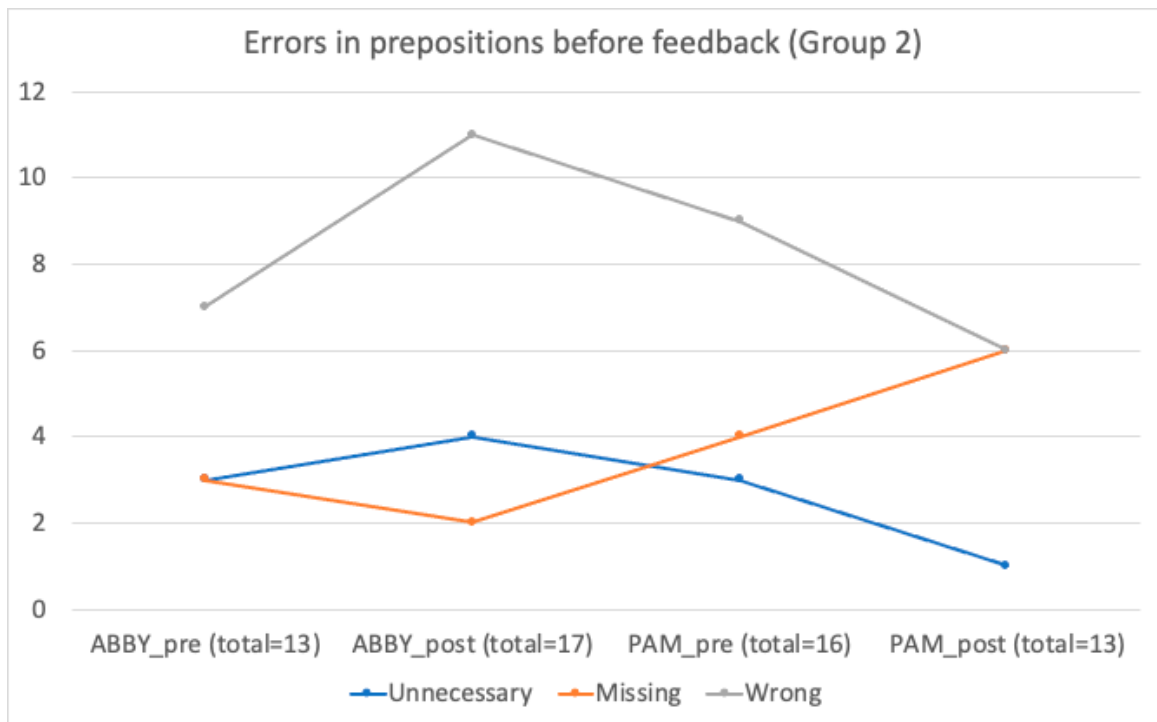


Figure 21. Errors in prepositions before feedback (Group 2)

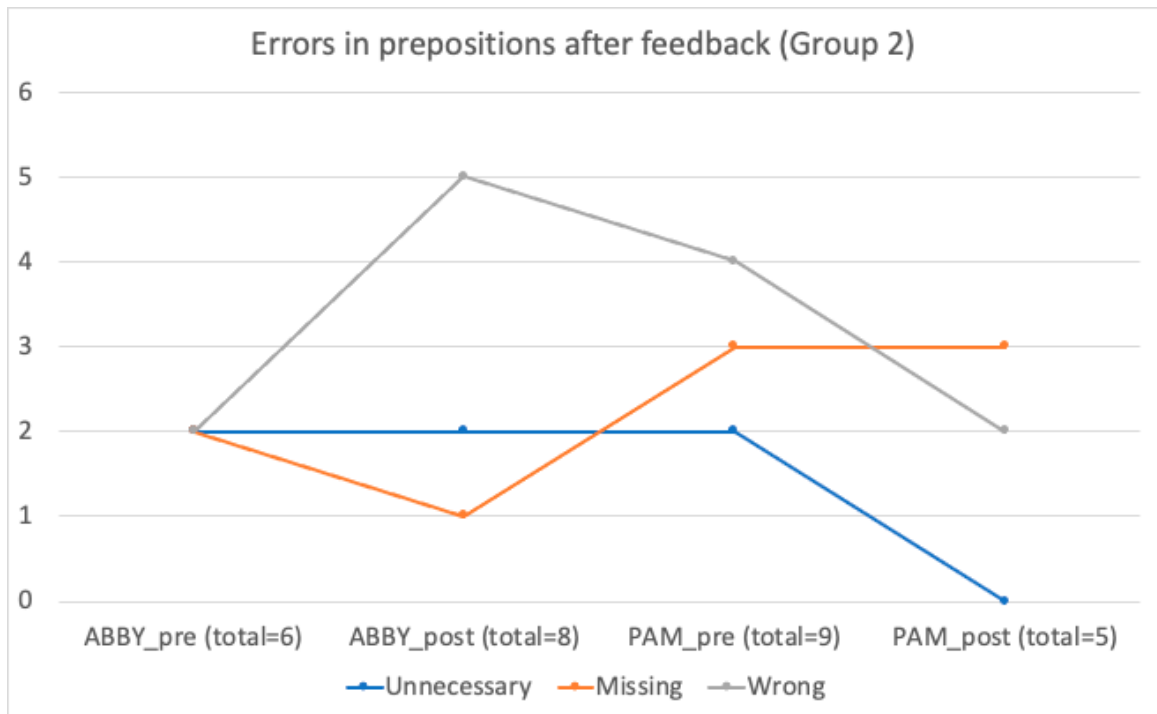


Figure 22. Errors in prepositions after feedback (Group 2)



On the one hand, Figures 21 and 22 show the same evolution as Figures 19 and 20, that is to say, a slight increase in the occurrence of errors from *ABBY\_pre* to *ABBY\_post* and a slight decrease between *PAM\_pre* and *PAM\_post*. On the other hand, after having received the teacher's feedback, a different picture from the computer-generated feedback emerged as the errors in all the drafts were reduced by half, especially in the wrong use of prepositions. Reduction of errors in the missing and unnecessary use of prepositions exists, but it hardly reaches significance. From the errors still committed after having received the teacher's feedback, the errors with the unnecessary use of prepositions were left uncorrected. As for errors with missing prepositions, only one was modified incorrectly in *ABBY\_pre* and the other ones were left uncorrected. Finally, from the errors with the wrong use of prepositions, two were modified incorrectly in *ABBY\_pre*, two in *ABBY\_post* and one in *PAM\_post*. The other ones were left uncorrected. It is interesting to note that, after having received the teacher's feedback, prepositional errors classified as missing or unnecessary were left uncorrected as students had problems to figure out the inconsistency. Perhaps, in case they had received some metalinguistic feedback as GC did, they could have amended the inaccuracy. On the contrary, when the preposition used was incorrect, they tried to modify it by providing an alternative even though it was inaccurate. In this case, by receiving some suggestions or metalinguistic comments the revision could have been even more successful. This fact may imply that, depending on the kind of error, different types of feedback can be applicable. For instance, for unnecessary prepositions, just by crossing them out would be understandable enough for the students. For missing prepositions, a symbol indicating there is something missing would be clearer as well. Finally, for wrong prepositions, some metalinguistic comments could be helpful enough to modify them correctly.

As for the progression between drafts, the unnecessary use of prepositions remained constant except in *PAM\_post* where no errors were made. The errors in missing prepositions underwent a slight increase in *PAM* versions but it was not significant. Finally, the wrong use of prepositions experienced an increase in *ABBY\_post* but returned to the same incidence in *PAM\_post*. Once again, this fluctuation was not considered significant.

In light of our findings, it can be concluded that, in contrast with computer-generated feedback, the students' performance improved after having received the teacher's feedback. Although the incidence of errors remained similar between drafts, the effects of the teacher's feedback was noticeable but not statistically significant (see Table 9).

Finally, Table 10 represents the analysis for Group 3 (control group). The results confirm the absence of statistically significant differences between the first and the second draft of each comic strip ( $Z_{(w)}=0.333$ ,  $p>.05$ ;  $Z_{(w)}=0.707$ ,  $p>0.5$ ) as the incidence of errors was very similar.

Table 10. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 3 and for prepositions

|                | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|----------------|------|-------|-----------|------|-------|
| ABBY_PREP_pre  | 0.88 | 1.126 | 0.333     | .739 | 0.126 |
| ABBY_PREP_post | 1.00 | 0.926 |           |      |       |
| PAM_PREP_pre   | 1.00 | 0.756 | 0.707     | .480 | 0.357 |
| PAM_PREP_post  | 1.38 | 1.302 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

In order to examine data more thoroughly, Figures 23 and 24 illustrate the number of errors for Group 3 before and after having revised their writings ( $n=34$ ,  $n=33$ , respectively). As mentioned before, this group did not receive any comments or marks on the three targeted grammatical features.

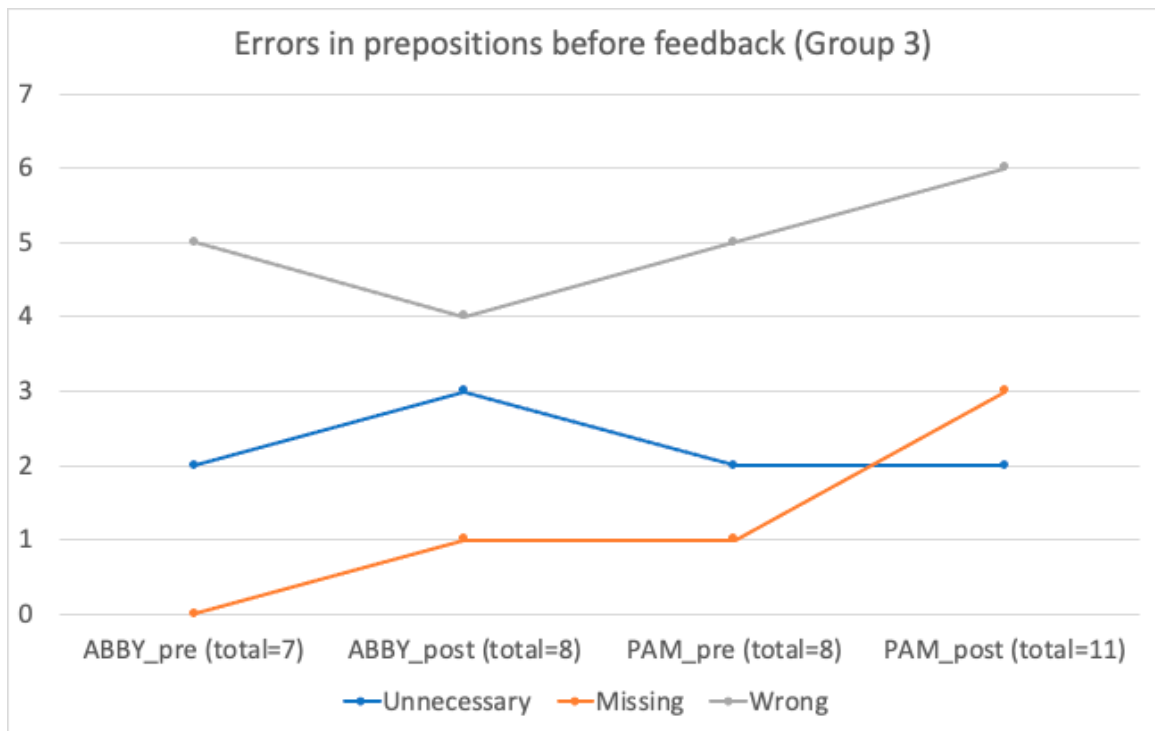


Figure 23. Errors in prepositions before feedback (Group 3)

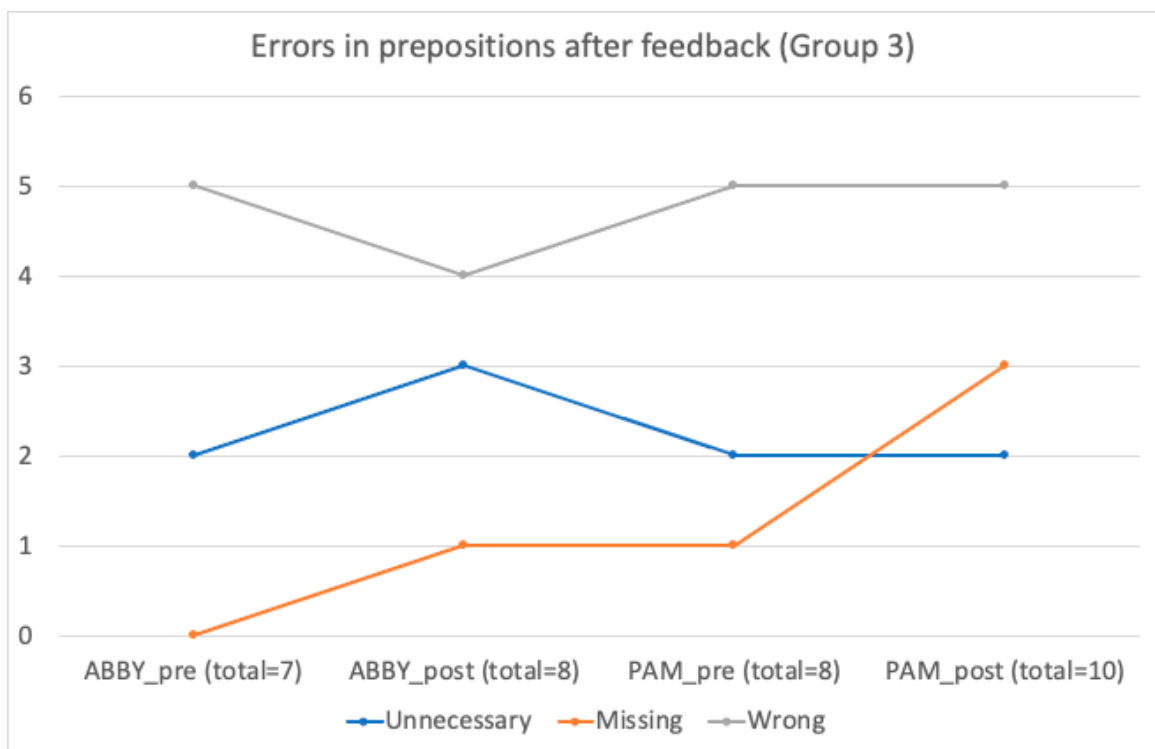


Figure 24. Errors in prepositions after feedback (Group 3)

As can be observed, the number of errors before and after having revised their writings is the same, except in PAM\_post where one student, although he did not receive any feedback on prepositions, corrected the wrong use of one preposition.

With regard to the progression between drafts, the incidence of errors remained similar. Only in the wrong use of prepositions it is worth mentioning that in ABBY\_pre no error was spotted, but there were three occurrences in PAM\_post. Still, these data were not statistically significant.

According to the literature, prepositions are highly used in a language and are difficult to master (Casas-Pedrosa, 2005; Chodorow, Tetreault & Han, 2007; Díez-Bedmar & Casas-Pedrosa, 2011; Nicholls, 2003; Sinclair, 1991; Watcyn-Jones & Allsop, 1990). In fact, Lindstromberg (2001) claims that less than 10% of advanced English learners can use and understand prepositions correctly. The morphological, syntactic and semantic characteristics of prepositions, which do not necessarily match from language to language (Celce-Murcia & Larsen-Freeman, 1999; James, 2007; Jie, 2008) and whose use in the FL may be influenced by the L1 (e.g., Alonso-Alonso, 1997; Bazo-Martínez, 2001; Benítez-Pérez & Simón-Granda, 1990; Cowan et al., 2003; García-Gómez & Bou-Franch, 1992; Jiménez-Catalán, 1996; Koosha & Jafarpour, 2006; Tanimura et al., 2004), may be part of the cause for prepositional errors. In line with previous studies that illustrate that the use of prepositions represents one of the most common sources of error for learners of English (e.g., Bueno-González, 1992; Dagneaux et al., 1998; Izumi et al., 2004; Nicholls, 2003; Tetreault & Chodorow, 2008), from the three grammatical errors analysed in the present study, the occurrence of prepositional errors is higher than errors with articles and tenses, as it will be shown in subsequent sections in this dissertation.

In order to better understand why some prepositional errors may be easier to identify and correct, some examples of the incorrect uses of prepositions by the participants in the study have been extracted from our data:

Example 1. Few mothers want a boy like him with\* their daughters (wrong preposition)

Example 2. But instead\* being rude, Richard showed that he was lovely (missing preposition)

Example 3. There was a boy called Richard of\*whom ABBY was in love with (unnecessary preposition)

Example 4. Based in\* his own experience he had a bad feeling (wrong preposition).

Errors in Examples 1 and 2 are difficult to explain because they share similarities between L1 and FL (for: *para*, instead of: *en vez de*). L1 influence from Spanish should have helped the students avoid those mistakes. As an alternative, we may think they were due to a lapse of concentration or typing slip in the case of error 2. Error 3 comes from L1 transfer. The student knows the verb is to fall *in love with* but even so she uses the preposition *of*, which is a literal translation from Spanish/Valencian.

Example 4 is a clear illustration of the semantic differences between Spanish/Valencian and English. Whereas there is only one possibility in Spanish/Valencian (*en*), there are several in English (e.g., *in*, *inside*, *on*, etc.), which increases the likelihood of committing an error and literal translation (Díez-Bedmar & Casas-Pedrosa, 2011). Thus, in the case of wrong use of prepositions, our findings confirm Ballesteros et al.'s (2005) claim that this type of error was the most frequent grammatical error, amounting to 65% of the total number of errors. What is more, the number of errors on prepositions could have been higher if the students had used two-word verbs (i.e., phrasal verbs) instead of the Latin counterpart (Celce-Murcia & Larsen-Freeman, 1999), since students consider them easier to employ in their writings.

Errors in Examples 3 and 4 are the most common ones (L1 transfer and English/Spanish semantic system), since their origin lies in the similarities and differences between languages. For these reasons, they are also the most difficult ones to master and, as mentioned before, the wrong use of prepositions has outnumbered the rest in this study.

Turning out attention to errors on articles, the same analysis was performed, the results of which are shown in the next section.

### 6.1.2 Articles

The fact that the students' L1 (Spanish/Valencian/Romanian) and their TL (English) are +ART languages may make the acquisition of English articles easier, since positive transfer may occur (Ekiert, 2004; Hawkins et al., 2006; Ionin & Montrul, 2010; Master, 1990, 1997;

Park, 1996; Romaine, 2003; Snape et al., 2013). On the contrary, as explained in Section 1.5.1, -ART languages can lead to negative transfer and, therefore, hinder acquisition when learning a language which does have one (Avery & Radišić, 2007; Chen, 2000; Ionin, Ko & Wexler, 2004; White, 2008; Zdorenko & Paradis, 2007, 2008, 2012). For example, dealing with definiteness may be difficult for speakers of -ART languages as they have to create this category. However, in the case of EFL Spanish learners, they seem to be quite accurate in the use of articles even from the beginning, as Ionin et al. (2008) claim.

In order to check the effectiveness of GC to improve the students' use of articles, a Wilcoxon test was conducted. Table 11 shows the results of the comparison between ABBY\_pre and ABBY\_post and PAM\_pre and PAM\_post for errors with articles in Group 1:

Table 11. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 1 and for articles

|               | Mean | SD    | $Z_{(w)}$ | $p$   | d     |
|---------------|------|-------|-----------|-------|-------|
| ABBY_ART_pre  | 1.13 | 1.356 | 0.000     | 1.000 | 0.000 |
| ABBY_ART_post | 1.13 | 2.100 |           |       |       |
| PAM_ART_pre   | 1.50 | 1.604 | 0.680     | .496  | 0.228 |
| PAM_ART_post  | 1.13 | 1.642 |           |       |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Table 11 shows the lack of statistical differences between the first and second draft of each comic strip ( $Z_{(w)}=0.000$ ,  $p>.05$ ;  $Z_{(w)}=0.680$ ,  $p>.05$ ), with no significant improvement. In addition, the effect size (Cohen's d) was also calculated to examine the magnitude of the difference in errors after having written the second draft. Therefore, taking into account the fact that Cohen's  $d=0.20$  is a small effect size,  $d=0.50$  a medium effect size and  $d=0.80$  a large effect size, Table 11 shows that the effect sizes range from small to medium. As with prepositions, in view of such a lack of effectiveness between drafts before receiving the computer feedback, a further analysis after having received it was conducted. Figures 25 and 26 show the distribution of errors students made with articles in each draft for Group 1 before and after receiving feedback ( $n=39$ ,  $n=38$ , respectively). Three different kinds of article errors were identified: unnecessary, missing and wrong articles.

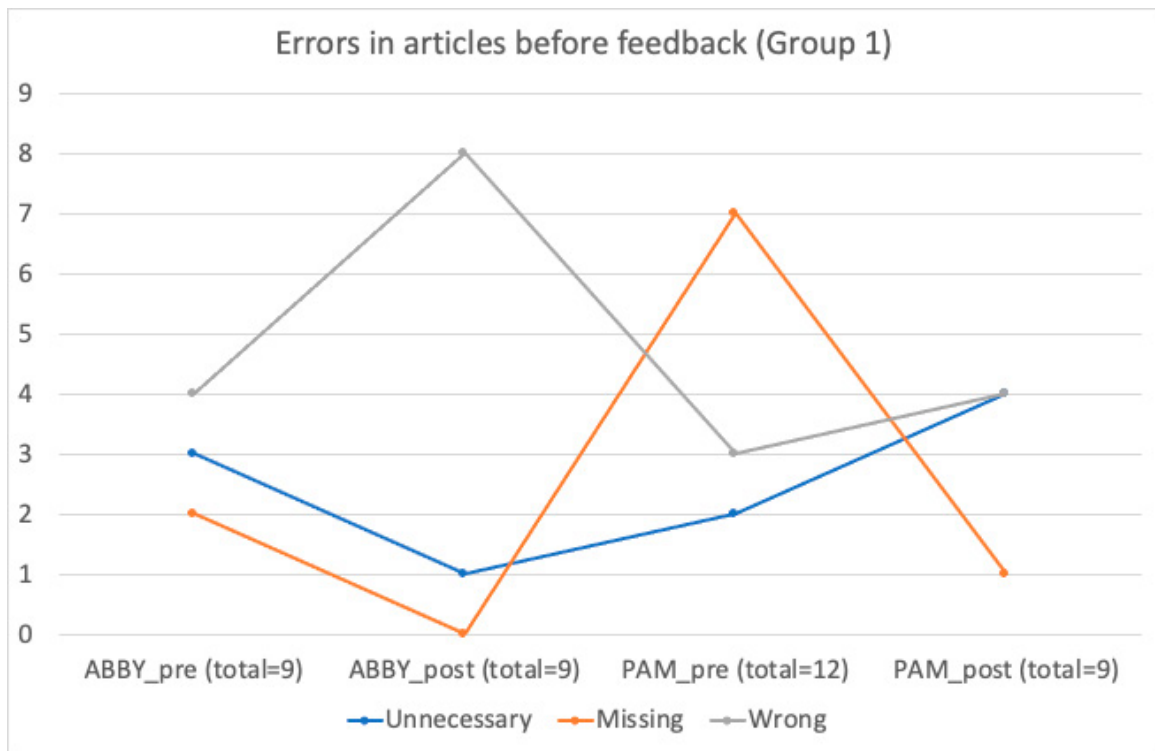


Figure 25. Errors in articles before feedback (Group 1)

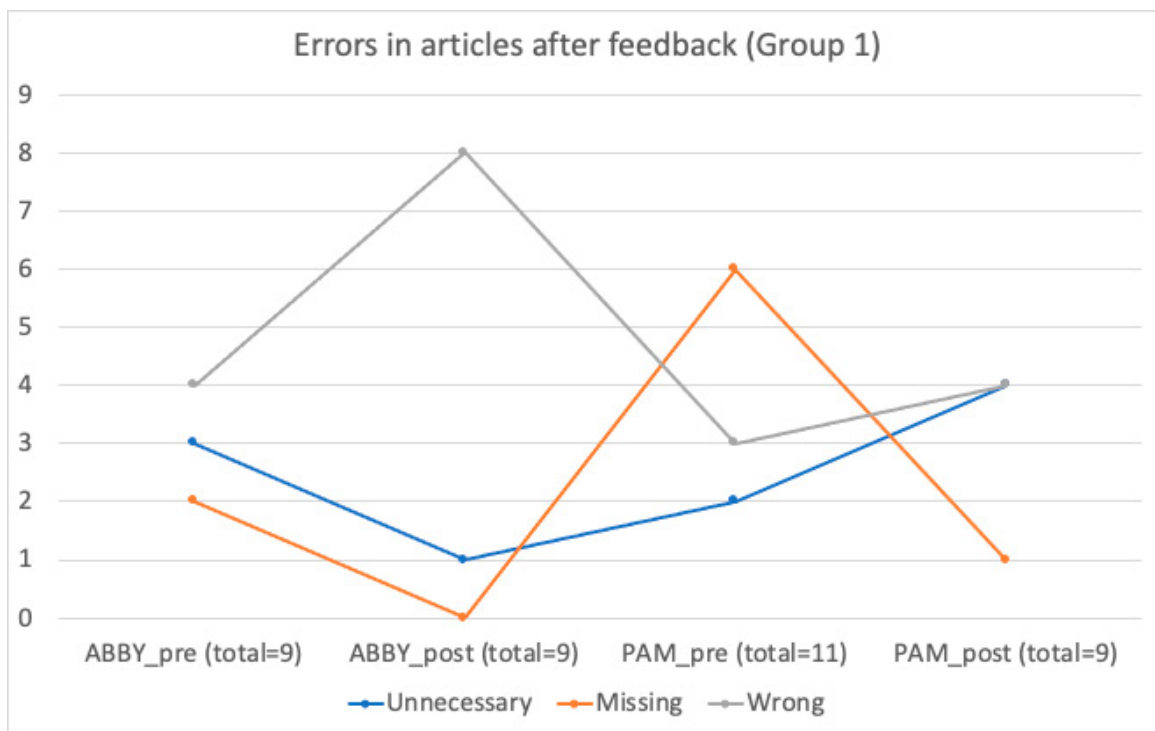


Figure 26. Errors in articles after feedback (Group 1)

Figures 25 and 26 do not represent any improvement in the use of articles as they amount to almost the same number of errors. Only some decrease can be noticed in PAM\_pre in the case of missing articles. After having analysed errors one by one before and after receiving feedback, we can draw the conclusion that GC did not mark most of the errors on articles and consequently students did not correct them. In order to understand why they were not highlighted even in yellow, all the article errors were clear inconsistencies such as “an advice” or “as waitress”. In comparison with errors in prepositions by the same group, the students seem to follow the same tendency, that is, no improvement.

As for Group 2, Table 12 presents the results of the comparison between ABBY\_pre and ABBY\_post and PAM\_pre and PAM\_post for errors with articles:

Table 12. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 2 and for articles

|               | Mean | SD    | $Z_{(w)}$ | $p$         | d     |
|---------------|------|-------|-----------|-------------|-------|
| ABBY_ART_pre  | 0.91 | 2.071 | 1.006     | .314        | 0.059 |
| ABBY_ART_post | 1.00 | 0.632 |           |             |       |
| PAM_ART_pre   | 1.36 | 1.286 | 2.232     | <b>.026</b> | 1.024 |
| PAM_ART_post  | 0.36 | 0.505 |           |             |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen’s d effect size. In bold, statistically significant differences

As observed in Table 12, there are statistically significant differences between PAM\_art\_pre and PAM\_art\_post ( $Z(W)=2.232$ ,  $p=.026$ ;  $d=1.024$ ), pointing to a significant decrease of errors. Figures 27 and 28 show the number of article errors in Group 2 ( $n=40$ ,  $n=27$ , respectively) before and after receiving feedback.



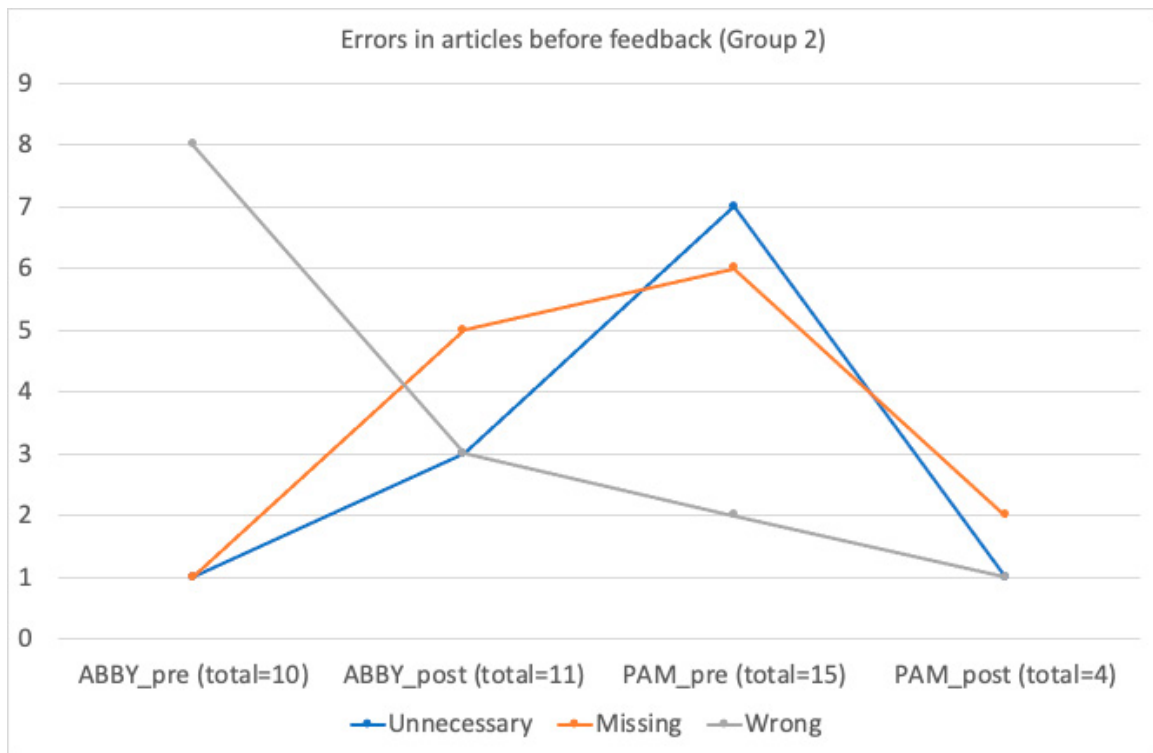


Figure 27. Errors in articles before feedback (Group 2)

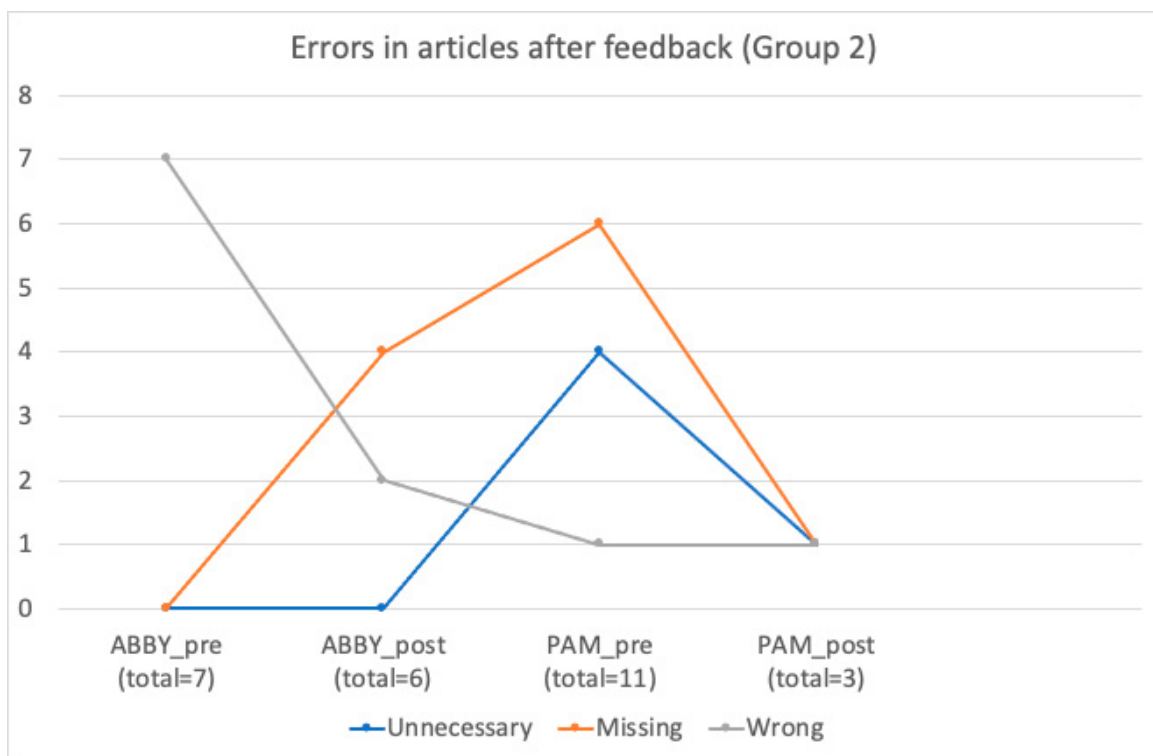


Figure 28. Errors in articles after feedback (Group 2)

As illustrated in the above figures, the three kinds of article errors experience some decrease after having received the teacher's feedback. If the four drafts are compared, the wrong use of articles experiences a steady decrease whereas the unnecessary use and the missing articles increase from ABBY\_pre to PAM\_pre and undergo an important decrease in PAM\_post. However, the number of errors committed in ABBY\_pre and PAM\_post are almost the same for unnecessary and missing articles, but much lower in the wrong use of articles. Then, despite the increase mentioned before, the teacher's feedback seems to have been effective, first, to avoid a further increase in unnecessary and missing article errors and, second, to decrease the wrong use of articles from the first draft.

As for Group 3, Table 13 shows the statistics of article errors between ABBY\_pre and ABBY\_post and PAM\_pre and PAM\_post:

Table 13. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 3 and for articles

|               | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|---------------|------|-------|-----------|------|-------|
| ABBY_ART_pre  | 0.50 | 0.756 | 0.816     | .414 | 0.304 |
| ABBY_ART_post | 0.75 | 0.886 |           |      |       |
| PAM_ART_pre   | 0.38 | 0.744 | 0.333     | .739 | 0.185 |
| PAM_ART_post  | 0.50 | 0.535 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

The results reveal the absence of statistical differences between drafts ( $Z_{(w)}=0.816$ ,  $p > .05$ ,  $Z_{(w)}=0.333$ ,  $p > 0.5$ ). Figure 29 presents the number of article errors in Group 3.

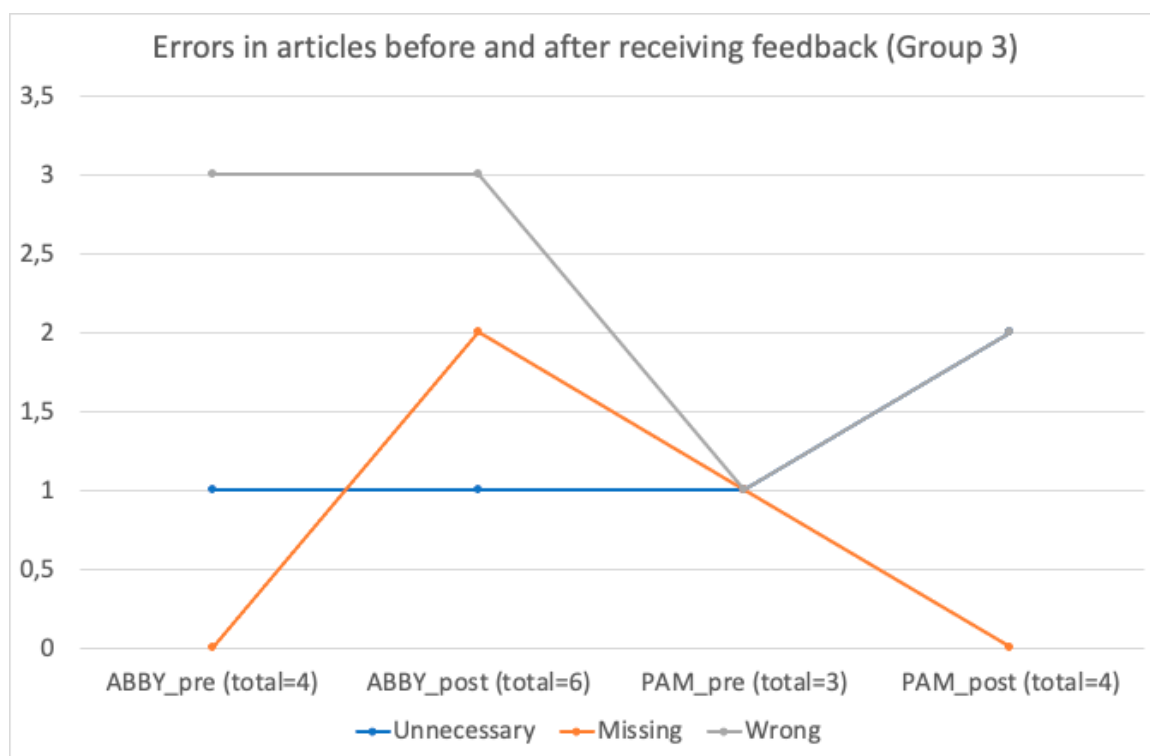


Figure 29. Errors in articles before and after receiving feedback (Group 3)

Despite the variability of each type of error (i.e., the incidence of unnecessary use of articles remains the same and slightly increases in PAM\_post, the incidence of missing articles increases in ABBY\_post and then decreases to zero, and the incidence of wrong use of articles decreases in PAM\_pre but then slightly increases in PAM\_post), their occurrence is very similar throughout the study. Considering that Group 3 did not receive any feedback on article errors, a lack of statistically significant difference was expected.

In regard to errors in articles, as explained in Section 1.5.1, article omissions in obligatory contexts, substitution of one article for another or oversupply are common errors among English students (García-Mayo, 2009; Ionin & Montrul, 2010; Ionin et al., 2008; Snape et al., 2009), as shown below in some examples of the incorrect use of articles extracted from our data:

Example 5. They invited grandma for a\* dinner. (unnecessary)

Example 6. She had \* high level of blood sugar. (missing)

Example 7. The\* grandmother was also reluctant to Pam's decision. (wrong)

The above examples represent the three categories examined in the study. Errors in Examples 5 and 6 were less frequent, whereas the error in Example 7 was much commonly found in our data. The analysis conducted to answer RQ1 partially confirms the results of previous studies (e.g., García-Mayo, 2009; Ionin & Montrul, 2010) that article errors committed by Spanish learners are mainly article omission and article misuse or substitution. Our findings corroborate those results, especially in ABBY\_pre and ABBY\_post.

According to the Article Choice Parameter and the Fluctuation Hypothesis, the fact that both Spanish and English have an article system makes transfer from L1 to L2 possible without the need to fluctuate (Ionin et al., 2008). In line with what Huebner (1983) and Master (1997) term “the-flooding”, Spanish speakers tend to overgeneralise the definite article *the*, especially with family words (e.g., the mother, the grandmother), instead of a possessive pronoun (e.g., her mother) or no article (e.g., grandma, mum).

In conclusion, contrary to Chodorow et al. (2010), who investigated short-term effects of another software (i.e., *Criterion*) on writing errors in articles and prepositions, and revealed a positive effect of *Criterion* on the use of feedback to make valid corrections, the present study did not identify either any significant differences or any relevant decrease in errors by means of GC. As explained before, the computer program did not mark many of the errors and, consequently, students did not make any modifications.

The next section deals with the third targeted grammatical feature. The same analyses were performed to analyse the effectiveness of GC with the dichotomy present perfect-simple past.

### 6.1.3 Tenses

Some evidence on the influence of L1 in the acquisition of tenses in English has been reported (e.g., Cai, 2001; Collins, 2004; Giacalone, 2002), pointing to the fact that speakers of languages with tenses (for example, Spanish, French and German) may find it easier to use them than speakers of tenseless languages (e.g., Chinese or Japanese). However, there is still scarce research into the acquisition of English perfect tenses and L1 influence.

To examine if the participants are effectively assisted by GC in order to use present/past perfect and past simple correctly, a Wilcoxon test was conducted. Table 14 shows the results of the comparison between ABBY\_pre and ABBY\_post and PAM\_pre and PAM\_post for errors with tenses in Group 1:

Table 14. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 1 and for tenses

|               | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|---------------|------|-------|-----------|------|-------|
| ABBY_TEN_pre  | 0.25 | 0.463 | 1.841     | .066 | 0.744 |
| ABBY_TEN_post | 1.50 | 2.330 |           |      |       |
| PAM_TEN_pre   | 0.13 | 0.354 | 0.577     | .564 | 0.292 |
| PAM_TEN_post  | 0.25 | 0.463 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Table 14 shows the lack of statistical differences between the first and second draft of each comic strip for tenses ( $Z_{(w)}=1.841$ ,  $p>.05$ ;  $Z_{(w)}=0.577$ ,  $p>0.5$ ). In addition, the effect size (Cohen's d) was also calculated to examine the magnitude of the difference in errors after having written the second draft. Table 14 shows that the effect size for ABBY\_TEN\_pre and ABBY\_TEN\_post is .744, a value that, although not statistically significant, has a large effect size.

The same analysis was conducted with Group 2 and Group 3 to check the effectiveness of computer-generated feedback (see Table 15).

Table 15. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 2 and for tenses

|               | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|---------------|------|-------|-----------|------|-------|
| ABBY_TEN_pre  | 1.09 | 1.758 | 1.725     | .084 | 0.713 |
| ABBY_TEN_post | 0.18 | 0.405 |           |      |       |
| PAM_TEN_pre   | 0.36 | 0.674 | 0.272     | .785 | 0.113 |
| PAM_TEN_post  | 0.27 | 0.905 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Despite the lack of statistically significant differences ( $Z_{(w)}=1.725$ ,  $p>0.5$ ;  $Z_{(w)}=0.272$ ,  $p>0.5$ ), the effect size was also calculated and similar to Group 1, the value obtained in ABBY\_TEN\_

pre vs. ABBY\_TEN\_post ( $d=.713$ ), although not statistically significant, stands out for having a value close to 0.80. That is to say, there exists a noticeable reduction of errors between ABBY\_pre and ABBY\_post. Finally, Table 16 illustrates the same analysis for Group 3.

Table 16. Wilcoxon signed-rank test for Pre vs. Post and effect size for Group 3 and for tenses

|               | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|---------------|------|-------|-----------|------|-------|
| ABBY_TEN_pre  | 0.38 | 0.518 | 0.378     | .705 | 0.185 |
| ABBY_TEN_post | 0.50 | 0.756 |           |      |       |
| PAM_TEN_pre   | 0.63 | 0.518 | 1.000     | .317 | 0.247 |
| PAM_TEN_post  | 0.38 | 0.744 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen’s d effect size

In this case, taking into account that this group did not receive any feedback on tenses, students did not show any significant improvement ( $Z_{(w)}=0.378, p >.05; Z_{(w)}=1.000, p>0.5$ ).

As with prepositions and articles, a further analysis after having received feedback was conducted. Figures 30 and 31 show the distribution of errors students made with tenses in each draft for each group before and after receiving feedback.

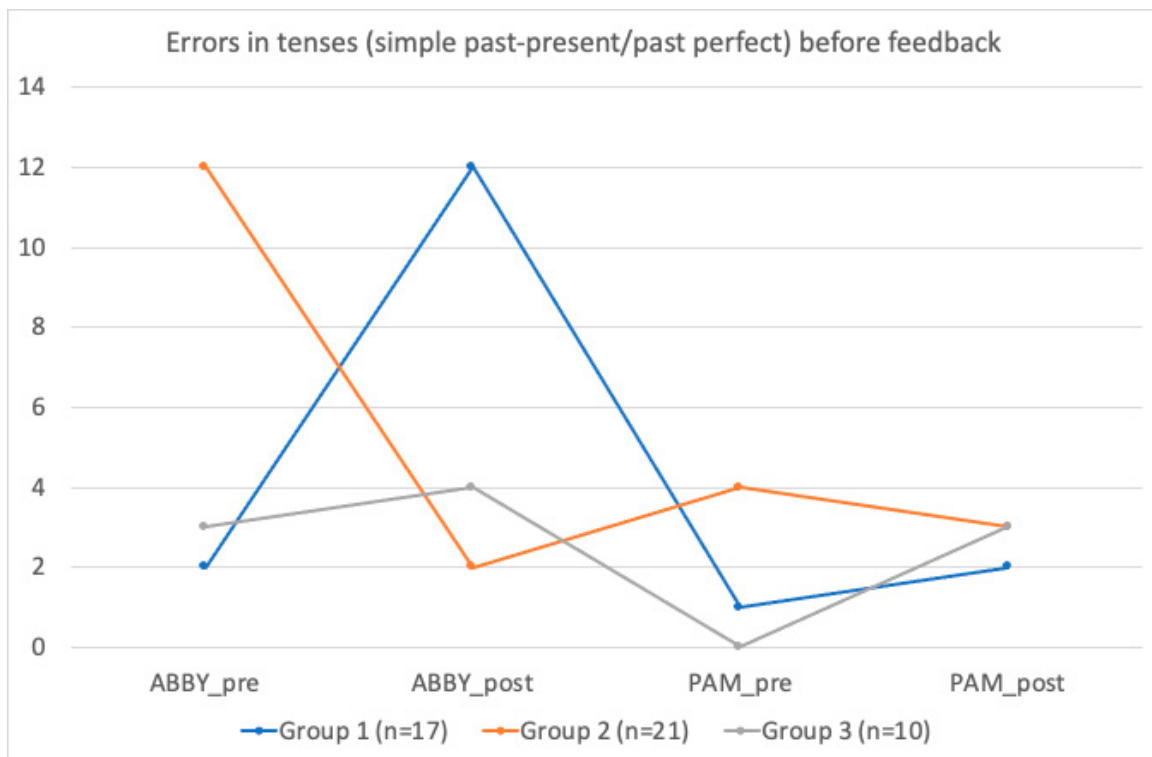


Figure 30. Errors in tenses (simple past-present/past perfect) before feedback.

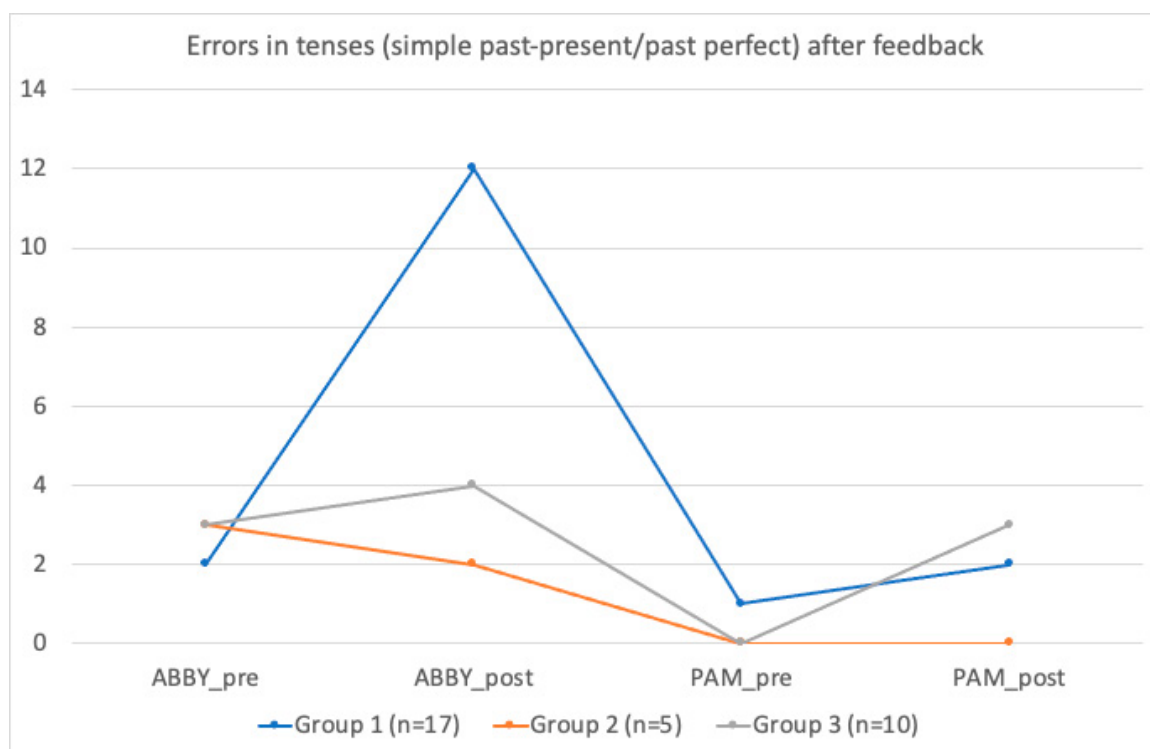


Figure 31. Errors in tenses (simple past-present perfect) after feedback

With regard to the errors students made with tenses (i.e., difference past-present/past perfect), it can be noticed that the incidence before correction between ABBY\_pre and PAM\_post in Group 1 and Group 3 remained the same and in Group 2 it decreased considerably. However, similar to articles and prepositions, the incidence shown in Figure 31 for Group 1 and Group 3 did not vary at all, whereas in Group 2 it decreased considerably in ABBY\_pre and slightly in PAM\_pre and PAM\_post. Again, we may attribute this finding to the fact that the incidence in Group 1 did not decrease because the errors were not marked by GC. The tenses used were grammatically correct but not appropriate for the context. Thus, the software was not able to spot those mistakes as it is based on bigrams and not meaning. As for Group 3, errors were not modified because they did not receive any feedback on tenses.

In regard to errors related to the difference between simple past-present/past perfect tense, it must be noted that they were the least frequent of the three targeted categories. Although the use of time expressions may facilitate the choice between present perfect and simple past (Bulut, 2011), their absence may complicate it too. In addition, as explained in Section 1.5.3, L1 transfer can help or hinder the acquisition of tenses as a result of the students' mother

tongue. Hereafter are some examples of the incorrect use of the simple past-present/past perfect extracted from our data:

Example 8: Grandma has advised\* Abby to be patient with her mother.

Example 9: After the concert, they have been talking and laughing\* ...

Example 10: I had to let Grandma know what happened\*.

Example 11: Pam wrote a short letter to ... telling her what concerned\* Pam lately.

The above examples represent some of the situations in which students have problems distinguishing between the simple and perfect aspects. In Examples 8 and 9 the action is past and finished and, therefore, using the present perfect makes no sense and confuses the reader. However, errors in Examples 10 and 11 typically occur when literal translation is used (i.e., negative transfer). Although the past perfect in Example 10 is also used in Spanish, Spanish speakers tend to use simple past more often, especially when there is no temporal particle. It is an error that has taken place often in the study. However, Example 11 does include a temporal word that forces the use of past perfect and still negative transfer occurs. Although the simple past would be used in Spanish, English rules require the past perfect in this situation.

After having analysed the errors committed in each draft, a further analysis was run globally, that is, we focused on the sum of the four drafts per group in order to examine data from another perspective which could draw different results.

### 6.1.4 Summation

An analysis of the sum of the total number of errors in the four drafts taking into account the three targeted grammatical features together per group was conducted in order to have a general perspective of the effectiveness of the two different types of feedback. The effect size was also calculated for each group (see Tables 17, 18 and 19).

Table 17. Wilcoxon signed-rank test and effect size for Group 1 for SUM\_Pre vs. SUM\_Post for *ABBY* and *PAM*

|              | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|--------------|------|-------|-----------|------|-------|
| ABBY_SUM_pre | 2.25 | 2.375 | 1.510     | .131 | 0.547 |



|               |      |       |       |      |       |
|---------------|------|-------|-------|------|-------|
| ABBY_SUM_post | 4.00 | 3.854 |       |      |       |
| PAM_SUM_pre   | 2.75 | 2.435 | 0.530 | .596 | 0.364 |
| PAM_SUM_post  | 2.00 | 1.604 |       |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Table 18. Wilcoxon signed-rank test and effect size for Group 2 for SUM\_Pre vs. SUM\_Post for *ABBY* and *PAM*

|               | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|---------------|------|-------|-----------|------|-------|
| ABBY_SUM_pre  | 3.18 | 3.430 | 0.157     | .876 | 0.170 |
| ABBY_SUM_post | 2.73 | 1.489 |           |      |       |
| PAM_SUM_pre   | 3.18 | 2.359 | 1.794     | .073 | 0.710 |
| PAM_SUM_post  | 1.82 | 1.328 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Table 19. Wilcoxon signed-rank test and effect size for Group 3 for SUM\_Pre vs. SUM\_Post for *ABBY* and *PAM*

|               | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|---------------|------|-------|-----------|------|-------|
| ABBY_SUM_pre  | 1.75 | 1.035 | 0.632     | .527 | 0.429 |
| ABBY_SUM_post | 2.25 | 1.282 |           |      |       |
| PAM_SUM_pre   | 2.00 | 1.512 | 0.271     | .786 | 0.149 |
| PAM_SUM_post  | 2.25 | 1.832 |           |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

According to Tables 17, 18 and 19, only PAM\_SUM\_pre vs. PAM\_SUM\_post in Group 2 stands out ( $d=0.710$ ), despite not being statistically significant ( $Z_{(w)}=1.794$ ,  $p>0.5$ ). Thus, regarding the global analysis of errors, the students who received the teacher's feedback seem to slightly outperform those who received GC feedback. Figure 32 compares the number of errors of the two drafts of *ABBY* with the two drafts of *PAM*.



Figure 32. Comparison of total number of errors between *ABBY* and *PAM*

As can be inferred from Figure 32, although the tendency and sort of errors (in the case of articles and prepositions) was different, both Group 1 and Group 2 reduced the number of errors, despite the fact that no statistical significance was reached. In turn, Group 3 also committed fewer errors but with a slight decrease of 3 errors. It is self-evident that feedback was more effective than no feedback, but no significant difference in effectiveness was noticed between the teacher's feedback and GC in the global analysis.

Despite the benefits of AWE programs discussed in Chapter 3, which point to the fact that they adapt to students' learning pace and allow them to be more autonomous (Cheng & Cheng, 2008), and contribute positively to the revision process by providing constructive feedback and reducing the teacher's pressure (Sundeen, 2015), in an attempt to answer RQ1, we may say that the results presented in this study do not confirm the hypothesis that GC will help students improve the use of the three targeted grammatical features significantly. Taking into account the analysis of the feedback received on the three grammatical features, GC was not as effective as expected, as it failed to highlight many of the errors. As a result, students could not even attempt to correct them.

The use of GC as a tool to improve the students' use of prepositions was not effective in our study, since the occurrence of errors before receiving feedback equaled the occurrence after receiving it. In addition, the amount of errors in *ABBY\_pre* (7) and *PAM\_post* (5) was almost the same, with an increase in *ABBY\_post* (10) and *PAM\_pre* (9). This could be explained in terms of U-shaped learning where learners follow a non-monotonic trajectory (good performance–bad performance–good performance) before language acquisition (Carlucci & Case, 2013). Similarly, Group 2 also revealed the U-shaped pattern but reduced errors to half.

As with prepositions, errors in articles committed by Group 1 were the same before and after receiving feedback and did not decrease from *ABBY\_pre* to *PAM\_post*. On the contrary, there was a considerable reduction of errors between *PAM\_pre* and *PAM\_post*. Finally, for tenses, Group 1 experienced the U-shaped pattern mentioned before and did not undergo any decrease in errors. Conversely, despite no statistical significance, Group 2 did experience a considerable decrease.

Overall, taking into account the total number of errors, only *PAM\_SUM\_pre* vs. *PAM\_SUM\_post* in Group 2 revealed a substantial reduction in errors despite not being statistically significant. Thus, contrary to previous research (e.g., Chen & Cheng, 2008; Chodorow, Gamon & Tetreault, 2010; Cotos, 2011; Grimes & Warschauer, 2006), it can be concluded that GC helped students reduce errors in the use of prepositions, articles and past simple-present/past perfect difference but results were not statistically significant so as to consider it an effective tool.

## 6.2 Results and Discussion Related to Hypothesis 2

In order to test H2, which predicted that Group 1 would at least equal Group 2 or even outperform Group 2 and Group 3, some between-groups and within-subjects tests were performed.

In the first place, four Kruskal-Wallis tests were run to compare each type of error of the first draft of *ABBY* and *PAM* between each group. The independent variables are each group and the dependent variables are the first draft of *ABBY* and *PAM* for each type of error and the *SUM\_pre* (sum of errors of the first draft). Table 20 shows the between-groups analysis:

Table 20. Kruskal-Wallis H test between groups for ABBY\_pre and PAM\_pre and effect size (partial eta square)

|               | Group | Mean | SD    | $X^2_2$ | p    | $\eta^2_p$ |
|---------------|-------|------|-------|---------|------|------------|
| ABBY_PREP_pre | 1     | 0.88 | 1.126 | 0.622   | .733 | 0.018      |
|               | 2     | 1.18 | 1.250 |         |      |            |
|               | 3     | 0.88 | 1.126 |         |      |            |
| ABBY_ART_pre  | 1     | 1.13 | 1.356 | 0.924   | .630 | 0.026      |
|               | 2     | 0.91 | 2.071 |         |      |            |
|               | 3     | 0.50 | 0.756 |         |      |            |
| ABBY_TEN_pre  | 1     | 0.25 | 0.463 | 2.300   | .317 | 0.105      |
|               | 2     | 1.09 | 1.758 |         |      |            |
|               | 3     | 0.38 | 0.518 |         |      |            |
| ABBY_SUM_pre  | 1     | 2.25 | 2.375 | 0.499   | .779 | 0.058      |
|               | 2     | 3.18 | 3.430 |         |      |            |
|               | 3     | 1.75 | 1.035 |         |      |            |
| PAM_PRE_pre   | 1     | 1.13 | 1.126 | 0.204   | .903 | 0.028      |
|               | 2     | 1.45 | 1.572 |         |      |            |
|               | 3     | 1.00 | 0.756 |         |      |            |
| PAM_ART_pre   | 1     | 1.50 | 1.604 | 4.098   | .129 | 0.140      |
|               | 2     | 1.36 | 1.286 |         |      |            |
|               | 3     | 0.38 | 0.744 |         |      |            |
| PAM_TEN_pre   | 1     | 0.13 | 0.354 | 4.082   | .130 | 0.121      |
|               | 2     | 0.36 | 0.674 |         |      |            |
|               | 3     | 0.63 | 0.518 |         |      |            |
| PAM_SUM_pre   | 1     | 2.75 | 2.435 | 1.037   | .595 | 0.054      |
|               | 2     | 3.18 | 2.359 |         |      |            |
|               | 3     | 2.00 | 1.512 |         |      |            |

Note: 1: “Computer”; 2: “Teacher”; 3: “Control”;  $X^2_2$ : chi square with two degrees of freedom;  $\eta^2_p$ : effect size (partial eta square)

As can be observed, after having conducted the Kruskal-Wallis tests, no statistically significant differences were noticed. In this case, the effect size was also considered to measure the results obtained beyond statistical significance. As there were three groups, the effect size was calculated through the partial eta-squared statistics<sup>13</sup> ( $\eta^2_p$ ), being around 0.01 a small effect size, around 0.06 a medium effect, and higher than 0.14 a large effect. Table 20 shows all the values are between a small and medium effect size with the exception of PAM\_ART\_pre with an eta-squared value of 0.14, which means that the errors in articles committed by the three groups are close to

13. Although some authors recommend the calculation of epsilon squared, it tends to be biased with small groups as it is the case of the present study.

belong to a different distribution, being 12 in Group 1, 4 in Group 2 and 3 in Group 3. As can be observed, although the incidence in Group 1 is higher than in the other groups, it cannot be considered statistically significant.

Similar to Table 20, the same Kruskal-Wallis tests for *ABBY\_post* and *PAM\_post* were conducted, in which the independent variables are each group and the dependent variables are *ABBY\_post* and *PAM\_post* measures for each sort of error and the *SUM\_post* (sum of errors of the second draft). Table 21 shows the results:

Table 21. Kruskal-Wallis H test between groups for *ABBY* post and *PAM* post and effect size (partial eta square)

|                       | Group | Mean | SD    | $X^2_2$ | p    | $\eta^2_p$ |
|-----------------------|-------|------|-------|---------|------|------------|
| <i>ABBY_PREP_post</i> | 1     | 1.38 | 1.302 | 0.842   | .656 | 0.041      |
|                       | 2     | 1.55 | 1.214 |         |      |            |
|                       | 3     | 1.00 | 0.926 |         |      |            |
| <i>ABBY_ART_post</i>  | 1     | 1.13 | 2.100 | 1.999   | .368 | 0.014      |
|                       | 2     | 1.00 | 0.632 |         |      |            |
|                       | 3     | 0.75 | 0.886 |         |      |            |
| <i>ABBY_TEN_post</i>  | 1     | 1.50 | 2.330 | 4.407   | .110 | 0.161      |
|                       | 2     | 0.18 | 0.405 |         |      |            |
|                       | 3     | 0.50 | 0.756 |         |      |            |
| <i>ABBY_SUM_post</i>  | 1     | 4.00 | 3.854 | 0.492   | .782 | 0.088      |
|                       | 2     | 2.73 | 1.489 |         |      |            |
|                       | 3     | 2.25 | 1.282 |         |      |            |
| <i>PAM_PRE_post</i>   | 1     | 0.63 | 1.061 | 2.228   | .328 | 0.069      |
|                       | 2     | 1.18 | 1.168 |         |      |            |
|                       | 3     | 1.38 | 1.302 |         |      |            |
| <i>PAM_ART_post</i>   | 1     | 1.13 | 1.642 | 1.684   | .431 | 0.109      |
|                       | 2     | 0.36 | 0.505 |         |      |            |
|                       | 3     | 0.50 | 0.535 |         |      |            |
| <i>PAM_TEN_post</i>   | 1     | 0.25 | 0.463 | 0.774   | .679 | 0.005      |
|                       | 2     | 0.27 | 0.905 |         |      |            |
|                       | 3     | 0.38 | 0.744 |         |      |            |

|  |   |      |       |       |      |       |
|--|---|------|-------|-------|------|-------|
| PAM_SUM_post   | 1 | 2.00 | 1.604 | 0.186 | .911 | 0.014 |
|  | 2 | 1.82 | 1.328 |       |      |       |
|  | 3 | 2.25 | 1.832 |       |      |       |
| Note: 1: “Computer”; 2: “Teacher”; 3: “Control”; $X^2_2$ : chi square with two degrees of freedom; $\eta^2_p$ : effect size (partial eta square) |   |      |       |       |      |       |

Table 21 shows that there are no statistically significant differences between the three groups, neither in *ABBY* nor in *PAM*, in any of the three targeted grammatical features and in the *SUM\_post*. The effect size was also taken into account to measure the results obtained beyond statistical significance and it can be claimed that all the values reveal either a small or medium effect size with the exception of *ABBY\_TEN\_post* with an eta-squared value of 0.16 (large effect size). That means that the distribution of tense errors between the three groups differs, being 12 in Group 1, 2 in Group 2 and 4 in Group 3. The incidence of errors in Group 1 is much higher than in Group 2 and Group 3.

In search of any further differences between drafts, a within-subjects analysis comparing *ABBY\_pre* vs *PAM\_pre* and *ABBY\_post* vs *PAM\_post* for the three targeted grammatical features in each of the three groups separately was conducted by means of a Wilcoxon test. In that way, instead of checking the progression between the first and the second draft of the same comic strip, a comparison between drafts of different comic strips was also performed to know whether the incidence of errors differed or not. Table 22 illustrates the results for Group 1:

Table 22. Wilcoxon test *ABBY* vs. *PAM* and Effect size for Group 1

|                      | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|----------------------|------|-------|-----------|------|-------|
| <i>ABBY_PREP_pre</i> | 0.88 | 1.126 | 0.355     | .722 | 0.222 |
| <i>PAM_PREP_pre</i>  | 1.13 | 1.126 |           |      |       |
| <i>ABBY_ART_pre</i>  | 1.13 | 1.356 | 0.680     | .496 | 0.249 |
| <i>PAM_ART_pre</i>   | 1.50 | 1.604 |           |      |       |
| <i>ABBY_TEN_pre</i>  | 0.25 | 0.463 | 0.577     | .564 | 0.291 |
| <i>PAM_TEN_pre</i>   | 0.13 | 0.354 |           |      |       |
| <i>ABBY_SUM_pre</i>  | 2.25 | 2.375 | 0.352     | .725 | 0.208 |
| <i>PAM_SUM_pre</i>   | 2.75 | 2.435 |           |      |       |

|                |      |       |       |      |       |
|----------------|------|-------|-------|------|-------|
| ABBY_PREP_post | 1.38 | 1.302 | 1.382 | .167 | 0.632 |
| PAM_PREP_post  | 0.63 | 1.061 |       |      |       |
| ABBY_ART_post  | 1.13 | 2.100 | 0.378 | .705 | 0.000 |
| PAM_ART_post   | 1.13 | 1.642 |       |      |       |
| ABBY_TEN_post  | 1.50 | 2.330 | 1.890 | .059 | 0.744 |
| PAM_TEN_post   | 0.25 | 0.463 |       |      |       |
| ABBY_SUM_post  | 4.00 | 3.854 | 1.581 | .114 | 0.678 |
| PAM_SUM_post   | 2.00 | 1.604 |       |      |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

Results show a lack of statistical significance between *ABBY* and *PAM* both for the first draft and for the second one, for the three grammatical features under analysis and the sum of errors of *ABBY* and *PAM* for the first draft and for the second one  $Z(W)=0.352$ ,  $p>0.5$ ). The effect size shows low values for most of the comparisons. Nevertheless, although not statistically significant, the fact that the comparison between *ABBY\_PREP\_post* and *PAM\_PREP\_post* ( $d=0.632$ ), *ABBY\_SUM\_post* and *PAM\_SUM\_post* ( $d=0.678$ ) and, especially *ABBY\_TEN\_post* and *PAM\_TEN\_post* ( $d=0.744$ ), are close to statistical significance ( $p=.059$ ) with a considerable effect size, could be kept in mind for further research on the topic (the number of errors of the comparisons just mentioned was 11 and 5, 32 and 16, 12 and 2, respectively). As can be observed, the incidence of errors in those categories decreased by at least half, which makes us think GC may have had a positive effect on new writings.

As for Group 2, the same analysis was conducted, and Table 23 shows the results:

Table 23. Wilcoxon test *ABBY* vs. *PAM* and Effect size for Group 2

|               | Mean | SD    | $Z_{(w)}$ | $p$   | d     |
|---------------|------|-------|-----------|-------|-------|
| ABBY_PREP_pre | 1.18 | 1.250 | 0.598     | .550  | 0.190 |
| PAM_PREP_pre  | 1.45 | 1.572 |           |       |       |
| ABBY_ART_pre  | 0.91 | 2.071 | 1.194     | .233  | 0.261 |
| PAM_ART_pre   | 1.36 | 1.286 |           |       |       |
| ABBY_TEN_pre  | 1.09 | 1.758 | 1.035     | .301  | 0.548 |
| PAM_TEN_pre   | 0.36 | 0.674 |           |       |       |
| ABBY_SUM_pre  | 3.18 | 3.430 | 0.000     | 1.000 | 0.000 |
| PAM_SUM_pre   | 3.18 | 2.359 |           |       |       |

|                |      |       |       |             |       |
|----------------|------|-------|-------|-------------|-------|
| ABBY_PREP_post | 1.55 | 1.214 | 0.686 | .493        | 0.311 |
| PAM_PREP_post  | 1.18 | 1.168 |       |             |       |
| ABBY_ART_post  | 1.00 | 0.632 | 2.111 | <b>.035</b> | 1.119 |
| PAM_ART_post   | 0.36 | 0.505 |       |             |       |
| ABBY_TEN_post  | 0.18 | 0.405 | 0.447 | .655        | 0.128 |
| PAM_TEN_post   | 0.27 | 0.905 |       |             |       |
| ABBY_SUM_post  | 2.73 | 1.489 | 1.611 | .107        | 0.645 |
| PAM_SUM_post   | 1.82 | 1.328 |       |             |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size. In bold, statistically significant differences

The results demonstrate the existence of statistically significant differences between ABBY\_ART\_post vs. PAM\_ART\_post ( $Z_{(w)}=2.111$ ;  $p=.035$ ;  $d=1.119$ ), where the effect size is large, which means the incidence of article errors is lower in PAM\_post in relation to ABBY\_post. This error decrease may be due to the treatment effect that has taken place from ABBY\_post to PAM\_post. There are no statistically significant differences in the other contrasts. The effect size was low in general; however, although they are not statistically significant, the values comparing ABBY\_SUM\_post vs. PAM\_SUM\_post ( $d=0.645$ ) and ABBY\_TEN\_pre vs. PAM\_TEN\_pre ( $d=0.548$ ) can be highlighted.

Finally, Table 24 shows the results of the same analysis for Group 3:

Table 24. Wilcoxon test *ABBY* vs. *PAM* and Effect size for Group 3

|                | Mean | SD    | $Z_{(w)}$ | $p$  | d     |
|----------------|------|-------|-----------|------|-------|
| ABBY_PREP_pre  | 0.88 | 1.126 | 0.276     | .783 | 0.125 |
| PAM_PREP_pre   | 1.00 | 0.756 |           |      |       |
| ABBY_ART_pre   | 0.50 | 0.756 | 0.378     | .705 | 0.160 |
| PAM_ART_pre    | 0.38 | 0.744 |           |      |       |
| ABBY_TEN_pre   | 0.38 | 0.518 | 0.816     | .414 | 0.483 |
| PAM_TEN_pre    | 0.63 | 0.518 |           |      |       |
| ABBY_SUM_pre   | 1.75 | 1.035 | 0.108     | .914 | 0.193 |
| PAM_SUM_pre    | 2.00 | 1.512 |           |      |       |
| ABBY_PREP_post | 1.00 | 0.926 | 0.966     | .334 | 0.336 |
| PAM_PREP_post  | 1.38 | 1.302 |           |      |       |
| ABBY_ART_post  | 0.75 | 0.886 | 0.816     | .414 | 0.342 |



|               |      |       |       |       |       |
|---------------|------|-------|-------|-------|-------|
| PAM_ART_post  | 0.50 | 0.535 |       |       |       |
| ABBY_TEN_post | 0.50 | 0.756 | 0.447 | .655  | 0.160 |
| PAM_TEN_post  | 0.38 | 0.744 |       |       |       |
| ABBY_SUM_post | 2.25 | 1.282 | 0.000 | 1.000 | 0.000 |
| PAM_SUM_post  | 2.25 | 1.832 |       |       |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size

As illustrated in Table 24, there were no statistically significant differences between *ABBY* and *PAM* neither in the first drafts nor in the second drafts for any of the three targeted grammatical features. Moreover, the effect size was low in all cases. Bearing in mind that Group 3 did not receive any feedback on the three grammatical features, the results obtained are in agreement with the initial expectations.

According to previous studies which show beneficial effects on the use of AWE (e.g., Attali, 2004; Cheng & Cheng, 2008; Chodorow et al., 2010; Coniam, 2009; Dikli, 2006; El Ebyary & Windeatt, 2010; Elliot & Mikulas, 2004; Hutchison, 2007; Li, Link, & Hegelheimer, 2015; Li et al., 2015; Nadasdi & Sinclair, 2007; Shi, 2012; Yang, 2004; Yang & Dai, 2015; Wang, Shang & Briody, 2013; Wei, 2015; Zhou, 2013) and the teacher's feedback (Chandler, 2003; Morch et al., 2017; O'Mahony et al., 2013; Wang & Liu, 2012; Warschauer & Ware, 2006), and despite the fact that both GC and the teacher's feedback have been effective to reduce the incidence of errors as shown in Figure 32, the analyses conducted to answer RQ2 partially confirm previous research. They reveal the lack of statistically significant differences between groups but close to statistical significance in some categories in the within-subjects analysis.

On the one hand, in the between-groups analysis, many errors were left unmodified because they were not marked by GC. Therefore, students did not notice them and the revision process was unsuccessful. Based on the opinion of some researchers (e.g., Anson, 2006; Condon, 2013; Herrington & Moran, 2001; Neal, 2011), computer programs fail to deal with aspects such as organisation, content, and meaning. They "process [words, sentences and paragraphs] as meaningless 'bits' or tiny fragments of the mosaic of meaning" (Ericsson & Haswell, 2006, p. 36). In fact, some of the errors that were not marked were grammatically correct (e.g., the difference between *in the end* and *at the end* as explained in Section 6.1.1, or the use of present/past perfect instead of simple past or vice versa as explained in Section

6.1.3) but not accurate in that context. This fact may explain our findings related to GC, in the sense that some errors were not marked because they were not recognised, and some others because they were grammatically correct but not in that context.

On the other hand, the within-subjects analysis which compares ABBY\_pre vs. PAM\_pre and ABBY\_post vs. PAM\_post seems to be more revealing and beneficial. The incidence of preposition and tense errors was reduced between ABBY\_post and PAM\_post as well as the SUM\_post. Students did not show any significant improvement in articles either in the first or second drafts but they did in tenses and prepositions in the second drafts. Therefore, for these two grammatical items, we may suggest that GC was more helpful over time than immediate revisions or second drafts. Students had written three compositions from ABBY\_post to PAM\_post (ABBY\_post, PAM\_pre and PAM\_post), two of which were revisions. During that time, they may also have benefited from instruction and input in their foreign language classes, which may have had an impact on their writings.

In line with previous studies, the teacher's feedback has proved to be beneficial both in revisions (e.g., Ashwell, 2000; Ferris, 2006; Ferris & Roberts, 2001; van Beuningen, Jong & Kuiken, 2008) and in new texts (e.g., Bitchener, 2008, 2009; Bitchener & Knoch, 2008, 2010a, 2010b; Ellis, Sheen, Takashima & Murakami, 2008; Sheen, 2007; Sheen, Wright & Moldawa, 2009). In our data, the teacher's feedback was superior in detecting errors than computer-generated feedback, in line with Otsu (2005) and Dikli and Bley (2014) and only 7 out of 119 were errors inaccurately modified. These cases, as Guo (2015) suggests, do not mean that the feedback provided has not been effective, but they reveal the need for more explicit feedback to reinforce learning. The feedback provided, in this case the colour code, may have not been sufficient so as to modify the error (e.g., Muñoz, 2011; Thouësny, 2011; Vyatkina, 2010). It is likely that by means of a complement such as a code or a metalinguistic explanation, those errors could have been solved.

Contrary to GC, the teacher's feedback has had a significant effect on articles in the comparison between ABBY\_post and PAM\_post, and it is close to significance in SUM\_post and TEN\_pre. Results illustrate different positive effects for GC and the teacher's feedback despite providing similar corrections.

Taking into account the fact that the incidence of errors in the between-groups analysis was only close to statistical significance in PAM\_ART\_pre (=0.14), and in ABBY\_TEN\_post (=0.16), results suggest that Group 1 and Group 2 showed positive effects and outperformed Group 3, which showed no significant differences in any of the drafts. With regard to the within-subjects analysis and taking into account the categories ABBY\_PREP\_post vs. PAM\_PREP\_post, ABBY\_SUM\_post vs. PAM\_SUM\_post and, especially ABBY\_TEN\_post vs. PAM\_TEN\_post in Group 1 are close to statistical significance, GC helped to reduce errors in those categories. As for Group 2, bearing in mind that the comparison between ABBY\_ART\_post vs. PAM\_ART\_post was statistically significant and the comparisons between ABBY\_SUM\_post vs. PAM\_SUM\_post ( $d=0.645$ ) and ABBY\_TEN\_pre vs. PAM\_TEN\_pre were close to being significant, we can claim that the teacher's feedback also helped to reduce the incidence of errors, especially after having written the first draft. Therefore, both GC and the teacher's feedback revealed to be equally useful in helping to reduce the incidence of errors in different categories.

In sum, taking into consideration that GC did not mark many errors, both grammatically correct but not suitable for that context and grammatically incorrect, relying on it as the only source of feedback seems too risky. The complexity of CF and the factors that influence its effectiveness can be the consequence of mixed results (Nassaji, 2017). Therefore, using computer-generated feedback as a supplement instead of as a substitute for teachers can be more conducive to learning than GC alone (Burstein, Chodorow & Leacock, 2003; Burstein & Marcu, 2003; Oladejo, 2005; Shermis & Burstein, 2003; Ware, 2005, 2011; Warschauer & Ware, 2006).

### **6.3 Results and Discussion Related to Hypothesis 3**

H3 predicted that students' gains after receiving WCF (either computer- or teacher-generated) would be maintained in the long term. A Wilcoxon test was conducted to test this hypothesis. The variables TAILOR\_pre (the sum of the errors made in the four drafts) and TAILOR\_post (the sum of the errors made in the tailor-made test six weeks after the last draft) per group and

per type of error as well as the sum of all of them were compared. Table 25 illustrates the results for Group 1 (computer-generated feedback).

Table 25. Wilcoxon test for TAILOR (Pre vs. Post) and Effect size for Group 1

|                  | Mean | SD    | $Z_{(w)}$ | $p$         | d     |
|------------------|------|-------|-----------|-------------|-------|
| TAILOR_prep_pre  | 3.50 | 2.268 | 1.604     | .109        | 0.827 |
| TAILOR_prep_post | 2.00 | 1.195 |           |             |       |
| TAILOR_art_pre   | 3.75 | 3.412 | 1.633     | .102        | 0.149 |
| TAILOR_art_post  | 3.25 | 3.105 |           |             |       |
| TAILOR_ten_pre   | 1.50 | 1.690 | 1.342     | .180        | 0.258 |
| TAILOR_ten_post  | 1.13 | 1.126 |           |             |       |
| TAILOR_sum_pre   | 8.75 | 5.922 | 2.041     | <b>.041</b> | 0.455 |
| TAILOR_sum_post  | 6.38 | 4.373 |           |             |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size. In bold, statistically significant differences

As can be observed, Table 25 shows statistically significant differences between TAILOR\_sum\_pre and TAILOR\_sum\_post ( $Z_{(w)}=2.041$ ,  $p=.041$ ). This value represents a significant reduction of errors in the TAILOR\_sum\_post in contrast with the TAILOR\_sum\_pre and a medium effect size. For the other comparisons, the effect size was small except for TAILOR\_prep\_pre and TAILOR\_prep\_post, which is really large ( $d=0.827$ ), despite not reaching statistical significance. The analysis suggests, then, that GC was effective in reducing the overall incidence of errors.

Table 26 illustrates the results of the same analysis for Group 2 (teacher's feedback).

Table 26. Wilcoxon test for TAILOR (Pre vs. Post) and Effect size for Group 2

|                  | Mean | SD    | $Z_{(w)}$ | $p$         | d     |
|------------------|------|-------|-----------|-------------|-------|
| TAILOR_prep_pre  | 5.18 | 2.892 | 2.699     | <b>.007</b> | 0.629 |
| TAILOR_prep_post | 3.55 | 2.252 |           |             |       |
| TAILOR_art_pre   | 2.91 | 1.814 | 1.633     | .102        | 0.250 |
| TAILOR_art_post  | 2.45 | 1.864 |           |             |       |
| TAILOR_ten_pre   | 1.45 | 1.968 | 2.264     | <b>.024</b> | 0.496 |
| TAILOR_ten_post  | 0.64 | 1.206 |           |             |       |

|                 |      |       |       |             |       |
|-----------------|------|-------|-------|-------------|-------|
| TAILOR_sum_pre  | 9.55 | 3.142 | 2.816 | <b>.005</b> | 0.904 |
| TAILOR_sum_post | 6.64 | 3.295 |       |             |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size. In bold, statistically significant differences

Statistically significant differences between TAILOR\_prep\_pre and TAILOR\_prep\_post ( $Z_{(w)}=2.699$ ,  $p=.007$ ); TAILOR\_ten\_pre and TAILOR\_ten\_post ( $Z_{(w)}=2.264$ ,  $p=.024$ ) and TAILOR\_sum\_pre and TAILOR\_sum\_post ( $Z_{(w)}=2.816$ ,  $p=.005$ ) can be observed. The number of errors detected in these three significant values was lower in the tailor-made test. In addition, the effect size was between medium and large. Results indicate that the feedback received by the teacher is an effective tool to reduce the number of errors in the long run, especially in prepositions and in the targeted tense (simple past-present/past perfect).

Finally, Table 27 presents the results of the same test for Group 3 (control group).

Table 27. Wilcoxon test for TAILOR (Pre vs Post) and Effect size for Group 3

|                  | Mean | SD    | $Z_{(w)}$ | $p$         | d     |
|------------------|------|-------|-----------|-------------|-------|
| TAILOR_prep_pre  | 4.88 | 4.486 | 1.841     | .066        | 0.262 |
| TAILOR_prep_post | 3.75 | 4.132 |           |             |       |
| TAILOR_art_pre   | 2.38 | 1.768 | 1.890     | .059        | 0.408 |
| TAILOR_art_post  | 1.75 | 1.282 |           |             |       |
| TAILOR_ten_pre   | 1.25 | 1.282 | 1.000     | .317        | 0.099 |
| TAILOR_ten_post  | 1.13 | 1.126 |           |             |       |
| TAILOR_sum_pre   | 9.63 | 4.406 | 2.388     | <b>.017</b> | 0.482 |
| TAILOR_sum_post  | 7.50 | 4.440 |           |             |       |

Note: SD: Standard Deviation;  $Z_{(w)}$ : Wilcoxon z-score; d: Cohen's d effect size. In bold, statistically significant differences

With regard to Group 3, significant differences were observed in TAILOR\_sum\_pre vs TAILOR\_sum\_post ( $Z_{(w)}=2.388$ ,  $p=.017$ ), where an important global reduction of errors took place. The other comparisons did not represent any relevant differences and the effect size was rather small. Despite not having received any feedback on the three grammatical features, these students kept on receiving instruction and exposure to the foreign language during the academic year, which may explain why students in Group 3 improved over time.

Contrary to the nativist view that CF plays little or no role in acquisition (e.g., Carroll, 1995; Cook, 1991; Krashen, 1982, 1985), and especially to Truscott's (1996) statement that "grammar correction has no place in writing courses and should be abandoned" (p. 328), this claim seems to be contradicted by the results of the present study. The language acquisition process is bound to the availability of language input (Gass & Selinker, 1994; Grady et al., 2011; Pica et al., 1987; VanPatten & Benati, 2010; VanPatten & Williams, 2007). Although SLA theories attach different importance to the role of input in SLA, they all acknowledge its relevance (Ellis, 2008). Part of this input is in the form of feedback, which is indispensable for SLA (Ellis, 2008; White, 1987) and it has been proven in our research, it plays a role in SLA, regardless of the type of CF provided.

Taking into account the fact that, unlike oral feedback, WCF is likely to be more noticeable (Bitchener, 2012; Ellis, 2003; Polio, 2012) and that writing is expected to further SLA (Manchón, 2011; Williams, 2012), WCF may have a positive effect on the students' written performance over time. Results of the analysis of Group 2 (statistically significant differences between TAILOR\_prep\_pre and TAILOR\_prep\_post; TAILOR\_ten\_pre and TAILOR\_ten\_post and TAILOR\_sum\_pre and TAILOR\_sum\_post) are in line with studies which suggest that WCF has a role in helping students improve the accuracy of their writing, since L2 learners cannot develop native-like accuracy on mere exposure to positive evidence (e.g., Bitchener, 2008; Bitchener & Knoch, 2008, 2010a, 2010b; Chandler, 2009; Ellis, 2007; Ellis, Sheen, Murakami & Takashima, 2008; Farrokhi & Sattapour, 2011; Ferris, 2006; Ferris & Helt, 2000; Ferris & Roberts, 2001; Ferris et al., 2000; Hanaoka & Izumi, 2012; Hyland, 2011; Russell & Spada, 2006; Sheen, 2006; Shintani, Ellis & Suzuki, 2014; van Beuningen, De Jong & Kuiken, 2012).

As for Group 1, previous studies on AWE (e.g., Chodorow, Gamon & Tetreault, 2010; Lee, 2015; Li, Link & Hegelheimer, 2015; Moon & Pae, 2011) reported beneficial effects on student writing. However, Group 2 performed better than Group 1 in the long term although it only exhibited significant improvement in TAILOR\_sum\_pre vs. TAILOR\_sum\_post and a large effect size ( $d=0.827$ ) in TAILOR\_prep\_pre vs. TAILOR\_prep\_post but not statistically significant. In addition, contrary to Link et al.'s (2020) study, long-term retention in Group 2 (only-teacher feedback) was higher than in Group 1 (computer feedback).

Both experimental groups outperformed the control group in the three targeted grammatical features (i.e., articles, tenses, and prepositions), as they did not exhibit any significant improvement. Significant differences were only observed in TAILOR\_sum\_pre vs TAILOR\_sum\_post, which means that the sum of the three categories did make a difference. The reduction of errors cannot be associated with CF, since no feedback on the three categories was provided. Thus, it can be due to mere instruction, with a possible effect on students' development.

Regarding the type of CF provided by the teacher and the computer, while some researchers have claimed that direct feedback is more effective because it overtly suggests how to correct the error (e.g., Bitchener, 2008; Bitchener, Young & Cameron 2005; Ellis et al., 2008; Nassaji, 2015; Sheen, 2007), some others have stated the opposite as indirect feedback engages students in problem-solving (e.g., Ferris 2003, 2006).

When comparing both, van Beuningen et al.'s (2008) study reveals that "while short-term effects were found for both direct and indirect corrective feedback, only direct feedback proved to have a significant long-term effect" (p. 279). On the contrary, in accordance with other previous studies (e.g., Ashoori Tootkaboni & Khatib, 2014; Bitchener & Knoch, 2008; Chandler, 2003; Erel, 2007; Ferris 2003; Ferris, Chaney, Komura, Roberts & McKee, 2000; Ferris & Helt, 2000; Li, 2010), indirect CF has proven to be more effective in the long-term than direct CF as it "requires pupils to engage in guided learning and problem solving" (Bitchener & Knoch, 2008, p. 415). Similarly, Ferris and Hedgcock (2014) explained that underlining seems to have "a more positive effect on long-term student improvement in accuracy and editing skills" than direct feedback (p. 206).

In line with Ferris' (2010) claims, rehearsing and repeating the drafts where students notice, by means of GC and the teacher's feedback in the present study, the inaccuracies they have to modify, may have also helped students to retain the noticed items in long-term memory. Thus, based on the quantitative results obtained in RQ3, it can be concluded that students' gains after receiving WCF (either computer- or teacher-generated) are maintained in the long term, especially those in Group 2, which outperformed Group 1.

## 6.4 Results and Discussion Related to Hypothesis 4

H4 states that students' attitude on the kind of feedback received (teacher or computer) will influence the results irrespective of the feedback provided. Indeed, it is a challenging suggestion to answer, since it is complex to ascertain how and to what extent the students' perceptions are influenced by the feedback received and how it can affect the results. The factors influencing the students' attitudes can be numerous, for example, the teacher's type of correction, the software characteristics, their technological abilities and social skills, learning preferences (teacher or computer interaction, on-site or distance), level of proficiency, teacher/computer dependence to learn, etc. What is evident is that the students' attitudes affect and mediate the process of WCF (Bitchener, 2012; Sheen, 2007; Storch & Wigglesworth, 2010). In fact, it has also been proven that a positive attitude and cognitive engagement may be related to positive academic results (e.g., Fredricks, 2013; Skinner & Pitzer, 2012).

Considering the participants' background, the fact that their mean age was nearly 40 years old, and that all of them had studied English for more than 6 years, they must have received, to a greater or lesser degree, some kind of feedback throughout their language training and university education and had some experience with computers as well.

From the questionnaire distributed at the end of the study (see Appendix 2), we become aware of the students' need to improve their writing abilities, since most of them consider they have an average level of 3 in a scale from 1 (excellent) to 5 (poor).

Students' concerns in EFL contexts are usually about the productive skills (i.e., speaking and writing). Putting their thoughts in a meaningful and proper form, being fluent, creative and comprehensible altogether is a challenging task for students, possibly resulting in anxiety. They tend to express themselves in their FL as they do in their L1 and, for some reasons (e.g., lack of vocabulary, coherence, or literal translation, among others) feel the need to improve their writing.

The participants' will to learn English goes hand in hand with their attitude towards the feedback received, an issue that may have influenced the results of the study. In line with Denton et al. (2008), who found that learners preferred computer-assisted feedback, when participants in



Group 1 were asked what tools they prefer to practise the writing skill, only one person chose the handwritten option, whereas the other participants opted either for the *computer* (4 participants) or for *both* (3 participants). The students' preferences for writing their work by hand or using a computer may depend on several factors. According to Wolfe and Manalo (2004), "examinees who speak a language that is not based on a Roman/Cyrillic alphabet, [...] have poorer English language skills [and exhibit] lower levels of computer experience and higher levels of computer anxiety are less likely to choose the word processor as their composition medium" (pp. 45-46). In this case, the participants' mother tongues are based on the Roman alphabet; however, taking into account the students' background, some other factors such as job and studies may have also influenced their choices.

In Group 1 most participants were in their 40s (75%), that is to say, they were not digital natives. Thus, the *computer* option was not expected to be chosen by a majority (if we add the option *both* to the other options, 4 subjects opted for *paper and pencil* and 7 for *computer*). However, if we consider their academic training, 75% of the participants held a Science Degree such as Technical Engineering or Computer Science, which may explain why they opted for the *computer* option and *both*. Supposedly, those students may not be too focused on linguistic features. Specifically, the student who chose the *paper and pencil* option had studied Labour Relations and was 42 at the moment of the study. Although these reflections are not conclusive since there was no personal interview between the teacher and the participants to confirm these rather subjective explanations, what is clear is that "the discipline and level of study might be a factor in the choice of media used for reading and writing" (Vincent, 2016, p. 100).

In addition, students felt positive about the use of computers in their learning process, as the following students' comments show (grammar errors have not been corrected in the students' comments):

*I had to use dictionaries, books of the subject or ask to some mates. Sometimes we did not have the best materials. Nowadays it is faster than before, you can find the solution quickly using Internet. We are lucky to have this option to improve ... a lot easier and less time consuming*  
*Very positive and it encourage you to write in English without fear*  
*Very useful*  
*It is needed*

According to Fredricks, Blumenfeld and Paris (2004), motivation through emotional and cognitive engagement is necessary for computer-generated feedback to be effective and facilitate language acquisition, the fact of having used a computer program to write essays should not have hindered the students' writing practice if their preferences are taken into account. Only the student who chose the *paper and pencil* option could have felt disappointed by being obliged to use a more innovative tool. However, the comments this student wrote in the questionnaire suggest that she had positive feelings about the use of computers, as she reported that *I want to learn good writing and this is a new way*.

In an attempt to go a step further, and based on the assumption that students feel the need to be corrected (e.g., Bang, 1999; Jang, 2003; Katayama, 2006; Schulz, 2001), question 13 in the questionnaire aimed at evaluating the influence of both types of feedback (GC and the teacher's) on the students' attitude and, consequently, on their performance. To do so, both types of feedback were analysed per group according to their individual questionnaires.

According to previous studies on the students' preference for correction (e.g., Al-Magid & Abdul, 2006; Katayama, 2007; Lee, 2008; Odalejo, 1993), students expect to be corrected mainly when errors impede communication. However, there are also other factors such as language proficiency or students' personal features which may influence how they prefer to be corrected (e.g., Lee, 2008). Sometimes the students' dissatisfaction during the learning process comes from a mismatch between teacher's and students' perceptions on correction, which may even lead to unsatisfactory learning outcomes (Schulz, 2001). For this reason, it is important to take into account what learners value when being corrected. Concerning the participants' preferences in Group 1 and in line with previous studies (e.g., Amrhein & Nassaji, 2010; Diab, 2005; Hamouda, 2011; Lee, 2004, 2005), when they were asked what, in their opinion, should be corrected if multiple mistakes are found in their writings, most (75%) answered that all errors should be marked (Option a). Only one student chose Option b (*Teacher should mark all major errors but not the minor ones*) and another one selected Options a, b and d (*Teacher should mark only the errors that interfere with communicating your ideas*). The majority of Option a suggests that students, in general, value being thoroughly corrected on, for example, organisation, grammar, content, punctuation, spelling and vocabulary.

Although it is evident the participants highly valued being corrected, they may not attribute the same relevance to all types of errors, something that may influence their attitude towards GC. For this reason, a more specific question on this issue was included, the results of which are illustrated in the figure below.

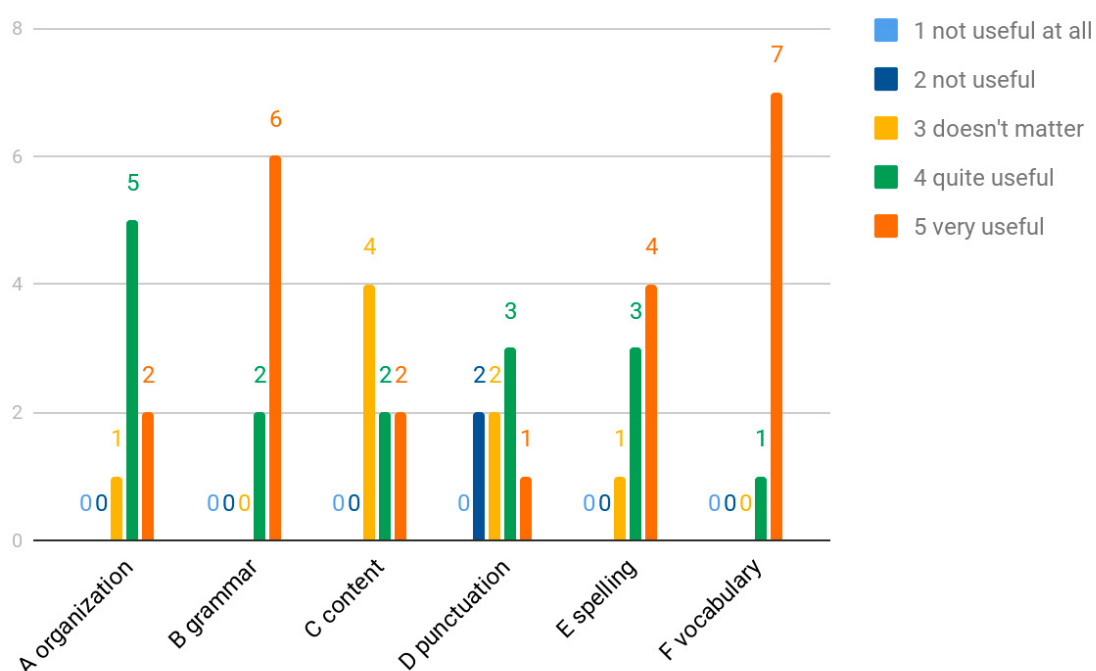


Figure 33. Most valued errors for correction (Group 1)

As observed, almost all aspects were given at least 3 points except for punctuation. Students highly value being corrected when they make grammar, spelling and vocabulary errors, giving less importance to content errors. Results suggest that although students value being corrected in general, they really appreciate being corrected on formal aspects. This may favour GC, as it only provides feedback on spelling, pairs filter (combination of words), incorrect sequences (the most frequent incorrect sequences in English), and problem words (words that many students of English use incorrectly). That is to say, it does not correct organisation, content or punctuation. Thus, taking into account that students wish to be thoroughly corrected, the fact that GC did not correct all their mistakes or even flagged correct utterances as incorrect may have had an effect on the students' attitudes. As previous studies have shown, students tend to trust auto-

mated feedback (e.g., Grimes & Warschauer, 2010; Li, Link, Ma, Yang & Hegelheimer, 2014; Rich, 2012) but the software and the teacher's approach also play a role (Chen & Cheng, 2008). In fact, the use of grammar checkers may increase the students' motivation and competence in grammar (Potter & Fuller, 2008).

In the present study, as reflected by the participants' comments (see below), they were conscious that GC has its limitations, but they saw the software as a useful tool to provide instant feedback on the aspects they value most (i.e., grammar, spelling and vocabulary). The teacher revised the remaining errors on content, organisation and punctuation.

*I think Grammar Checker is a good tool to practise the writing [...] doesn't correct completely the essay [...] the teacher has to check the essay again*  
*It shows you fast and easily yours majors errors but sometimes it can't grab your ideas and it mark them as an error*  
*You managed to correct yourself the minor issues in your texts, in an easy, friendly and fast way. The teacher can concentrate in improving other aspects of the writings, be it vocabulary, general structure, tone, style meeting the purpose, etc.*  
*... very advisable [...] you can check your errors instantly. It also brings benefits for the teacher, who has to check less mistakes in the compositions and saves a lot of time. [...] the students can check their grammar and spelling mistakes easily. However, in other skills such as vocabulary it is not as useful as the teacher correction*  
*... was a good method, useful and easy of using*  
*Really good. It helped me in a visual way with no so common mistakes like collocations, prepositions, spelling of tricky words and the like*  
*The use of this computerized-feedback is a quick way to receive the feedback but, probably, not so good as the human teacher.*

As some of their comments reflect, students are conscious that computer-generated feedback saves time and helps teachers to focus on higher-order errors (Denton et al., 2008). However, it is difficult to measure the extent to which learners trust AWE given the fact that its effectiveness is somehow questionable (Lai, 2010).

Computer literacy may also be influential and can be considerably determining when dealing with computer programs. Thus, although Matsumara and Hann's (2004) research revealed that both high- and low-anxiety students improved when using the computer, a complementary question to know their level of satisfaction with computer-generated feedback was included in the questionnaire. The participants were asked to rate how challenging the experience of using GC was. Figure 34 illustrates the results.

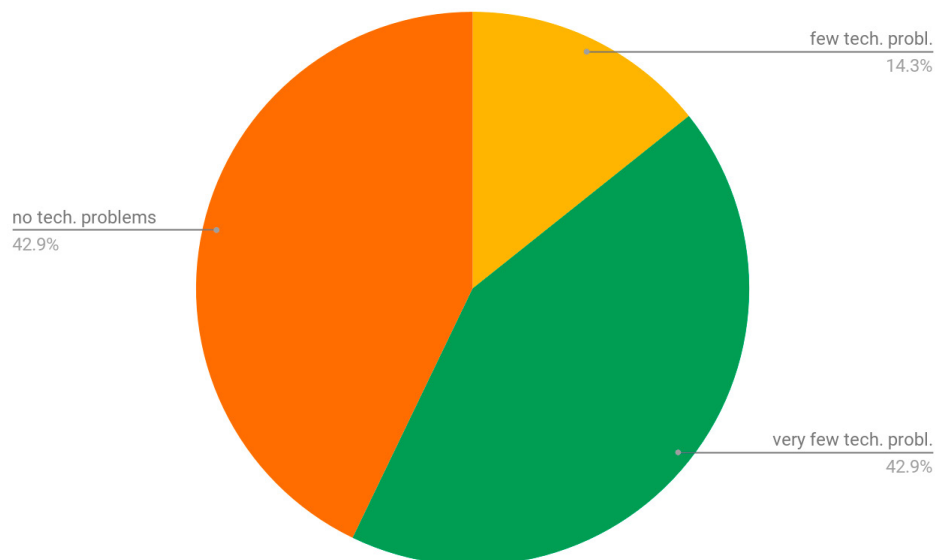


Figure 34. Technological challenge (Group 1)

As illustrated in Figure 34, although three of the participants did not answer this question, the results obtained were predominantly positive. Despite being their first experience with GC, only 14.3% (1 participant) had few technological problems. In other words, students could manage the computer program without much difficulty and described it as *very useful, positive, new way, streamlined, very positive, and easy of using*.

All in all, taking into account Figures 36, 37 and 38 and in line with Dikli and Bleyle (2014), despite being aware of the shortcomings and weaknesses of GC, students tend to trust computer-generated feedback. In addition, although it seems face-to-face feedback is still a priority (Guardado & Shi, 2007; Liu & Sadler, 2003; Schultz, 2000; Tuzi, 2004), computer-generated feedback is also seen as positive, corroborating previous research on this issue (e.g., El Ebyary & Windeatt, 2010; Elliot & Mikulas, 2004; Elola & Oskoz, 2016; Liu & Sadler, 2003; Lu & Bol, 2007; Rich, 2012; Yang, 2004; Yu & Yeh, 2003).

As for the teacher's feedback, 57.1% of the students in Group 2 reported that they preferred paper and pencil as the means to write their essays, as opposed to Group 1. That is to say, a minority opted for the options *both* (28.6%) or *computer* (14.3%). Aspects such as age or job may have played a role in the participants' choice of writing medium. However, age ranges are quite balanced among groups, so they cannot have influenced having chosen *Paper and pencil* as the

major option. Thus, contrary to Vershinskaya (2014), who stated “the age matters” (p. 4), in this study it seems we cannot support this argument.

Regarding their professions, if we compare Group 1 and Group 2, it is interesting to note that most participants in Group 2 were teachers, whereas most students in Group 1 held a Science Degree. It could be suggested that because of the fact of being teachers in their 40s (71,5%) they were more comfortable using paper and pencil. New technologies were not so highly used when they started working and as primary/secondary teachers, so they are conscious of the importance of handwriting to learn and retain knowledge (Mueller & Oppenheimer, 2014) as well as of its intimate emotions and sensorial qualities (Vincent, 2016).

In sum, the participants in both groups did not have any constraint to choose a specific tool to write their compositions. For this reason, the freedom of using any means for their writings should not have influenced the students’ writing practice negatively, since they were given more autonomy and self-confidence.

Turning to the students’ expectations about the teacher’s role, Group 2 is comparable to Group 1. Most students (71.5%) mainly agreed on the fact that the teacher should mark all errors except for two participants (28.5%), who preferred being corrected on major but not on minor errors.

In the present study, the teacher/researcher corrected all the students’ errors in the same way *Grammar Checker* did, both minor and major. However, unlike Group 1, Group 2 did receive feedback on all aspects. That is to say, not only grammar and spelling but also content, structure, punctuation, etc. In the case of Group 1, it was the teacher who provided that global feedback. For the different aspects to be corrected, Figure 35 illustrates what type of errors are most valued by students in Group 2.

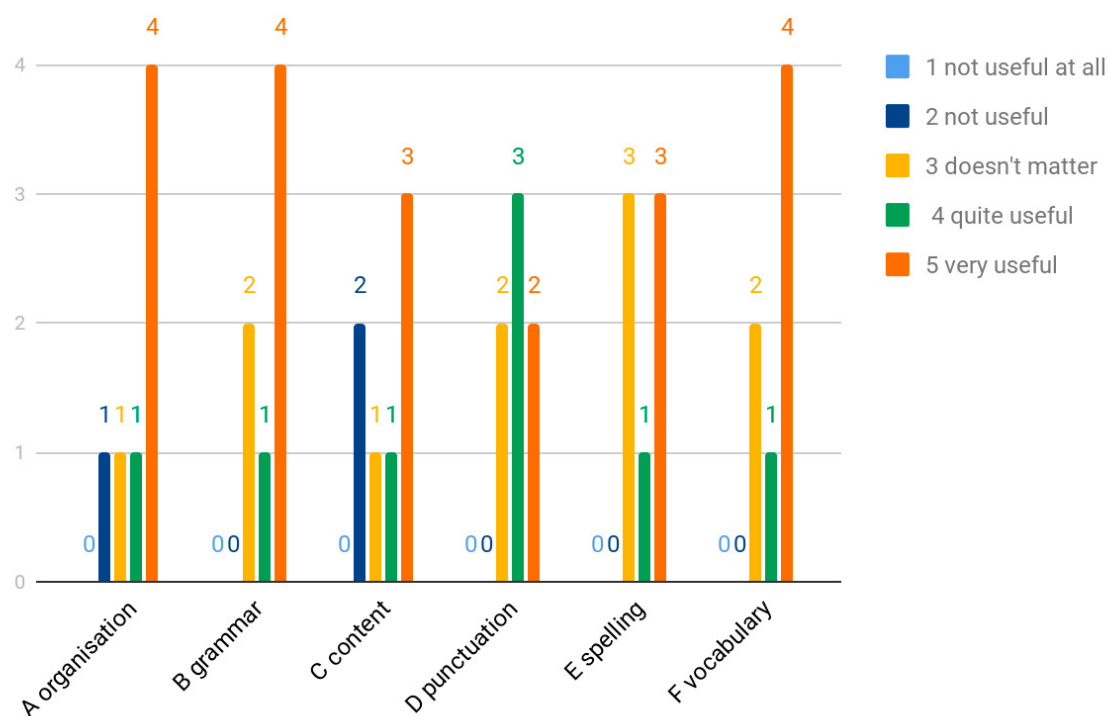


Figure 35. Most valued errors for correction (Group 2)

As can be observed, participants in Group 2 highly valued being corrected on aspects such as organisation, grammar, and vocabulary, but not so much on punctuation, which is still highly valued but less than the three previous ones. As for content, one student thinks it is not important and two students believe it is not even useful. Concerning spelling, the options *Doesn't matter* and *Very useful* were chosen by three participants, which means there is no consensus on its relevance. Hence, taking into account Figure 35, the teacher's role should not have had a negative influence on the students' attitudes when being corrected but a favourable one, since students were provided with feedback on the highly valued aspects (i.e., organisation, grammar, vocabulary and punctuation) and also on the less valued ones (i.e., spelling and content). Theoretically, disappointment or any other negative feeling that could influence the students' attitude towards the teacher's feedback should only appear, if so, in content, where two people chose the option 'not useful'.

Mixed opinions emerged in Group 3. As Figure 36 illustrates, half of the participants preferred using paper and pencil for their written assignments, whereas almost the other half opted for the computer and one participant chose both options.

Group 3 could also choose the means they preferred to write their essays and they mainly opted for *paper and pencil*, followed by *computer* and *both*. Probably, their age and jobs may have played a role in this selection, as in this group they were mainly in their 30's (50%) and their professions were quite mixed (lawyer, teacher, architect, etc.).

With regard to the participants' preferences for feedback provision, Group 3 was as demanding as Group 1 and Group 2. 67% of the students think all errors should be marked, whereas only 17% think only major ones should be corrected and another 17% only those errors that interfere with communicating content.

To get a further insight on the topic, the figure below illustrates the specific aspects students valued most to be corrected on.

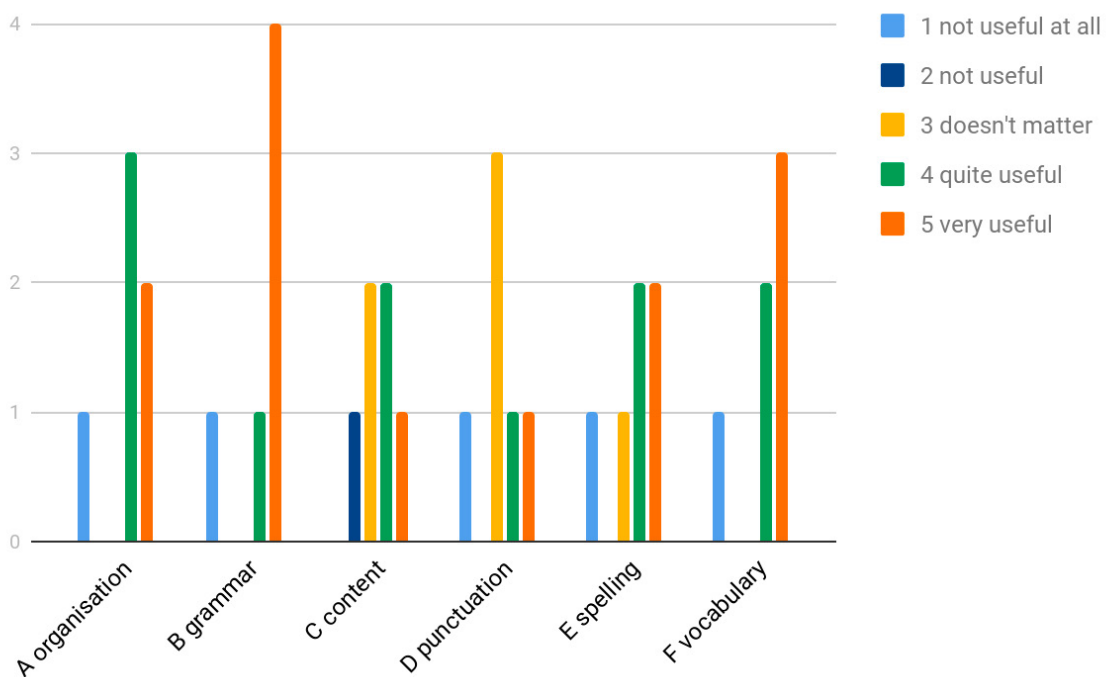


Figure 36. Most valued errors for correction (Group 3)

What students valued most was the correction of organisation, spelling, vocabulary, and, especially, grammar, as in previous studies on the topic (e.g., Ashwell, 2000; Chiang, 2004; Saito, 1994). Content obtained mixed results, as half of the participants considered it to be useful and the other half regarded it as unimportant and even useless. In comparison with Group 1 and



Group 2, it is surprising that punctuation was seen as insignificant. In addition, the option *not useful at all* was reflected for the first time in almost all the aspects but it was not chosen by the same person. Then, as far as Group 3 is concerned, the participants received feedback on all the aspects they valued most as well as on those they did not value so much. Therefore, in an attempt to answer RQ4, we may claim that students' preferences should not have negatively influenced their attitudes to the feedback received, since their desires were satisfied. Their errors on articles, prepositions and the past simple-present/past perfect difference were not proofread, something they were not aware of because they were not told. Otherwise, it could have somehow interfered in the study.

From the answers obtained after the analysis of the questionnaires, the fact that the three groups are comparable becomes evident. Most of them are Spanish students between their 30's and 50's (question 3), have had experience with other languages (question 5), have studied English for more than 5 years (question 6), think learning English is important (question 7), have an average writing level of 3 according to them (question 11), are studying the 2B2 level after having passed the 1B2 level, hold a degree (question 9) and highly value correction (option a in question 13). In addition, all the three groups benefited from the medium used to write their compositions. Group 1 preferred using the *computer* or *both* as the second option, Group 2 opted for *paper and pencil* and Group 3 was between the *computer* and *paper and pencil*. Thus, as far as questions 1 to 13 are concerned, there is no concern which could have influenced their attitudes towards feedback.

On the contrary, with regard to question 14, which queried about the importance given to the corrections on the different features, some differences emerged. As for organisation errors, all the students agreed that correction is quite relevant. Regarding grammar errors, all groups valued it with almost the maximum score, except two students in Group 2 who gave it an average score. In general, the correction of grammar is considered important.

As for content, the score in Group 1 and Group 3 was close, since half of the participants considered it to be irrelevant whereas the other half thought it was quite useful. In regards to punctuation, half of the participants in Group 1 thought it was relevant, whereas the other half did not value its correction. Group 2 appraised its correction, whereas Group 3 did not do it to

a large extent. Concerning spelling errors, Group 1 and Group 3 considered it a relevant aspect while Group 2 obtained similar answers for *Doesn't matter* and *Very useful*. Finally, vocabulary was highly valued in general, especially in Group 1 and Group 3.

In conclusion, taking into account the analysis of the participants' answers to the questionnaire, feedback does not seem to have influenced the students' performance in any of the four drafts negatively, in line with previous research that has proven the effectiveness and usefulness of computer-generated feedback (e.g., Chodorow, Gamon & Tetreault, 2010; Cotos, 2011; Grimes & Warschauer, 2006) by helping and motivating students to revise and improve their texts.

Undoubtedly, we live in the digital era where print and digital technology coexist but more research into the impact that both of them have on the students' learning and language acquisition should be undertaken. Likewise, the extent to which the students' attitude and involvement influence language acquisition and writing development is an under-researched area (Handley, Price & Millar, 2011).

## 6.5 Chapter Summary

Chapter 6 included the statistical analysis aimed at answering the 4 RQs proposed in Section 4.2. Taking into account the small sample size ( $n=27$ ), nonparametric tests were employed as they make no assumptions about the probability distributions of the variables being assessed. For the comparison within subjects, a Wilcoxon test was performed. When the comparison was between groups and more than two, as it is the case, the Kruskal-Wallis test was performed. Table 28 illustrates all the tests used in order to support or refute the four hypotheses suggested in the present study.

Table 28. Summary of main findings of the study

| Hypotheses  | Data Collection Procedure   | Data Collection Time | Data Analysis                                    | Hypothesis confirmed |
|---|---|----------------------|--|----------------------|
| H1: <i>Grammar Checker</i> will provide proper CF so as to improve students' accuracy on the use of prepositions, articles and past simple-present/past perfect difference.           | Wilcoxon signed-rank test + Cohen's d effect size   | During the study     | Quantitative: between groups                     | No                   |
| H2: Students using <i>Grammar Checker</i> will at least equal those students who receive the teacher's feedback and will outperform those receiving no feedback in terms of accuracy. | Kruskal-Wallis one-way analysis of variance + Effect size (partial eta square) or Wilcoxon signed-rank test + Cohen's d effect size | During the study     | Quantitative: between groups and within subjects | Partially            |
| H3: The students' gains will be maintained in the long-term irrespective of the feedback provided.  | Wilcoxon signed-rank test + Cohen's d effect size   | Tailor-made test     | Quantitative: between groups                     | Yes                  |
| H4: The students' attitude on the kind of feedback received (teacher and computer) will influence the results.  | Questionnaire   | After the study      | Qualitative                                      | No                   |



## Chapter 7. Conclusions

In this final chapter, a summary of the main findings related to the field of computer-generated feedback are provided in Section 7.1. In turn, the limitations of the study are presented in Section 7.2. The contribution of the study and further research that can be conducted in order to widen the body of research on computer-generated feedback are discussed in Section 7.3. Finally, section 7.4 describes some pedagogical implications which derive from the present study.

### 7.1 Summary of the Main Findings

As explained in Chapter 4, the idea of conducting the present research emerged as a result of the researcher's experience as an English teacher at an Official School of Languages. In view of the difficulties students had when writing, the teacher's workload during examination periods and, in general, the students' need to practise the production skills, the search for a tool that assisted students in this sense became significant. In addition, taking into account the fact that research into the implementation of computer-generated feedback showed inconclusive results, since some studies reported favourable results (e.g., Coniam, 2009; Hutchison, 2007) while some others showed the opposite (e.g., Lai, 2010; Lee et al., 2009; Tuzi, 2004), the aim of the present dissertation was to contribute to a better understanding of computer-generated feedback and to shed light on the effectiveness of AWE tools to improve the students' written work by means of the software GC in comparison to teacher's feedback. In order to do so, four RQs were proposed and different analyses were conducted.

RQ1 intended to examine GC effectiveness in providing proper CF so as to improve students' accuracy on the use of prepositions, articles and past simple-present/past perfect difference. Taking into account previous research on AWE programs (e.g., Attali, 2004; Burston, 2008; Chen & Cheng 2006; Coniam, 2009; Hutchison, 2007; Nadasdi & Sinclair, 2007; Potter & Fuller, 2008; Shi, 2012; Yang, 2004; Yang & Dai, 2015; Wei, 2015; Zhou, 2013) and the benefits reported when students' L1 and FL share similarities (Cai, 2001; Collins, 2004; Ekiert, 2004; Giacalone, 2002; Hawkins et al., 2006; Ionin & Montrul, 2010; Ionin, Zubizarreta &

Maldonado, 2008; Master, 1990, 1997; Park, 1996; Romaine, 2003; Snape et al., 2013), H1 predicted beneficial results in provision of individualised feedback on specific aspects such as grammar by means of GC.

Focusing on the three targeted grammatical features analysed, prepositions, in accordance with Chodorow, Tetreault & Han (2007) and Díez-Bedmar & Casas-Pedrosa (2011), were highly used and, therefore, where most errors were committed (Ballesteros et al., 2005; Izumi et al., 2004; Nicholls, 2003; Tetreault & Chodorow, 2008). As for articles, in accordance to some authors' claims, (e.g., García-Mayo, 2009; Ionin & Montrul, 2010; Ionin et al., 2008; Snape et al., 2009), the main article errors committed in this study are article omission in obligatory contexts, substitution of one article for another and oversupply. The “the-flooding” effect (Huebner, 1983; Master, 1997) mentioned in Section 1.5.1 where Spanish speakers tend to overgeneralise the definite article *the* has also been noticeable when using family words.

Finally, in relation to tenses, errors referring to this category amount to half the errors in articles and one third of preposition errors. It may be due to positive L1 transfer or a less complex system in comparison to prepositions.

Taking into account the statistical analysis, as expected, Group 3 did not show any improvement in any of the three targeted grammatical features in the short term (prepositions:  $n=34$ ,  $n=33$ ; articles:  $n=17$ ; tenses:  $n=10$ ). Surprisingly, the same happened with Group 1. GC did not show statistically significant results in the comparison of ABBY\_pre vs ABBY\_post and PAM\_pre vs PAM\_post in either of the three grammatical elements. Only the value shown in ABBY\_TEN\_pre and ABBY\_TEN\_post ( $d = .744$ ) could be stressed as it is close to being statistically significant. This grammatical feature experienced a sudden increase but later dropped, and ABBY\_TEN\_pre equalled PAM\_TEN\_post.

In turn, considering the fact that the feedback that Group 2 received was similar to the one provided by GC, the results of the comparison of ABBY\_pre vs ABBY\_post and PAM\_pre vs PAM\_post in the three grammatical features makes us think something happened with GC. Although the results were only statistically significant in PAM\_art\_pre and PAM\_art\_post ( $p=.026$ ;  $d=1.024$ ) and close to statistical significance in ABBY\_TEN\_pre vs. ABBY\_TEN\_post

( $d=.713$ ), the incidence of errors was substantially reduced (prepositions:  $n=59$ ,  $n=28$ ; articles:  $n=40$ ,  $n=27$ ; tenses:  $n=21$ ,  $n=5$ ).

Students tried to correct the wrong use of prepositions but left the missing and unnecessary prepositions uncorrected. As for articles, the wrong use experienced a constant decrease, whereas the unnecessary use and the missing articles increased from *ABBY\_pre* to *PAM\_pre* and underwent an important decrease in *PAM\_post* almost equalling the incidence committed in *ABBY\_pre*. This fact may imply that, depending on the kind of error, different types of feedback can be applicable.

The global analysis of errors depicts a similar picture where only *PAM\_SUM\_pre* vs. *PAM\_SUM\_post* in Group 2 revealed a substantial reduction in errors despite not being statistically significant. Bearing in mind the incidence of errors before and after feedback (Group 1:  $n=84$ ,  $n=61$ ; Group 2:  $n=78$ ,  $n=57$ ; Group 3:  $n=57$ ,  $n=52$ ), errors in Group 1 and Group 2 reduced in one quarter. That means GC helped students reduce errors but results were not statistically significant in the present study.

Therefore, taking into account that the feedback provided in Group 1 and Group 2 was similar, the advantage of teacher's feedback over computer-generated feedback must be the result of the teacher marking all errors, whereas GC only highlights some of them, thus hindering correction and, possibly, acquisition. After a deep analysis of each error, an important lack of feedback provision was revealed. This may be partly due to the fact that errors were not related to the grammatical structure but contextual, and partly because GC did not even spot the errors. Most likely, the reason why GC did not exhibit any effectiveness was because its corpus was limited and, therefore, inefficient. It did not mark many of the errors and, therefore, did not even offer students the possibility to modify them.

In light of our findings, it can be concluded that, in contrast with computer-generated feedback, the students' performance improved after having received the teacher's feedback. Nevertheless, although noticeable, the effect of the teacher's feedback was not statistically significant. The use of GC as a tool to improve the students' use of the three targeted grammatical features was not as effective as expected because of the reasons aforementioned. However, it could be a helpful instrument for students and teachers if updated in the near future.

RQ2 aimed at analysing whether students obtaining feedback by means of GC would outperform those getting the teacher's feedback. Taking into account previous research (Dikli, 2006), H2 predicted that computer-generated feedback would be more effective than no feedback and at least as effective as the teacher's feedback.

The results of the between-groups analysis conducted revealed no statistically significant differences between the groups in ABBY\_pre, PAM\_pre, ABBY\_post and PAM\_post with the exception of PAM\_ART\_pre with an eta-squared value of 0.14, which means that the errors in articles committed by the three groups are close to belong to a different distribution, being 12 in Group 1, 4 in Group 2 and 3 in Group 3, and ABBY\_TEN\_post with an eta-squared value of 0.16 (large effect size). That means that the distribution of tense errors between the three groups differs, being 12 in Group 1, 2 in Group 2 and 4 in Group 3. The incidence of errors in Group 1 was much higher than in Group 2 and Group 3.

The results of the within-group analysis revealed no statistically significant differences when comparing ABBY\_pre vs PAM\_pre and ABBY\_post vs PAM\_post for the three targeted grammatical features in each of the three groups. A significant reduction in errors was only found in ABBY\_ART\_post vs. PAM\_ART\_post in Group 2 ( $Z_{(w)}=2.111$ ;  $p=.035$ ;  $d=1.119$ ). Results to take into account despite not being statistically significant are ABBY\_PREP\_post and PAM\_PREP\_post ( $d=0.632$ ), ABBY\_SUM\_post and PAM\_SUM\_post ( $d=0.678$ ) and, especially ABBY\_TEN\_post and PAM\_TEN\_post ( $d=0.744$ ) in Group 1 and ABBY\_SUM\_post vs. PAM\_SUM\_post ( $d=0.645$ ) and ABBY\_TEN\_pre vs. PAM\_TEN\_pre ( $d=0.548$ ) in Group 2.

As previously stated, the fact that GC left many errors unmarked made students not pay attention to them so modification did not take place. They were processed as meaningless bits and, sometimes, despite being grammatically correct, some errors were not structurally accurate (Ericsson & Haswell, 2006).

In line with previous studies, the teacher's feedback has proved to be beneficial (e.g., Bitchener, 2008, 2009; Bitchener & Knoch, 2008; Ellis, Sheen, Takashima & Murakami, 2008; Ferris, 2006; Sheen, Wright & Moldawa, 2009; Van Beuningen, Jong & Kuiken, 2008) and also



more effective than computer-generated feedback in detecting errors (e.g., Otoshi, 2005; Dikli & Bleyle, 2014). However, none of them has proved to improve students' performance.

Group 1 and Group 2 showed positive effects and outperformed Group 3, which showed no significant differences in any of the drafts. Although only the value for *ABBY\_ART\_post* vs. *PAM\_ART\_post* in Group 2 was statistically significant, other values such as *ABBY\_PREP\_post* vs. *PAM\_PREP\_post*, *ABBY\_SUM\_post* vs. *PAM\_SUM\_post* and, especially *ABBY\_TEN\_post* vs. *PAM\_TEN\_post* in Group 1, and *ABBY\_SUM\_post* vs. *PAM\_SUM\_post* ( $d=0.645$ ) and *ABBY\_TEN\_pre* vs. *PAM\_TEN\_pre* in Group 2 have a medium effect size. They can be regarded as interesting values which, although they do not show the expected effectiveness, they may point to a tendency towards improvement.

Similar to previous studies (e.g., Link et al., 2014; Tuzi, 2004; Wang, Shang & Briody, 2013), GC helped students improve their written accuracy and they also noticed the benefits of the software but, according to the statistical analyses, the results are not sufficiently significant. Both GC and the teacher's feedback proved to be equally useful in helping to reduce the incidence of errors, depending on different categories.

RQ3 attempted to explore whether the student's gains after having received both GC and the teacher's feedback would be maintained in the long term. Based on previous studies (e.g., Bitchener & Knoch, 2008; Link et al., 2020), H3 anticipated the long-term success of both types of feedback.

The analysis showed statistically significant differences between *TAILOR\_sum\_pre* and *TAILOR\_sum\_post* ( $p=.041$ ), and a large effect size between *TAILOR\_prep\_pre* and *TAILOR\_prep\_post* ( $d=0.827$ ) in Group 1. This value represents a significant reduction of errors in the *TAILOR\_sum\_post* in contrast with the *TAILOR\_sum\_pre*, which confirms the usefulness of GC in reducing the incidence of errors and the long-term effect of the feedback provided. In Group 2 results were even higher, since statistically significant differences were found in *TAILOR\_prep\_pre* vs. *TAILOR\_prep\_post*, *TAILOR\_ten\_pre* vs. *TAILOR\_ten\_post* ( $p=.024$ ) and *TAILOR\_sum\_pre* vs. *TAILOR\_sum\_post*. Results indicate that the teacher's feedback was effective in reducing the amount of errors, especially in prepositions and in the present/past perfect-simple past difference, maintaining the effect in the long term. In line

with previous studies (e.g., Ashoori Tootkaboni & Khatib, 2014; Bitchener & Knoch, 2008; Chandler, 2003; Erel, 2007; Ferris 2003; Ferris et al., 2000; Ferris & Helt, 2000; Li, 2010), the fact that the feedback provided by GC and the teacher engaged students in guided learning and problem solving may have helped students retain that knowledge in the long term.

Surprisingly, Group 3 also experienced a significant global reduction of errors in the tailor-made test ( $p=.017$ ). Improvement may be due to instruction and exposure to language in that period (Ferris, 2010). This finding confirms the relevant role input has in SLA (Ellis, 2008; Gass & Selinker, 1994; Grady et al., 2011; Patten & Benati, 2010; VanPatten & Williams, 2007).

Finally, RQ4 focused on the effect the students' attitude on the kind of feedback received has on the results by analysing the students' background, their technological abilities, teacher/computer dependence to learn, learning preferences, level of proficiency, the teacher's type of correction, and the characteristics of the software. Based on previous research (e.g., Bai & Hu, 2017; Chapelle et al., 2015; Cheng & Cheng, 2008; Dikli, 2006; Li et al., 2014; Warschauer & Ware, 2006), GC was expected to have a positive effect and engage students in the revision process to improve their writings.

To make the most of the feedback, it is important to have a positive attitude (Fredricks, 2013; Skinner & Pitzer, 2012). According to the questionnaire, participants felt comfortable with the means they received feedback with the exception of one student in Group 1 who chose *paper and pencil*, but she was satisfied with trying a new method. As shown in some of the comments in Section 6.4, students in Group 1 appreciated the use of computers in their learning process, something that, according to previous research (Fredricks, Blumenfeld & Paris, 2004), may have had an influence on the effectiveness of computer-generated feedback. In addition, they also reported not having had any technological problems with GC. As for Group 2, the participants did not have any constraint to choose a specific tool to write their compositions. They had the freedom of using any means for their writings, which should have influenced the students' writing practice positively.

As for the participants' preferences, both Group 1 and Group 2 stated that they appreciated being corrected all errors, with the exception of one student in Group 1 who chose Option b (*Teacher should mark all major but not the minor ones*) and another one who selected Options a,

b and d (*Teacher should mark no errors and respond only to the ideas and content*), and 2 students in Group 2 who chose Option b. In the case of the teacher, she corrected all types of errors as students expected, but in the case of GC, it only corrected spelling, combination of words and incorrect sequences. It was the teacher who corrected the other aspects after having received the computer feedback. Despite being aware of the limitations of GC, the fact that it worked as a first filter before handing in the final version may have motivated students in the revision process.

Aspects such as grammar, spelling and vocabulary in Group 1 were given high values, except for punctuation and organisation, which obtained lower scores. Students really appreciated being corrected on formal aspects, which may favour GC as it does not provide feedback on punctuation and organisation/content. Thus, the limitations of GC and the fact that students like being thoroughly corrected may have had a positive and/or negative effect on the students' attitudes. On the one hand, positive attitudes may have emerged because, as explained in the previous paragraph, GC may increase the students' motivation to improve their writings (Potter & Fuller, 2008); on the other hand, students' negative feelings could be caused because they usually trust automated feedback (e.g., Grimes & Warschauer, 2010; Li, Link, Ma, Yang & Hegelheimer, 2014; Rich, 2012), and so they may have overlooked some of the corrections proposed by GC.

As for Group 2, all aspects except content and spelling were highly valued. In this line, two participants even claimed that content correction was not useful. Therefore, taking into account these values, the teacher's role should not have had a negative influence on the students' attitudes when being corrected but a favourable one, since students were provided with feedback both on the highly valued aspects (i.e., organisation, grammar, vocabulary and punctuation) and on the less valued ones (i.e., spelling and content).

In sum, according to the students' background and opinions, GC limitations do not seem to have created significant obstacles so as to interfere negatively in the correction process. Students were conscious of those limitations but also of the positive aspects computer-generated feedback offers (e.g., El Ebyary & Windeatt, 2010; Elola & Oskoz, 2016; Liu & Sadler, 2003; Lu & Bol, 2007; Rich, 2012).

On the other hand, the fact that Group 1 also received the teacher's feedback on other aspects than the three grammatical features included in the study may have satisfied the students' needs. However, if GC is used alone without any other type of support may result in students' disappointment and demotivation, since it flags correct utterances as incorrect and omits errors. From the students' answers to the questionnaire, GC helps teachers to alleviate them in terms of time constraints, and students to receive immediate feedback and produce multiple drafts at their own pace (Chen & Cheng, 2008; Dikli, 2006; Hyland & Hyland, 2006). Thus, computer-generated feedback should be seen as a supplement and not as a teacher replacement (Brent & Townsend, 2006; Ware, 2011).

Finally, regarding the three aims explained in Chapter 4, it can be concluded that the present study has, on the one hand, contributed to a better understanding of computer-generated feedback and especially to GC on its effectiveness and weaknesses and the way it can help teachers and students in the learning and teaching process (Bai & Hu, 2017; Chen & Cheng, 2008; Dikli, 2006, Warschauer & Ware, 2006), and, on the other hand, it has shed light on the effectiveness of computer-generated feedback versus teacher feedback.

## 7.2 Limitations of the Study

The present study aimed at comparing the impact of teacher and computer feedback on specific students' B2 errors (i.e., articles, prepositions and past simple-present/past perfect difference) over an extended period of time. Due to the small sample, a fact that we will discuss below as the main limitation, the results and conclusions must be taken with caution. However, the study was conducted in a real context and, therefore, as far as its ecological validity is concerned, it complies with the veridicality and verisimilitude needed to generalise the study findings to real-life settings. The advantage of working in a real context can sometimes be unfavourable, since the students' commitment is not always the desired one. That means that some students' writings could not be analysed because, for example, they either failed to hand in some of them, submit the questionnaire, or even use the required tool (i.e., GC).

Apart from the inevitable obstacles of a real context, other limitations were found. The most remarkable one is the sample size. Out of the initial 44 students, only 27 participated in the

study, and this fact limits the generalisability of the results. However, in the period of time the study was conducted allowed students to hand in four essays, which deliberately required the use of the three targeted grammatical features. That allowed for constant and precise monitoring of these features so as to get a set of data to be analysed. This limitation could have been addressed by having recruited more students, but it was impossible because, as mentioned before, the study was conducted in a real setting where the teacher had only two B2 groups in that academic year and a specific syllabus to comply with.

The limitations of *Grammar Checker* were another constraint. Although it was chosen as the most appropriate software for our context at the time the study was conducted, a perfect AWE program to provide feedback does not exist yet. As mentioned before, GC is a low-cost program designed for Spanish students learning English, which served the purposes of our investigation. Nevertheless, the effectiveness of this software has proven to be insufficient. This fact is out of the teacher and students' control and, in some way, it is also part of the study as it tests students on their self-confidence and linguistic command. However, as mentioned in Section 6.4 and taking into account the students' questionnaires, any limitation of the software may have only influenced the participants' attitudes and performance to a small extent.

Another limitation refers to the questionnaire. After having conducted the statistical analysis and having interpreted the data, we may conclude that some of the questions could have been more precise in order to obtain better-quality data to be analysed. Interesting insights were gathered, but some claims were somehow vague due to the lack of further questions to obtain more specific information. This constraint could have been solved with more precise questions or even with a personal interview between the students and the researcher after each draft or at the end of the study. However, asking students to devote extra time to have a personal interview would have been unsuitable, since they come to class in their free time after their workday and sometimes have problems to come or even hand in their work.

### **7.3 Contribution of the Study and Further Research**

Despite the limitations described in the previous section, the present study has contributed to widen the body of research on computer-generated feedback, since it has tried to capture the

difficulties of real foreign language instruction and assessment as well as shed light on the existing literature of teacher's feedback and on the effectiveness of AWE programs. Many studies deal with AWE or feedback provided by the teacher, but hardly any, to our knowledge, combine both in the same study. The comparison of both types of feedback in the same study represent an innovative piece of research in the field of computer-generated feedback.

In spite of the low number of participants, this study took place in a real context with participants who were not conscious they were being part of an investigation on feedback. Additionally, a control group, who did not receive any feedback on the three targeted grammatical features, was included. Many studies fail to provide accurate results due to the lack of a control group. However, in the present study a control group was included to test whether gains were due to intervention.

The comparison between teacher's feedback and computer-generated feedback has also helped to better understand the effectiveness of AWE systems and the role that both kinds of feedback play in SLA. In addition, the error analysis conducted for each error category has brought to light the relevance it has for SLA to know the reason why EFL students commit errors (e.g., knowledge of the language, L1 transfer, -ART language, lack of input, lack of one-form-one-function correspondence, etc.), and give an appropriate solution for each type of error.

Future research regarding AWE systems should, on the one hand, attempt to give support to GC designers in order to improve its effectiveness and be able to recommend this tool to Spanish students of English, and, on the other, to continue research on feedback adequacy, that is to say, the relation between error and type of feedback (Ferris, 1999). The required type of CF may depend on a wide range of aspects, such as the nature of the error, the learner's existing knowledge or their cognitive processing. Using the same type of feedback for all types of errors or restricting one type of CF to certain errors may be too simplistic. Most likely, using different ways of providing feedback or even a combination can be more enriching.

In addition, the current study identifies a thought-provoking gap in how the students' use of new technologies can influence their willingness and preference to receive teacher's or computer's feedback. The questionnaire conducted after the study has contributed to know the possible effects the teacher's feedback and AWE can have on the students' attitudes. Thus,

future research on this topic could focus more closely on the insights into computer-generated feedback before and after using it so as to provide more exhaustive and comprehensive evidence.

## 7.4 Pedagogical Implications

Notwithstanding the handicap of the small sample size, the results obtained have provided valuable insights in the field of computer-generated feedback on student writing and, more specifically, on the use of GC as a complement to the teacher's corrections.

Grammar transfer errors are common when learning a language, as revealed in this study. However, errors occur as a natural part of the learning process (Corder, 1981) and as such they should be addressed correctly. For this reason, the nature of errors is an important aspect to tackle before providing feedback. For example, as explained in Section 1.5, not all the prepositions or articles are acquired in the same way, because some may be acquired more easily owing to, for instance, L1 transfer, nature of the word, meaning or use of the word (e.g., Bong, 2010; 2011; 2012; Hayashi 2001; 2008). Hence, some prepositions can be better taught in chunks instead of isolatedly so that students notice they are imposed by the previous or subsequent elements (Díez-Bedmar & Casas-Pedrosa, 2011).

CF is another relevant feature that should be dealt with carefully in the EFL classroom. Taking into account the number of students in a group, the diversity of errors in each piece of writing and the frequent practice students need, providing written CF is a hard and time-consuming task. For this reason, according to Ferris (2011) and Guénnete (2012), some strategies to avoid teachers' burnout can be applied (e.g., setting realistic goals, being selective, training learners to self-edit, etc.). In addition, different types of CF (i.e., teacher, computer-generated, peer feedback, as well as implicit and explicit feedback, etc.) can be provided depending on the type of error to make the most of it. In order to do so, it is also interesting to have some background information about the students (e.g., mother tongue, proficiency level, computer literacy, etc.).

The debate whether direct or indirect CF is more effective in the long or short term has provided contradictory results. While some studies report positive evidence of indirect feedback in the long term (Chandler, 2003; Ferris, 2003, 2006; Fratzen, 1995) others suggest only direct feedback is effective in the long term (van Beuningen et al., 2008). Thus, taking into account

that the feedback used by the teacher and GC was the colour code, that is, indirect, the present research has contributed to this debate.

The results of the analyses performed through the data gathered during the study may help identify what tools can be beneficial for students and teachers in the EFL context. Despite its ecological validity, had this study been conducted with a bigger sample size, the results could have shed even more light on the effects of both types of feedback, teacher and computer. Computer-generated feedback benefits both teachers and students, but it is key for success to decide what role it has (e.g., teacher complement or substitute, only grammar/spelling corrector, etc.) in the language classroom and how often it should be employed.

To sum up, the present study may be a significant contribution to the field of WCF, especially computer-generated feedback, but still further research is necessary in order to examine the effectiveness of AWE software thoroughly.



## References

- AbuSeileek, A. F. (2012). The effect of computer-assisted cooperative learning methods and group size on the EFL learners' achievement in communication skills. *Computers & Education*, 58(1), 231-239. Elsevier Ltd. Retrieved October 30, 2019 from <https://www.learntechlib.org/p/50707/>.
- Ahmadvand, M. (2008). Analyzing errors of Iranian EFL learners in their written productions. Retrieved from <http://moslem17.googlepapers.com/AnalysingerrorsofIranianEFLlearners.pdf> March 15, 2019.
- Akay, C., and Akbarov, A. (2011). Corrective feedback on the oral production and its influence in the intercultural classes, *IBU Repository*. Retrieved October 30, 2021 from <https://omeka.ibu.edu.ba/items/show/2700>.
- Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language acquisition* (pp. 259-302). Honolulu: University of Hawaii Press.
- Allaei, S. K., and Connor, U. (1990). Using performative assessment instruments with ESL student writers, in Hamp-Lyons, L. (Ed.) *Assessing Second Language Writing in Academic Contexts* (pp. 227-240), Norwood, NJ: Ablex.
- Allwright, R. L. (1975). Problems in the study of the language teacher's treatment of learner error. In M. K. Burt and H. C. Dulay (Eds.), *On TESOL'75: New Directions in Second Language Learning, Teaching and Bilingual Education* (96-109). Washington, D. C.
- Allwright, D., and Bailey, K. M. (1991). *Focus on the Language Learner*. Cambridge: Cambridge University Press.
- Alonso-Alonso, R. (1997). Language transfer: Interlingual Errors in Spanish Students of English as a Foreign Language. *Revista Alicantina de Estudios Ingleses*, 10, 7-14.
- Alonso-Alonso, R., Cadierno, T. and Jarvis, S. (2016). Crosslinguistic Influence in the Acquisition of Spatial Prepositions in English as a Foreign Language. In R. Alonso-Alonso (Ed.), *Crosslinguistic Influence in Second Language Acquisition* (pp. 93-120). Bristol, Multilingual Matters. <https://doi.org/10.21832/9781783094837-008>
- Al-Magid, M. A. (2006). The Effect of Teachers' Attitudes on the Effective Implementation of the Communicative Approach in ESL Classroom. Master's thesis, University of South Africa. Retrieved May 7, 2020 from <http://uir.unisa.ac.za/bitstream/handle/10500/1747/dissertation.p4>.
- Ammar, A., and Spada, N. (2006). One size fits all?: Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28(4), 543-574. <https://doi.org/10.1017/S0272263106060268>
- Amrhein, H. and Nassaji, H. (2010). Written corrective feedback: What do students and teachers prefer and why?. *Canadian Journal of Applied Linguistics*, 13, 95-127.
- Andersen, R., and Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in Second Language Acquisition*, 16, 133-156.
- Andersen, R., and Shirai, Y. (1996). Primacy of aspect in first and second language acquisition: The pidgin/creole connection. In W. C. Ritchie and T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 527-570). San Diego, CA: Academic Press.
- Anderson, T. (2010). The effects of tiered corrective feedback on second language academic writing. Unpublished MA thesis. University of British Columbia. [https://www.researchgate.net/publication/263385090\\_Anderson\\_T\\_2010\\_The\\_effects\\_of\\_tiered\\_corrective\\_feedback\\_on\\_second\\_language\\_academic\\_writing\\_Unpublished\\_MA\\_thesis\\_University\\_of\\_British\\_Columbia](https://www.researchgate.net/publication/263385090_Anderson_T_2010_The_effects_of_tiered_corrective_feedback_on_second_language_academic_writing_Unpublished_MA_thesis_University_of_British_Columbia)

- Anson, C. (2006). Can't touch this: Reflections on the servitude of computers as readers. In P. F. Ericsson, and R. Haswell (Eds.), *Machine scoring of human essays* (pp. 38-56). Logan, UT: Utah State University Press.
- Ashcraft, M. H., and Radvansky, G. A. (2009). *Cognition*. Upper Saddle River, N. J.: Pearson Education.
- Ashoori Tootkaboni, A. and Khatib, M. (2014). The Efficacy of Various Kinds of Error Feedback for Improvement in the Writing Accuracy of Iranian EFL Learners. *Bellaterra Journal of Teaching & Learning Language & Literature*, 7, 30. doi: 10.5565/rev/jtl3.529.
- 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), 227-257.
- Asma, T. (2010). Transfer of simple prepositions from standard Arabic into English: The case of third year LMD students of English language at Mentouri University-Constantine. Unpublished M.A. Thesis. Mentouri University-Constantine, Algeria. Retrieved October 29, 2021 from: <https://bu.umc.edu.dz/theses/anglais/TEM1143.pdf>.
- Attali, Y. (2004). Exploring the Feedback and Revision Features of Criterion. *Journal of Second Language Writing*, 14. Retrieved August 8, 2021 from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.214.9047&rep=rep1&type=pdf>
- Avery, P., and Radišić, M. (2007). Accounting for variability in the acquisition of English articles. In A. Belikova, L. Meroni, and Umeda, M. (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 1-11). Somerville, MA: Cascadilla Proceedings Project.
- Ayedh, A., and Khaled, E. (2011). EFL teachers' feedback to oral errors in EFL classroom: Teachers' perspectives. *Arab World English Journal*, 2, (1), 214-232.
- Ayoun, D., and Salaberry, M. (2008). Acquisition of English tense-aspect morphology by advanced French instructed learners. *Language Learning*, 58, 555-595.
- Azar, B. (2007). Grammar-based teaching: A practitioner's perspective. *Teaching English as a Second or Foreign Language* 11(2), 1-12.
- Azizollah, D. (2008). A Comparison of the Effects of Implicit and Explicit Corrective Feedback on Learners' Performance in Tailor-Made Tests. *Journal of Applied Sciences*, 8, 1-13.
- Bai, L., and Hu, G. (2017). In the face of fallible AWE feedback: how do students respond? *Educational Psychology*, 37, 67-81. doi:10.1080/01443410.2016.1223275.
- Baleghizadeh, S. and Dadashi, M. (2011). The effect of direct and indirect corrective feedback on students' spelling errors. *Profile: Issues in Teachers' Professional Development*, 13(1), 129-137. Retrieved September 11, 2017 from <https://revistas.unal.edu.co/index.php/profile/article/view/20572/36835>.
- Ballesteros F., Rica, J. P., Neff Van Aertselaer, J. and Díez Prados, M. (2005). The ICLE error tagging project: analysis of Spanish EFL writers. In C. Mourón Figueroa, and T. I. Moralejo Gárate (Eds.) *Studies in Contrastive Linguistics. Proceedings of the 4th International Contrastive Linguistics Conference* (pp. 89-97). Santiago de Compostela: Servizo de Publicacións e Intercambio Científico.
- Bang, Y. (1999). Reactions of EFL students to oral error correction. *Journal of Pan-Pacific Association of Applied Linguistics*, 3, 39-51.
- Barcelos, A., and Kalaja, P. (2003). Exploring possibilities for future research on beliefs about SLA. In Kalaja, P., and Ferreira Barcelos, A. M. (Eds.) *Beliefs about SLA: new research approaches* (231-238). Dordrecht: Kluwer Academic Publishers.
- Bardovi-Harlig, K. (1992). The relationship between form and meaning: A cross-sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied*

- Psycholinguistics*, 13, 253-278.
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.
- Bardovi-Harlig, K. (2002). Another piece of the puzzle: The emergence of the present perfect. *Language Learning* 47(3), 375-422.
- Bataineh, R. (2005). Jordanian undergraduate EFL students' errors in the use of the indefinite. *The Asian EFL Journal Quarterly*, 7(1), 56-76.
- Batstone, R. (1994). *Grammar*. Oxford: Oxford University Press.
- Bazo Martínez, P. (2001). A qualitative analysis of the interlanguage found in compulsory secondary education students in the Canary Islands. *Revista Canaria de Estudios Ingleses*, 43, 209-218.
- Benati, A. (2017). The role of input and output tasks in grammar instruction: Theoretical, empirical and pedagogical considerations. *Studies in Second Language Learning and Teaching*, 7, 377-396.
- Benítez-Pérez, P., and Simón-Granda, J. (1990). Adquisición de la preposición en la primera y segunda lengua. *Revista Española de Lingüística Aplicada*, 6, 139-148.
- Berg, C. E. (1999). The effects of trained peer response on ESL students' revision types and writing quality. *Journal of Second Language Writing*, 8(3), 215-241.
- Birdsong, D., and Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44(2), 235-249. <https://doi.org/10.1006/jmla.2000.2750>.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17, 102-118. doi: 10.1016/j.jslw.2007.11.004.
- Bitchener, J. (2009). Measuring the effectiveness of written corrective feedback: A response to "Overgeneralization from a Narrow Focus: A Response to Bitchener (2008)." *Journal of Second Language Writing*, 18(4), 276-279. doi: <https://doi.org/10.1016/j.jslw.2009.06.001>
- Bitchener, J. (2012). Written corrective feedback for L2 development: Current knowledge and future research. *TESOL Quarterly*, 46(4), 855-860. Retrieved August 23, 2020, from <http://www.jstor.org/stable/43267894>
- Bitchener, J. and Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. New York: Routledge.
- Bitchener, J., and Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research Journal*, 12(3), 409-431.
- Bitchener, J., and Knoch, U. (2009). The value of a focused approach to written corrective feedback. *ELT Journal*, 63. 10.1093/elt/ccn043.
- Bitchener, J., and Knoch, U. (2010a). The contribution of written corrective feedback to language development: A ten month investigation. *Applied linguistics*, 31(2), 193-214. doi:10.1093/applin/amp016.
- Bitchener, J., and Knoch, U. (2010b). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207-217. doi:10.1016/j.jslw.2010.10.002
- Bitchener, J., Young, S., and Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 9, 227-258.
- Black, P., and William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5(1), 7-73.
- Black, P., and William, D. (2009). Developing the Theory of Formative Assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31. <https://doi.org/10.1007/s11092-008-9068-5>.
- Block, D. (2003). *The social turn in second language acquisition*. Edinburgh, Scotland: Edinburgh University Press.

- Bloomfield, L. (1933). *Language*. New York: Holt, Reinhart and Winston.
- Bolt, P. (1992). An evaluation of grammar-checking programs as self-help learning aids for learners of English as a foreign language. *Computer Assisted Language Learning*, 5 (1-2), 49-91.
- Bong, H. K. (2005). *Economical Parameter-Setting in Second Language Acquisition: Japanese Speaking Learners of English*. Doctoral Dissertation. University of Cambridge.
- Bong, H. K. (2008). *Semantics of Types of English Prepositions: Acquisition of English Prepositions*. Presented in ELSJ Spring Forum 2008. Manuscript. 4.
- Bong, H. K. (2009). *A Minimalist Model of Language Acquisition*. UK: VDM Verlag.
- Bong, H. K. (2010). Acquisition of English prepositions. *Proceedings of the 15th International Conference of Pan-Pacific Association of Applied Linguistics (PAAL)*, 158-165.
- Bong, H. K. (2011). *Lemmatic transfer in the second language acquisition of English prepositions*. Proceedings of the 16th Conference of Pan-Pacific Association of Applied Linguistics. Hong Kong: Pan-Pacific Association of Applied Linguistics, 109-116.
- Bong, H. K. (2012). Acquisition of the English Preposition 'at'. *Journal of Humanities and Social Sciences*, 6, 148-164.
- Bong, H. K. (2013). Acquisition of the English Preposition 'on': Assessing the Prototypicality Hypothesis. *Journal of Humanities and Social Sciences*, 7, 142-156.
- Bourke, R., and Mentis, M. (2011). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, 17(8), 854-867. doi:10.1080/13603116.2011.602288.
- Brent, E., and Townsend, M. (2006). Automated essay grading in the sociology classroom. In P. Freitag Ericsson, & R. H. Haswell, (Eds.) *Machine Scoring of Student Essays: Truth and Consequences?* (pp. 177-198). Utah State University Press.
- Bromely, K. (2005). *Technology and Writing. The International Handbook of Literacy and Technology: Volume II*. New York: Routledge.
- Brown, A. V. (2009). Students' and teachers' perceptions of effective foreign language teaching: a comparison of ideals. *The Modern Language Journal*, 93, (1), 46-60.
- Brown, H. D. (2007). *Principles of language learning and teaching*. White Plains, NY: Pearson Education.
- Bruno, I., and Santos, L. (2010). Written comments as a form of feedback. *Studies in Educational Evaluation*, 36, 111-120.
- Bueno-González, A. (1992). Errores en la elección de palabras en inglés por alumnos de Bachillerato y C.O.U. In Bueno-González, A., Carini-Martínez, J. A, & Linde-López, Á., (Eds.) *Análisis de errores en inglés: tres casos prácticos* (pp. 39-105). Granada: Servicio de Publicaciones de la Universidad de Granada.
- Bulut, T. (2011). The licensing of the English present perfect tense by Turkish adults. *International Journal of Humanities and Social Science*, 1(15), 221-227.
- Burke, D. (2014). "Feedback was very helpful": the benefits of feedback on drafts for students' learning. *DISKUS*, 14. doi: 10.18792/diskus.v14i0.18.
- Burt, M. K. (1975). Error analysis in the adult EFL classroom. *TESOL Quarterly*, 9, 53-63.
- Burstein, J., and Chodorow, M. (1999). Automated essay scoring for nonnative English speakers. In *Proceedings of the ACL99 Workshop on Computer-Mediated Language Assessment and Evaluation of Natural Language Processing*. College Park, MD.
- Burstein, J., and Chodorow, M. and Leacock, C. (2003). CriterionSM: Online essay evaluation: An application for automated evaluation of student essays. *Proceedings of the Fifteenth Annual Conference on Innovative Applications of Artificial Intelligence*.
- Burstein, J., and Marcu, D. (2003). Developing Technology for Automated Evaluation of Discourse Structure in Student Essays. In M. D. Shermis, & J. C. Burstein (Eds.), *Automated Essay Scoring: A Cross-disciplinary Perspective* (pp. 209-230). Hillsdale, NJ: Lawrence Erlbaum.

- Burston, J. (2001). Computer-Mediated Feedback in Composition Correction. *CALICO Journal*, 19, 37-50. doi:10.1558/cj.v19i1.
- Burston, J. (2008). BonPatron: An online spelling, grammar, and expression checker. *CALICO Journal*, 25(2), 337-347.
- Butler, Y. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition*, 24, 451-480.
- Butler, A. C., Karpicke, J. D., and Roediger, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology*, 13, 273-281.
- Butzkamm, W., and Caldwell, J. (2009). *The Bilingual Reform: A paradigm shift in foreign language teaching*. Tübingen: Gunter Narr Verlag.
- Byrd, D. (2003). Practical tips for implementing peer editing tasks in the foreign language classroom. *Foreign Language Annals*, 26(3), 434-441.
- Cai, J. T. (2001). *The influence of Chinese aspect markers on the use of tense and aspect in English as a foreign language*. Paper presented at the 2nd International Conference on English Teaching in Asia, Nanjing.
- Calvo, R., and Ellis, R. (2010). Students' conceptions of tutor and automated feedback in professional writing. *Journal of Engineering Education*, 99, 427-438.
- Camajoan, L. (2006). The Aspect Hypothesis: Development of morphology and appropriateness of use. *Language Learning*, 56, 201-268.
- Carranza, L. M. V. (2007). Correction in the ESL classroom: What teachers do in the classroom and what they think they do. *Revista Pensamiento Actual*, 7(8-9), 84-95.
- Carlucci, L., and Case, J. (2013). On the necessity of U-shaped learning. *Top Cognitive Science*, 5(1), 56-88. doi: 10.1111/tops.12002. PMID: 23335574.
- Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry, & C. W. Stansfield (Eds.), *Language Aptitude Reconsidered* (pp. 11-29). Englewood Cliffs, NJ: Prentice Hall Regents.
- Carroll, S. (1995). The irrelevance of verbal feedback to language learning. In L. Eubank, L. Selinker, and M. S. Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford* (pp. 73-88). Philadelphia: John Benjamins.
- Carroll, S. (1999). Putting "input" in its proper place. *Second Language Research*, 15, 337-388.
- Carroll, S. (2001). *Input and evidence: The raw material of second language acquisition*. Amsterdam: John Benjamins.
- Carroll, S., and Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalization. *Studies in Second Language Acquisition*, 15, 357-386. <http://dx.doi.org/10.1017/S0272263100012158>.
- Carson, J. (2001). Second language writing and second language acquisition. In T. Silva, & P. Matsuda (Eds.), *On second language writing* (191-200). Mahwah, NJ: Lawrence Erlbaum.
- Carson, J. and Nelson, G. (1996). Chinese students' perception of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1-19.
- Casas-Pedrosa, A. V. (2005). Teaching English prepositions to Spanish speakers: some aspects to bear in mind. *Greta*, 13(1-2), 94-101.
- Castro, M. (2013). An analysis of prepositional errors of college students. *Foreign Language Learning and Testing Conference. Proceedings by Language Institute Thammasat University*, 2 (1), 97-108. Retrieved June 25, 2018 from <http://litu.tu.ac.th/2012/index.php/flt-conferences/521-flt-2013-conference-proceedings>
- Catalan, R. J. (1996). Frequency and variability in errors in the use of English prepositions. *Miscellanea: A Journal of English and American Studies*, 17, 171-187.
- Cavaleri, M., and Dianati, S. (2016). You want me to check your grammar again? The usefulness

- of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, 10(1), 223-236.
- Cavanaugh, A. J., and Song, L. (2014). Audio feedback versus written feedback: Instructors' and students' perspectives. *Journal of Online Learning and Teaching*, 1, 122-138.
- Celce-Murcia, M. (2001). *Teaching English as a Second or foreign Language*. Boston: Heinle & Heinle.
- Celce-Murcia, M., and Larsen-Freeman, D. (1999). *The Grammar Book: An ESL/EFL Teacher's Course*. Boston, MA: Heinle and Heinle.
- Chacón, R. (2016). *Cloud computing en entornos Start-up educativos*. Retrieved February 26th, 2018 from <https://www.emadridnet.org/index.php/es/28-eventos-y-seminarios/525-cloud-computing-en-entornos-start-up-educativos-el-ejemplo-de-la-herramienta-grammar-checker-como-medio-para-aprender-ingles>
- Chan, H., Jessica, F., Willie, C., and Yasuhiro, S. (2012). L2 acquisition of tense-aspect morphology. In Filipovic, L. & Jaszczolt, K. M. (Eds.) *Space and Time in Languages and Cultures* (pp. 181-204). Amsterdam: John Benjamins Publishing Company. <http://dx.doi.org/10.1075/hcp.36.12chan>.
- 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, 267-296. doi:10.1016/S1060-3743(03)00038-9
- Chandler, J. (2009). Response to Truscott. *Journal of Second Language Writing*, 18, 57-58. doi: 10.1016/j.jslw.2008.09.002.
- Chapelle, C. A. (1998). Multimedia CALL: Lessons to be learned from research on instructed SLA. *Language Learning & Technology*, 2(1), 21-39. Retrieved January 18, 2017 from <http://llt.msu.edu/vol2num1/pdf/article1.pdf>
- Chapelle, C. A., and Chung, Y. A. (2010). The promise of NLP and speech processing technologies in language assessment. *Language Testing*, 27(3), 301-315.
- Chapelle, C. A., Cotos, E., and Lee, J. Y. (2015). Validity arguments for diagnostic assessment using Automated Writing Evaluation. *Language Testing*, 32, 385-405. <https://doi.org/10.1177/0265532214565386>
- Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning*, 27, 29-46.
- Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition*, 7(1), 1-14. doi:10.1017/S027226310000512X.
- Chaudron, C. (1988). *Second Language Classroom: Research on Teaching and Learning*. New York: Cambridge University Press.
- Chaudron, C., and Parker, K. (1990). Discourse markedness and structural markedness: The acquisition of English noun phrases. *Studies in Second Language Acquisition*, 12, 43-64.
- Chen, H. C. (2000). Error Analysis of some features of English article usage. *Journal of Wu-Feng Applied Linguistics*, 8, 282-296.
- Chen, L. L. (2006). The effect of L1 and CAI on grammar learning: An error analysis of Taiwanese beginning EFL learners' English essays. *Asian EFL Journal*, 9(2), 1-19.
- Chen, C. and Cheng, W. (2006). *The Use of a Computer-Based Writing Program: Facilitation or Frustration?* Paper presented at the 23 International Conference on English Teaching and Learning, Wenzao Ursuline College of Languages, Kaohsiung.
- Chen, C. and Cheng, W. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning and Technology*, 12(2), 94-112.
- Chesterman, A. (1991). *On definiteness: A study with special reference to English and Finnish*. Cambridge: Cambridge University Press.

- Chiang, K. K. (2004). An investigation into students' preferences for and responses to teacher feedback and its implication for writing teachers. *Hong Kong Teachers' Centre Journal*, 3, 98-115.
- Chodorow, M., Gamon, M., and Tetreault, J. (2010). The utility of article and preposition error correction systems for English language learners: Feedback and assessment. *Language Testing*, 27(3), 419-436. <https://doi.org/10.1177/0265532210364391>
- Chodorow, M., Tetreault, J. & Han, N. (2007). Detection of grammatical errors involving prepositions, *Proceedings of the 4th ACL-SIGSEM Workshop on Prepositions*, (pp. 25-30). doi: 10.3115/1654629.1654635.
- Chomsky, N. (1964). *Current issues in linguistic theory*. Papers presented at the Ninth International Congress of Linguists, Cambridge, Massachusetts.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA.: MIT Press.
- Chou, H.-N., Moslehpour, M., and Chung-Yu Y. (2016). *My Access!* and Writing Error Corrections of EFL College Pre-intermediate Students. *International Journal of Education*, 8, 144-161. 10.5296/ije.v8i1.9209.
- Chua, R. M., Ferrer, R., Quijano, L. J., and Santos, J. A. (2015). Interference between Filipino and English prepositions: A comparative analysis of male and female Filipino bilingual learners in the World Englishes paradigm in the Philippine setting. *DLSU Research Congress*, 3, 1-9.
- Ciesielkiewicz, M., and Márquez, E. D. (2015). Error analysis and its relevance to teaching ESL composition. *International Journal of Linguistics*, 7, 119-138. Retrieved August 12, 2020 from <http://www.macrothink.org/journal/index.php/ijl/article/view/8076>
- Clarke, S. (2005). *Formative assessment in action: Weaving the elements together*. London: Hodder and Stoughton.
- Clarke, S., Timperley, H., and Hattie, J. A. (2003). *Assessing formative assessment*. Auckland, New Zealand: Hodder Moa Beckett.
- Collins, L. (2002). The roles of L1 influence and lexical aspect in the acquisition of temporal morphology. *Language Learning*, 52, 43-94.
- Collins, L. (2004). The Particulars on Universals: A Comparison of the Acquisition of Tense-Aspect Morphology among Japanese- and French-Speaking Learners of English. *The Canadian Modern Language Review*, 61(2), 251-274. doi:10.1353/cml.2005.0002.
- Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.
- Condon, W. (2013). Large-scale assessment, locally-developed measures, and automated scoring of essays: Fishing for red herrings. *Assessing Writing*, 18(1), 100-108.
- Coniam, D. (2009). Experimenting with a computer essay-scoring program based on ESL student writing scripts. *ReCALL*, 21(2), 259-279. doi:10.1017/S0958344009000147
- Cook, V. (1991). *Second Language Learning and Second Language Teaching*. Oxford: Blackwell.
- Cook, V. (1993). *Linguistics and Second Language Acquisition*. London: Macmillan.
- Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5(1-4), 161-170. doi:10.1515/iral.1967.5.1-4.161
- Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics*, 5, 161-170.
- Corder, S. P. (1973). *Introducing applied linguistics*. Middlesex, Penguin.
- Corder, S. P. (1974). *Error Analysis: Perspectives on Second Language Acquisition*. London: Longman.
- Corder, S. P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- Cotos, E. (2011). Potential of Automated Writing Evaluation Feedback. *English Publications*, 58. From [https://lib.dr.iastate.edu/engl\\_pubs/58](https://lib.dr.iastate.edu/engl_pubs/58) (retrieved February 16th 2021)

- Cotos, E. (2012). Towards effective integration and positive impact of automated writing evaluation in L2 writing. Retrieved February 15, 2021 from [https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1062&context=engl\\_pubs](https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=1062&context=engl_pubs)
- Cotos, E. (2015). Automated Writing Analysis for writing pedagogy: From healthy tension to tangible prospects. *Writing & Pedagogy. English Publications*, 63. doi: [https://lib.dr.iastate.edu/engl\\_pubs/63](https://lib.dr.iastate.edu/engl_pubs/63)
- Cowan, R., H. E. Choi & D. H. Kim. (2003). Four questions for error diagnosis and correction in CALL. *CALICO Journal*, 20(3), 451-463.
- Creel, S. C., Newport, E. L., and Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(5), 1119-1130. <https://doi.org/10.1037/0278-7393.30.5.1119>
- Cunningham, D. (1998). 25 years of technology in language teaching: A personal experience. *Babel: Journal of the Australian Federation of Modern Language Teachers' Associations*, 33(1), 4-7.
- Dagen, A., Matter, C., Rinehart, S., and Ice, P. (2008). Can you hear me now? Providing feedback using audio commenting technology. *College Reading Association Yearbook*, 29, 152-166.
- Dagneaux, E., Denness, S. and Granger, S. (1998). Computer-aided error analysis. *System*, 26, 163-174. doi: 10.1016/S0346-251X(98)00001-3.
- Declerck, R. (2006). *The Grammar of the English Tense System: A Comprehensive Analysis*. Berlin: Mouton de Gruyter.
- DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *Modern Language Journal*, 77, 501-514.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Oxford: Blackwell.
- DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55, 1-25.
- Denton, P., Madden, J., Roberts, M., and Rowe, P. (2008). Students' response to traditional and computer-assisted formative feedback: A comparative case study. *British Journal of Educational Technology*, 39(3), 486-500.
- De Grez, L., Valcke, M., and Roozen, I. (2012). How effective are self- and peer assessment of oral presentation skills compared with teachers' assessments? *Active Learning in Higher Education*, 13(2), 129-142. doi:10.1177/1469787412441284
- Diab, R. L. (2005). EFL university students' preference for error correction and teacher feedback on writing. *TESL Reporter*, 38(1), 27-51.
- Diab, R. L. (2006). Error correction and feedback in the EFL writing classroom: Comparing instructor and student preferences. *English Teaching Forum*, 44(3), 2-13.
- Diab, N. M. (2010). *Effects of peer versus self-editing on students' revision of language errors in revised drafts*. *System*, 38, 85-95. <http://dx.doi.org/10.1016/j.system.2009.12.008>.
- Díez-Bedmar, M., and Casas-Pedrosa, A. (2011). The use of prepositions by Spanish learners of English at University level: a longitudinal analysis. In N. Kübler (Ed.). *Corpora, Language, Teaching, and Resources: From Theory to Practice* (pp. 199-218). Bern: Peter Lang. Retrieved Aug 4, 2021, from [https://www.academia.edu/7740605/The\\_use\\_of\\_prepositions\\_by\\_Spanish\\_learners\\_of\\_English\\_at\\_University\\_level\\_a\\_longitudinal\\_analysis](https://www.academia.edu/7740605/The_use_of_prepositions_by_Spanish_learners_of_English_at_University_level_a_longitudinal_analysis)
- DiGiovanni, E., and Nagaswami, G. (2001). Online peer review: An alternative to face-to-face? *ELT Journal*, 53(3), 263-272.
- Dikli, S. (2006). An Overview of Automated Scoring of Essays. *Journal of Technology, Learning, and Assessment*, 5(1), 1-35.
- Dikli, S., and Bley, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing*, 22, 1-17.



- Dochy, F., Segers, M., and Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24(3). Retrieved October 15, 2018 from <http://www.tandfonline.com/doi/abs/10.1080/03075079912331379935?src=recsys#.VOGj7S5UMsI>
- Doughty, C. (1994). Fine-tuning of feedback by competent speakers to language learners. In J. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics 1993: Strategic Interaction and Language Acquisition: Theory, Practice, and Research* (pp. 96-108). Washington, D. C.: Georgetown University Press.
- Doughty, C. (2001). Cognitive underpinnings of focus on form. In Robinson, P. (Ed.), *Cognition and second language instruction* (pp. 206-257). New York: Cambridge University Press.
- Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In Doughty, C., and Long, M. H. (Eds.), *The handbook of second language acquisition* (pp. 257-310). Oxford: Blackwell.
- Doughty, C., and Long, M. H. (2003). *The handbook of second language acquisition*. Oxford: Blackwell.
- Doughty, C., and Varela, E. (1998). Communicative focus on form. In C. Doughty, & J. Williams (Eds.) *Focus on form in classroom second language acquisition* (pp. 114-138). New York: Cambridge University Press.
- Doughty, C., and Williams, J. (Eds.). (1998). *Focus on forms in classroom second language acquisition*. New York: Cambridge University Press.
- Dulay, H. and Burt, M. (1974). Natural sequence in child second language acquisition. *Language Learning*, 24, 37-53.
- Dulay, H., and Burt, M. (1977). *Viewpoints on English as a Second Language*. New York: Regents.
- Dulay, H., Burt, M., and Krashen, S. (1982). *Language two*. New York: Oxford University Press.
- Dunning, D., Heath, C., and Suls, J. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69-106. Retrieved January 28, 2021, from <http://www.jstor.org/stable/40062350>
- Egbert, J. L. (2005). Conducting research on CALL. In J. L. Egbert, & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 3-8). Mahwah, NJ: Lawrence Erlbaum.
- Ekiert, M. (2004). Acquisition of the English article system by speakers of Polish in ESL and EFL settings. Teachers College, *Columbia University Working Papers in TESOL & Applied Linguistics*, 4(1). Retrieved March 10, 2020 from <https://academiccommons.columbia.edu/doi/10.7916/D8P84BCK>.
- Elashri, I. I. (2013). The impact of the direct teacher feedback strategy on the EFL secondary stage students' writing performance. Retrieved February 3, 2017 from <https://files.eric.ed.gov/fulltext/ED540631.pdf>
- Elliot, S. (2003). Intellimetric™: From here to validity. In M. D. Shermis and J.C. Burstein (Eds.), *Automatic essay scoring: A cross-disciplinary perspective* (pp. 71-86). Mahwah, NJ: Lawrence Erlbaum Associates.
- Elliot, S., and Mikulas, C. (2004, April). *The impact of My Access use on student writing performance: A technology overview and four studies*. Paper presented at the Annual meeting of the American Educational Research Association, San Diego, California.
- Ellis, N. C. (2002). Frequency Effects in Language Processing. *Studies in Second Language Acquisition*, 2(24), 143-188.
- Ellis, R. (1991). *The interaction hypothesis: A critical evaluation*. Paper presented at the Regional Language Center Seminar, Singapore, April 22-28, 1991. (ERIC document no. ED338037).
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics*, 16,

- 409-441.
- Ellis, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1998). Teaching and research: Options in grammar teaching. *TESOL Quarterly*, 32 (1),
- Ellis, R. (2001). Investigating form-focused instruction. *Language Learning*, 51, 1-46.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). Principles of instructed language learning. *System*, 7, 209-224. doi: 10.1016/j.system.2004.12.006.
- Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. *AILA Review*, 19, 18-41.
- Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 339-360). Oxford: Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63, 97-107.
- Ellis, R. (2010). A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition*, 32(2), 335-349.
- Ellis, R. (2015). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Ellis, R., Basturkmen, H., and Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51, 281-318.
- Ellis, R., and He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21(2), 285-301. doi:10.1017/S0272263199002077
- Ellis, R., Loewen, S., and Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339-368.
- Ellis, R., Sheen, Y., Murakami, M., and Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36, 353-371. doi: 10.1016/j.system.2008.02.001.
- Elola, I., and Oskoz, A. (2016). Supporting second language writing using multimodal feedback. *Foreign Language Annals*, 49, 58-74. doi: 10.1111/flan.12183.
- El-Ebyary, K., and Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, 10(2), 121-142. <https://doi.org/10.6018/ijes/2010/2/119231>
- Ene, E., and Upton, T. (2014). Learner uptake of teacher electronic feedback in ESL composition. *System*. 46. 80-95.
- Erdogan, V. (2005). Contribution of error analysis to foreign language teaching. *Mersin University Journal of the Faculty of Education*, 1(1), 261-270.
- Erel, S. (2007). Error treatment in L2 writing: A comparative study of direct and indirect coded feedback in Turkish EFL context. *Sosyal Bilimler Enstitüsü Dergisi Sayı*, 22(1), 397-415.
- Ericsson, P., and Haswell, R. 2006. Machine scoring of student essays: Truth and consequences. *All USU Press Publications*. 139. [https://digitalcommons.usu.edu/usupress\\_pubs/139](https://digitalcommons.usu.edu/usupress_pubs/139)
- Erlam, R., and Loewen, S. (2010). Implicit and explicit recasts in L2 oral French interaction. *Canadian Modern Language Review*, 66(6), 877-905.
- Evans, C. A. (2013). Making Sense of Assessment Feedback in Higher Education. *Review of Educational Research*, 83, 120-70.
- Evans, N. W., Hartshorn, K. J., McCollum, R. M., and Wolfersberger, M. (2010). Contextualizing corrective feedback in L2 writing pedagogy. *Language Teaching Research*, 14, 445-463.
- Evans, N. W., Hartshorn, K. J., and Strong-Krause, D. (2011). The efficacy of dynamic written

- corrective feedback for university matriculated ESL learners. *System*, 39, 229-239.
- Evans, V., and Tyler, A. (2005). Applying cognitive linguistics to pedagogical grammar: the English prepositions of verticality. *Revista Brasileira de Linguística Aplicada*, 5(2), 11-42. <https://dx.doi.org/10.1590/S1984-63982005000200002>
- Fang, A. C. (2000). A lexicalist approach towards the automatic determination for the syntactic functions of prepositional phrases. *Natural Language Engineering*, 6(2), 183-20.
- Fang, Y. (2010). Perceptions of the computer-assisted writing program among EFL college learners. *Educational Technology & Society*, 13, 246-256. Retrieved August 8, 2021 from <https://citeserx.ist.psu.edu/viewdoc/download?doi=10.1.1.174.1176&rep=rep1&type=pdf>.
- Farrokhi, F., and Sattarpour, S. (2011). The effects of focused and unfocused written corrective feedback on grammatical accuracy of Iranian EFL learners. *Theory and Practice in Language Studies*, 1(12), 1797-1803. doi: 10.4304/tpls.1.12.1797-1803.
- Fathman, A. K., and 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: Cambridge University Press.
- Fazio, L. (2001). The effect of corrections and commentaries on the journal writing accuracy of minority-and majority-language students. *Journal of Second Language Writing*, 10(4), 235-249.
- Ferdouse, F. (2013). Learning from mistakes: Using correction code to improve students' writing skill in English composition class. *Stamford Journal of English*, 7, 62-86.
- Ferris, D. R. (1999). The case of grammar correction in L2 writing classes: A response to Truscott 1996. *Journal of Second Language Writing*, 8, 1-10.
- Ferris, D. R. (2002). *Treatment of error in second language student writing*. Michigan: University of Michigan Press.
- Ferris, D. R. (2003). *Response to student writing: Implications for second language students*. Lawrence Erlbaum Associates: Mahwah, NJ.
- Ferris, D. R. (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. doi: 10.1016/j.jslw.2004.04.005.
- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on 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: Cambridge University Press.
- Ferris, D. R. (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. (2011). *Treatment of error in second language student writing*. Ann Arbor, MI: The University of Michigan Press.
- Ferris, D. R., Chaney, S. J., Komura, K., Roberts, B. J., and McKee, S. (2000). *Perspectives, problems, and practices in treating written error*. Colloquium presented at International TESOL Convention. (March 14-18, 2000).
- Ferris, D. R., and Helt, M. (2000). *Was Truscott right? New evidence on the effects of error correction in L2 writing classes*. Paper presented at the American Association of Applied Linguistics Conference, Vancouver, BC.
- Ferris, D. R., and B. Roberts. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10, 161-184.
- Ferris, D. R., and Hedgcock, J. S. (2014). *Teaching L2 Composition. Purpose, Process, and Practice*. New York: Routledge.
- Foo, C. (2007). An analysis of the preposition errors of time, place and direction among Chinese secondary school students (Master Thesis). Universiti Pendidikan Sultan Idris, Tanjong Malim, Malaysia.

- Fourcin, A. (1975). Language development in the absence of expressive speech. In E. Lenneberg, & E. Lenneberg (Eds.), *Foundations of Language Development* (pp. 263-268). New York: Academic Press.
- Frank, M. (1972). *Modern English: A practical reference guide*. Englewood Cliffs, NJ.: Prentice-Hall.
- Fratzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *The Modern Language Journal*, 79, 329-344.
- Frear, D., and Chiu, Y. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System*, 53, 24-34.
- Fredricks, J. A. (2013). Behavioral engagement in learning. In J. Hattie, & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 42-44). New York: Routledge.
- Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: Potential of the concept. State of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- Fiddienika, A. (2020). Proofing tools in Microsoft Office Word 2019: The application in Indonesian language. *TEKNOSASTIK*, 18, 111-119. doi: 10.33365/ts.v18i2.490.
- García-Gómez, E., and Bou-Franch, P. (1992). Estudio experimental sobre interferencias lingüísticas. In R. Losada-Durán, & M. Mansilla García (Eds.) *Actas del VIII Congreso Nacional de Lingüística Aplicada* (pp. 279-292). Vigo: AESLA.
- García-Mayo, M. P. (2008). The acquisition of four non-generic uses of 'the' by EFL learners. *System*, 36, 550-565.
- García-Mayo, M. P. (2009). Article choice in L2 English by Spanish speakers: Evidence for full transfer. In M. P. Garcia-Mayo, & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications*, (pp. 13-35). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/lald.49.05pil>
- Gass, S. M. (1988). Second language vocabulary acquisition. *Annual Review of Applied Linguistics*, 9, 92-106. doi:10.1017/S0267190500000829.
- Gass, S. M. (1991). Grammar instruction, selective attention, and learning processes. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, and M. Swain (Eds.), *Foreign/second language pedagogy research* (pp. 134-141). Clevedon, UK: Multilingual Matters.
- Gass, S. M. (1997). *Input, interaction, and the second language learner*. Hillsdale, NJ: Erlbaum.
- Gass, S. M. (2001). Sentence matching: A reexamination. *Second Language Research*, 17, 421-441.
- Gass, S. M. (2003). Input and interaction. In C. Doughty and M. Long (Eds.), *The handbook of second language acquisition* (pp. 224-255). Oxford: Blackwell
- Gass, S. M., and Mackey, A. (2006). Input, Interaction and Output: An Overview. *AILA Review*, 19, 3-17. doi:10.1075/aila.19.03gas.
- Gass, S. M., and Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 175-200). London: LEA.
- Gass, S. M., and Selinker, L. (1949). *Second language acquisition: an introductory course*. Hilldale, NJ: Lawrence Erlbaum Associates.
- Gass, S., and Selinker, L. (1992). *Language transfer in language learning*. Amsterdam: John Benjamins.
- Gass, S. M., and Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York: Routledge.
- Gass, S. M., and Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, 16, 283-302. doi:10.1017/S0272263100013097
- Gebhart, A. L., Aslin, R. N., and Newport, E. L. (2009). Changing Structures in Midstream:

- Learning Along the Statistical Garden Path. *Cognitive Science*, 33(6), 1087-1116. <https://doi.org/10.1111/j.1551-6709.2009.01041.x>.
- Gialcone, A. (2002). How do learners acquire the classical three categories of temporality? Evidence from L2 Italian. R. Salaberry, & Y. Shirai (Eds.) *The L2 Acquisition of Tense-Aspect Morphology*. Amsterdam: John Benjamins.
- Gibbs, G., and Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-31.
- Goad, H., and White, L. (2004). Ultimate attainment of L2 inflections: Effects of L1 prosodic structure. In S. Foster-Cohen, M. Ota, M. A. Sharwood Smith, & A. Sorace (Eds.) *EUROSLA Yearbook*, 4 (pp. 119-145). Amsterdam: John Benjamins.
- Goldstein, L. (2004). Questions and answers about teacher written commentary and student revision: Teachers and students working together. *Journal of Second Language Writing*, 13(1), 63-80.
- Goad, H., and White, L. (2009). Articles in Turkish/English interlanguage revisited: Implications of vowel harmony. In M. P. García-Mayo, & R. Hawkins (Eds.) *Second language acquisition of articles: Empirical findings and theoretical implications* (pp 201-232). Amsterdam: John Benjamins.
- Good, T. L., and Brophy, J. (2000). *Looking in classrooms*. New York: Longman.
- Grady, W., Lee, S. O., and Lee, J. H. (2011). Practical and theoretical issues in the study of heritage language acquisition. *Heritage Language Journal*, 8(3) (online).
- Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63, 665-703. doi: 10.1111/lang.12018.
- Gregg, K. R. (1984). Krashen's Monitor and Occam's Razor. *Applied Linguistics*, 5(2), 79-100.
- Gregg, K. R. (2001). Learnability and second language acquisition theory. In P. Robinson (Ed.) *Cognition and second language instruction* (pp. 152-182). Cambridge: Cambridge University Press.
- Greenbaum, S., and Nelson, G. (2002). *An introduction to English grammar*. London: Pearson Education Limited.
- Grimes, D., and Warschauer, M. (2006). *Automated essay scoring in the classroom*. Paper presented at the American Educational Research Association (AERA) Annual Conference, San Francisco, CA.
- Grimes, D., and Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology*, 8(6), 1-44.
- Guardado, M., and Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24, 443-461. doi: 10.1016/j.compcom.2007.03.002.
- Guénette, D. (2013). The pedagogy of error correction: Surviving the written corrective feedback challenge. *TESL Canada Journal*, 30, 117-126.
- Guo, Q. (2015). *The effectiveness of written corrective feedback for L2 development: A mixed method study of written CF types, error categories and proficiency levels*. Unpublished doctoral thesis, AUT University, Auckland, New Zealand.
- Gvarishvili, Z. (2012). Interference of L1 prepositional knowledge in acquiring of prepositional usage in English. *Procedia-Social and Behavioral Sciences*, 70, 1565-1573.
- Hall, G., and Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching*, 45, 271-308. doi: 10.1017/S0261444812000067.
- Halliday, M. A. K. (1994). *Functional Grammar*. London: Edward Arnold.
- Halsey, S. (2007). Embracing emergent technologies and envisioning new ways of using them for literacy learning in the primary classroom. *English Teaching: Practice and Critique*, 6, 99-107.

- Hamouda, A. (2011). A study of students and teachers' preferences and attitudes towards correction of classroom written errors in Saudi EFL context. *English Language Teaching*, 4(3), 128-141.
- Hamp-Lyons, L. (2001). Fourth generation writing assessment. In T. Silva, & P. Matsuda, (Eds.), *On second language writing*. Mahwah: Lawrence Erlbaum Associates.
- Han, J., and Hong, H. (2015). The acquisition problem of English present perfect to Korean adult learners of English. *English Language and Linguistics*, 21(3), 141-164.
- Hanaoka, O., and Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, 21, 332-347. doi: 10.1016/j.jslw.2012.09.008.
- Handley, K., Price, M., and Millar, J. (2011). Beyond 'doing time': Investigating the concept of student engagement with feedback. *Oxford Review of Education*, 37, 543-560. doi: 10.1080/03054985.2011.604951.
- Harmer, J. (2001). *The Practice of English Language Teaching*. Harlow: Longman.
- Harmer, J. (2004). *How to Teach Writing*. London: Longman.
- Hart, B., and Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes Publishing.
- Hartshorn, K. J. (2008). *The effects of manageable corrective feedback on ESL writing accuracy*. Unpublished dissertation, Brigham Young University, Provo, UT. Retrieved May 12, 2018 from <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=2521&context=etd>
- Hartshorn, K. J., and Evans, N. W. (2012). The differential effects of comprehensive corrective feedback on L2 writing accuracy. *Journal of Linguistics and Language Teaching*, 3(2), 217-248.
- Hatch, E. (1983). Simplified input and second language acquisition. In R.W. Andersen (Ed.), *Pidginization and Creolization as Language Acquisition* (pp. 64-88). Rowley, MA: Newbury House.
- Hattie, J., and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>.
- Havranek, G., and Cesnik, H. (2003). Factors affecting the success of corrective feedback. In S. Foster-Cohen, & A. Nizegorodzew (Eds.), *EUROSLA Yearbook* (pp. 99-122). Amsterdam: John Benjamins.
- Hawkins, R. (2009). Second language acquisition of morphosyntax. In W. Ritchie, & T. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 211-236). Bingley: Emerald Group.
- Hawkins, R., Al-Eid S., Almahboob, I., Athanasopoulos, P., Chaengchenkit, R., Hu, J., Rezai, M., Jaensch, C., Jeon, Y., Jiang, A., Leung, I., Matsunaga, K., Ortega, M., Sarko, G., Snape, N., and Velasco-Zárate, K. (2006). Accounting for English article interpretation by L2 speakers. In S. H. Foster-Cohen (Ed.), *EUROSLA Yearbook*, 6, (pp. 7-25). Amsterdam: John Benjamins.
- Hayashi, M. (2001). The acquisition of the prepositions 'in' and 'on' by Japanese learners of English. *JACET Bulletin*, 33, 29-42.
- Hayashi, M. (2008). *Second language acquisition of English prepositions*. Tokyo: Eihosha.
- Heift, T. (2004). Corrective feedback and learner uptake in CALL. *ReCALL*, 16, 416-431. doi: 10.1017/S0958344004001120.
- Heift, T., and Hegelheimer, V. (2017). Computer-assisted corrective feedback and language learning: Research, Theory, Applications, Implications. In H. Nassaji, & E. Kartchava, (Eds.) *Corrective Feedback in Second Language Teaching and Learning: Research, Theory, Applications, Implications*, (pp. 51-65). New York: Routledge.
- Heift, T., and Schulze, M. (2007). *Errors and intelligence in Computer-Assisted Language Learning. Parsers and Pedagogues*. New York: Routledge.
- Herrington, A., and Moran, C. (2001). What happens when machines read our students' writing?

- College English*, 63, 480-499.
- Heydari, P., and Bagheri, M. (2012). Error Analysis: Sources of L2 learners' errors. *Theory and Practice in Language Studies*, 2, 1583-1589. doi: 10.4304/tpls.2.8.1583-1589.
- Hinkel, E. (1992). L2 tense and time reference. *TESOL Quarterly*, 26(3): 556-572.
- Horney, J. E. (1998). *An Error of English Composition Written by Taiwanese students*. Unpublished Doctoral Dissertation, University of Texas, Austin.
- Horwitz, E. K., Horwitz, M. B., and Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132. <https://doi.org/10.2307/327317>
- Housen, A. (2002). The development of tense-aspect in English as a second language and the variable influence of inherent aspect. In R. Salaberry, & S. Yasuhiro (Eds.), *The L2 acquisition of tense-aspect morphology* (pp. 155-197). Amsterdam: John Benjamins.
- Hsin, Ai-li. (2003). Indefinite subject NPs between English and Chinese: An error analysis. Retrieved August 11, 2020 from <https://eric.ed.gov/?id=ED475018>.
- Hsu, A., Charter, N., and Vitányi, P. (2013). Language learning from positive evidence, reconsidered: A simplicity-based approach. *Topics in Cognitive Science*, 5, 35-55.
- Hu, G. (2007). Developing an EAP writing course for Chinese ESL students. *RELC Journal*, 38(1), 67-86.
- Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Michigan: Ann Arbor.
- Hunt, R., and Aslin, R. (2002). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology*, 130, 658-80. doi: 10.1037//0096-3445.130.4.658.
- Hunt, R., and Aslin, R. (2010). Category induction via distributional analysis: Evidence from a serial reaction time task. *Journal of Memory and Language*, 62 2, 98-112.
- Hunter, J. D., (2007). *The question of error*. Unpublished Ph.D. Dissertation, University of Birmingham, UK.
- Hutchison, D. (2007). An evaluation of computerised essay marking for national curriculum assessment in the UK for 11-year-olds. *British Journal of Educational Technology*, 38(6), 977-989.
- Hyland, F. (2003). Focusing on form: student engagement with teacher feedback. *System*, 31, 217-230.
- 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). Amsterdam: John Benjamins.
- Hyland, F., and Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing*, 10(3), 185-212.
- Hyland, K. (2000). *Disciplinary discourses. Social interaction in academic writing*. London: Longman.
- Hyland, K. (2002). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 115-130). London: Longman.
- Hyland, K. (2003). *Second language writing*. Cambridge: Cambridge University Press.
- 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* (pp.17-35). Amsterdam: John Benjamins.
- Hyland, K., and Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101. doi:10.1017/S0261444806003399.
- Ibáñez Moreno, A. (2011). *Secondary school students' errors when writing in English: Analysis of the interaction of age, language, levels and types of errors*. Retrieved August 12, 2020 from [https://www.academia.edu/1214929/SECONDARY\\_SCHOOL\\_STUDENTS\\_ERRORS\\_](https://www.academia.edu/1214929/SECONDARY_SCHOOL_STUDENTS_ERRORS_)

- WHEN WRITING IN ENGLISH Analysis of the interaction of age language level linguistic levels and types of error.
- Ionin, T. (2003). *Article Semantics in Second Language Acquisition*. PhD. dissertation, MIT. Retrieved March 2, 2018.
- Ionin, T., Ko, H., and Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12, 3-69. doi: 10.1207/s15327817la1201\_2.
- Ionin, T., and Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning*, 60, 877-925. doi: 10.1111/j.1467-9922.2010.00577.x.
- Ionin, T., Zubizarreta, M. L., and Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118(4), 554-576. <https://doi.org/10.1016/j.lingua.2006.11.012>.
- Iwashita, N. (2003). Negative feedback and positive evidence in task based interaction: Differential effects on L2 development. *Studies in Second Language Acquisition*, 25, 1-36. doi:10.1017/S0272263103000019.
- Izquierdo, J., and Collins, L. (2008). The facilitative role of L1 influence in tense-aspect marking: A comparison of Hispanophone and Anglophone learners of French. *The Modern Language Journal*, 92, 350-368. doi: 10.1111/j.1540-4781.2008.00751.x.
- Izumi, E., Uchimoto, K., and Isahara, H. (2004). The overview of the SST speech corpus of Japanese learner English and evaluation through the experiment on automatic detection of learners' errors. *Proceedings of Language Resource and Evaluation Conference (LREC)*, (pp. 1435-1438), Portugal.
- 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., Bigelow, M., Fujiwara, M., and 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.
- Izumi, Y., and Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *Canadian Modern Language Review*, 60, 587-609.
- James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. London: Longman.
- James, M. (2007). Interlanguage variation and transfer of learning. *International Review of Applied Linguistics in Language Teaching*, 45, 95-118.
- Jang, J. D. (2003). Korean EFL learners' attitude toward error correction. *The Journal of Linguistic Science*, 26, 345-364.
- Jansson, H. (2006). *Native Swedish Speakers' Problems with English Prepositions* (Bachelor's thesis). Retrieved July 5, 2018 from <http://www.diva-portal.org/smash/search.jsf/get/diva2:137490/FULLTEXT01.pdf>
- Japos, G. V. (2013). Effectiveness of coaching interventions using Grammarly software and plagiarism detection software in reducing grammatical errors and plagiarism of undergraduate research. *JPAIR Institutional Research Journal*, 1(1). Retrieved April 2, 2017 from <http://ejournals.ph/form/cite.php?id=7652>
- Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, 24, 387-418. doi: 10.1017/S0272263102003029.
- Jiang, M. C. (1995). An analysis of Chinese learners' errors in prepositions. *Journal of National Chiayi Institute of Agriculture*, 41, 187-201.
- Jiang, L., and Chen, J. (2013). The effects of written corrective feedback on second language



- acquisition of English generic NPs by Chinese-speaking learners. *Contemporary Foreign Languages Studies*, 11, 31-35.
- Jiménez-Catalán, R. (1996). Frequency and variability in errors in the use of English prepositions. *Miscelánea: A Journal of English and American Studies*, 17, 171-188.
- Jie, X. (2008). Error theories and second language acquisition. *US-China Foreign Language*, 6, 35-42.
- Jones, W. (2012) Assessing students' grammatical ability. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyhoff (Eds.) *The Cambridge Guide to Second Language Assessment* (pp. 247-256). Cambridge: Cambridge University Press.
- Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., and Doughty, C. (1995). Does textual enhancement promote noticing? A think aloud protocol analysis. In R. W. Schmidt (Ed.) *Attention and Awareness in Foreign Language Learning* (pp. 183-216). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Kang, E., and Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *The Modern Language Journal*, 99, 1-18.
- Karpava, S., and Agouraki, G. (2013). L2 acquisition of English present perfect interpretations. *Journal of Portuguese Linguistics*, 12(1), 145-175.
- Karyautry, L. (2018). Grammarly as a tool to improve students' writing quality: Free online-proofreader across the boundaries. *Jurnal Sains Sosial dan Humaniora*, 2, 83-89.
- Katayama, A. (2006). Perceptions of JFL students toward correction of oral errors. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT 2005 Conference Proceedings*. Tokyo: JALT.
- Katayama, A. (2007). Japanese EFL students' preferences toward correction of classroom oral errors. *The Asian EFL Journal Quarterly*, 9, 289-305.
- Kaufman, S., DeYoung, C., Gray, J., Jiminez, L., Brown, J., and MacKintosh, N. (2010). Implicit learning as an ability. *Cognition*, 116, 321-340.
- Kearns, K. (2011) *Semantics*. London: Macmillan.
- Keeler, C. M. (1996). Networked instructional computers in the elementary classroom and their effect on the learning environment: A qualitative evaluation. *Journal of Research on Computing in Education*, 28(3), 329-345. Retrieved August 8, 2021 from <https://www.learntechlib.org/p/80963/>.
- Kennedy, S. (2010). Corrective feedback for learners of varied proficiency levels: A teacher's choices. *TESL Canada Journal*, 27(2), 31-50. <https://doi.org/10.18806/tesl.v27i2.1054>
- Kepner, C. (1991). An experiment in the relationship of types of written feedback to the development of second language writing skills. *The Modern Language Journal*, 75(3), 305-313.
- Keshavarz, M. H. (1999). *Contrastive Analysis & Error Analysis*. Tehran: Rahnama Press.
- Keshavarz, M. H. (2003). *Error Analysis and Contrastive Analysis*. Tehran: Rahnama Publications.
- Khodabandeh, F. (2007). Analysis of students' errors: The case of headlines. *The Asian ESP Journal*, 3(1), 6-21.
- Kim, H. (1987). *An analysis of learners' errors made in their English composition especially at the high school level*. Unpublished master's thesis. Chungbuk National University, Chungju, Korea.
- Kim, I. (1988). *A study of the errors in the use of the English verbs with special reference to tense, mood, and voice*. Unpublished master's thesis. Busan National University, Busan, Korea.
- Kim, S. (2001). An error analysis of college students' writing: Is that really konglish? *Studies in Modern Grammar*, 25, 159-174.
- Kinzer, C. K., and Verhoeven, L. (2008). *Interactive literacy education: Facilitating literacy environments through technology*. New York: Lawrence Erlbaum.
- Kirkgoz, Y. (2010). An analysis of written errors of Turkish adult learners of English. *Procedia. Social and Behavioral Sciences*, 2, 4352-4358. doi: 10.1016/j.sbspro.2010.03.692.

- Kirschner, P. A., van den Brink, H., and Meester, M. (1991). Audio tape feedback for essays in distance education. *Innovative Higher Education*, 15(2), 185-195. doi:10.1007/BF00898030.
- Klein, W. (1994). *Time in language*. London: Routledge.
- Ko, K. (2011). Grammar versus content: KFL/ESL teachers' trends in feedback on college student writing. *The Korean Language in America*, 16, 41-72. Retrieved January 4, 2021, from <http://www.jstor.org/stable/42922289>
- Koffi, E. (2010). *Applied English syntax: Foundations for word, phrase, and sentence analysis*. Dubuque, Iowa: Kendall Hunt Publishing Company.
- Koosha, M., and Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal*, 8, 192-209.
- Kozna, R. B., and Johnston, J. (1991). The technological revolution comes to the classroom. *Change*, 23(1), 10- 22.
- Krashen, S. D. (1978). Individual variation in the use of the Monitor. In W. Ritchie (Ed.) *Principles of Second Language Learning* (pp. 175-183). New York: Academic Press.
- Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon Press.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.
- Krashen, S. D. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.
- Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73, 440-464.
- Krashen, S. D. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-77). London: Academic Press.
- Krashen, S. D. (1998). Comprehensible Output. *System*, 26(2), 175-182.
- Krashen, S. D. (2003). *Explorations in Language Acquisition and Use: The Taipei Lectures*. Portsmouth, NH: Heinemann.
- Krashen, S. D. (2008). Language education: Past, present and future. *RELC Journal*, 39(2), 178-187. <https://doi.org/10.1177/0033688208092183>.
- Lai, C., Fei, F., and Roots, R. (2008). The contingency of recasts and noticing. *The Computer Assisted Language Instruction Consortium Journal*, 26(1), 70-90. Retrieved May 4, 2016 from [https://calico.org/html/article\\_723.pdf](https://calico.org/html/article_723.pdf)
- Lai, Y-H. (2010). Which do students prefer to evaluate their essays: Peers or computer programs? *British Journal of Educational Technology*, 41(3), 432-454.
- Lalande, J. F. (1982). Reducing composition errors: An experiment. *The Modern Language Journal*, 66(2), 140-149.
- Lam, Y. (2009). Applying cognitive linguistics to teaching the Spanish prepositions. *Language Awareness*, 18, 2-18. doi: 10.1080/09658410802147345.
- Leacock, C. (2004). Scoring free-responses automatically: A case study of a large-scale assessment. *Examens*, 1(3).
- Le Bruyn, B. S. W. (2019). Definiteness across languages and in L2 acquisition. In A. Aguilar-Guevara, J. Pozas-Loyo, & V. Vázquez-Rojas (Eds.), *Definiteness across languages* (pp. 201-219). Berlin: Language Science Press. doi:10.5281/zenodo.3265953.
- Lee, C., Wong, K., Cheung, W., and Lee, F. (2009). Web-based essay critiquing system and EFL students' writing: A quantitative and qualitative investigation. *Computer Assisted Language Learning*, 22, 57-72. <http://dx.doi.org/10.1080/09588220802613807>.
- Lee, E. P. (2001). *Error Analysis on medical students' writing*. Retrieved August 10, 2020, from <https://studylib.net/doc/8272198/error-analysis-on-medical-students—writing>
- Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong.

- Journal of Second Language Writing*, 13(4), 285-312.
- Lee, I. (2005). Error correction in the L2 writing classroom: What do students think? *TESL Canada Journal*, 22(2), 1-16.
- Lee, I. (2008). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17, 69-85.
- Lee, L., and O'Rourke, M. (2006). Information and communication technologies: transforming views of literacies in early childhood settings. *Early Years*, 26(1), 49-62.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37-63.
- Leeman, J. (2007). Feedback in L2 learning: Responding to errors during practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 111-137). Cambridge: Cambridge University Press.
- Leki, I. (1990). Coaching from margins: Issues in written response. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom* (pp. 57-68). Cambridge: Cambridge University Press.
- Leki, I. (1991). The preferences of ESL students for error correction in college level writing classes. *Foreign Language Annals*, 24, 203-218.
- Lengo, N. (1995). What is an error? *English Teaching Forum*, 33(3), 20-24.
- Lennon, P. O. (1991). Error: Some problems of definition, identification, and distinction. *Applied Linguistics*, 12, 180-196.
- Leow, R. P. (1997). Attention, awareness, and foreign language behavior. *Language Learning*, 47, 467-505.
- Leow, R. P. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis. *Applied Psycholinguistics*, 19, 133-159.
- Leow, R. P. (2000). A study of the role of awareness in foreign language behavior. *Studies in Second Language Acquisition*, 22, 557-584.
- Levy, M. (1997). *CALL: Context and Conceptualisation*. Oxford: Oxford University Press.
- Levy, M., and Hubbard, P. (2005). Why call CALL "CALL"? *Computer Assisted Language Learning*, 18, 143-149. doi: 10.1080/09588220500208884.
- Lewis, M. (1993). *The lexical approach: The state of ELT and the way forward*. Hove, England: Language Teaching Publications.
- Li, J. R., Link, S., and Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, 27, 1-18.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309-365. doi: 10.1111/j.1467-9922.2010.00561.x.
- Li, S., Ellis, R., and Shu, D. (2016). *The differential effects of immediate and delayed feedback on learners of different proficiency levels*. Retrieved from <https://www.semanticscholar.org/paper/The-differential-effects-of-immediate-and-delayed-Li-Ellis/e842767af04af3592ee27e1f02fcd27b5e33ae48> (14-08-2020)
- Li, Z., Link, S., Ma, H., Yang, H. and Hegelheimer, V. (2014). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, 44, 66-78. doi: 10.1016/j.system.2014.02.007.
- Lieven, E. V., and Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective: In P. Robinson, & N. Ellis (Eds.) *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 168-196). London: Routledge.
- Lightbown, P. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 177-196). Cambridge: Cambridge University Press.

- Lightbown, P., and Spada, N. (1990). Focus-on-Form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition*, 12, 429-448. <http://dx.doi.org/10.1017/S0272263100009517>.
- Lim, J. M.-H. (2007). Crosslinguistic influence versus intralingual interference: A pedagogically motivated investigation into the acquisition of the present perfect. *System*, 35(3), 368-387. <https://doi.org/10.1016/J.SYSTEM.2007.03.002>.
- Lindstromberg, S. (2001). Preposition entries in UK monolingual learners' dictionaries: Problems and possible solutions. *Applied Linguistics*, 22, 79-103. doi: 10.1093/applin/22.1.79.
- Link, S., Dursun, A., Karakaya, K., and Hegelheimer, V. (2014). Towards best ESL practices for implementing automated writing evaluation. *CALICO Journal*, 31(3), 322-344.
- Link, S., Mehrzad, M., and Rahimi, M. (2020). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 1-30. doi:10.1080/09588221.2020.1743323.
- Listiani, L. (2017). Students' perception toward teacher's written corrective feedback in writing 3 class. *4th AECon Proceedings*, doi: <https://dx.doi.org/10.2991/aecon-17.2017.30>. Retrieved November 10, 2017 from <https://www.atlantis-press.com/proceedings/aecon-17/25884235>.
- Littlewood, W., and Yu, B. (2011). First language and target language in the foreign language classroom. *Language Teaching*, 44(1), 64-77. doi: 10.1017/S0261444809990310.
- Liou, H.-C. (1994). Practical considerations for multimedia courseware development: An EFL IVD experience. *CALICO Journal*, 11(3), 47-74.
- Liu, D., and Gleason, J. L., (2002). Acquisition of the article *the* by nonnative speakers of English. An analysis of four nongeneric uses. *Studies in Second Language Acquisition* 24, 1-26.
- Liu, J., and Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2(3), 193-227. [https://doi.org/10.1016/S1475-1585\(03\)00025-0](https://doi.org/10.1016/S1475-1585(03)00025-0)
- Lochtman, K. (2002). Oral corrective feedback in the foreign language classroom: How it affects interaction in analytic foreign language teaching. *International Journal of Educational Research*, 37, 271-283.
- Lococo, V. (1976). A comparison of three methods for the collection of second language data. *Working Papers in Bilingualism*, 8, 99, 59-86.
- Loewen, S. (2004). Uptake in incidental focus on form in meaning-focused ESL lessons. *Language Learning*, 54, 153-187.
- Londoño-Vásquez, D. (2008). Error analysis in a written composition. Profile issues in teachers' professional development. *Profile*, 10, 135-146. <https://dialnet.unirioja.es/descarga/articulo/4858537.pdf>
- Logan, E. (2009). Self and peer assessment in action. *Practitioner Research in Higher Education*, 3(1), 29-35.
- Long, M. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences*, 379, 259-278.
- Long, M. (1982). Native speaker/non-native speaker conversation in the second language classroom. In M. Long, & C. Richards (Eds.), *Methodology in TESOL: A book of readings* (pp. 339-354). New York: Newbury House.
- Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, C. Kramsch and R. Ginsburg (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M. (1996). The role of the linguistic environment on second language acquisition. In W. Ritchie, & T. Bhatia (Eds.) *Handbook of second language acquisition* (pp. 413-468). New York: Academic Press.

- Long, M., Inagaki, S., and Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82, 357-370. doi:10.1111/j.1540-4781.1998.tb01213.x
- Long, M., and Robinson, P. (1998). Focus on form: Theory, research and practice. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (15-41). Cambridge: Cambridge University Press.
- Loschky, L. (1994). Comprehensible input and language acquisition: What is the relationship? *Studies in Second Language Acquisition*, 16(3), 303-323. doi:10.1017/S0272263100013103.
- Lu, R. L., and Bol, L. (2007). A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback. *Journal of Interactive Online Learning*, 6(2), 100-115.
- Lynch, T. (2001). Seeing what they mean: transcribing as a route to noticing. *ELT Journal*, 55(2), 124-132.
- Lyster, R. (1994). La negociation de la forme: Stratégie analytique en classe d'immersion (The Negotiation of Form: Analytical Strategy in Immersion Class). *Canadian Modern Language Review*, 50(3), 446-465.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81.
- Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 51, 265-301. doi: 10.1111/j.1467-1770.2001.tb00019.x.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432.
- Lyster, R., and Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28, 269-300.
- Lyster, R., and Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66.
- Lyster, R. and Saito, K. (2010). Effects of oral feedback in SLA classroom research: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265-302.
- Lyster, R., Saito, K., and Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46, 1-40.
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27, 405-530.
- Mackey, A., Gass, S., and McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition*, 22(4), 471-497.
- Mackey, A., and Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: a series of empirical studies* (pp. 407-453). Oxford: Oxford University Press.
- Mackey, A., and Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338-356.
- Manchón, R. M. (Ed.). (2011). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/llt.31>
- Mangelsdorf, K., and Schlumberger, A. L. (1992). ESL student response stances in a peer-review task. *Journal of Second Language Writing*, 1(3), 235-254.
- Martin, D. (2008). The authors' gallery: A meaningful integration of technology and writing. *Middle School Journal*, 39, 13-17.
- Martínez, J. D. (2008). Linguistic risk-taking and corrective feedback. In J. D. Martínez (Ed.), *Oral Communication in the EFL Classroom* (pp. 165-193). Sevilla: Ediciones Alfar.

- Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system* (Unpublished doctoral dissertation). University of California at Los Angeles. Retrieved February 19, 2016 from <https://www.sjsu.edu/faculty/pmaster/Master%20Dissertation.pdf>.
- Master, P. (1988). *Acquiring the English article system: a cross-linguistic interlanguage analysis*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, Chicago.
- Master, P. (1990). Teaching the English articles as a binary system. *TESOL Quarterly*, 25, 461-478.
- Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, 25, 215-232.
- Master, P. (2002). Information structure and English article pedagogy. *System*, 30, 331-348.
- Matsumura, S., and Hann, G. (2004). Computer anxiety and students' preferred feedback methods in EFL writing. *The Modern Language Journal*, 88, 403-415.
- McCormick, D. C. (2008). Teaching the English present perfect in the second language classroom. University of Toronto PhD Dissertation. Retrieved September 11, 2018 from [https://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/NR39811.PDF?is\\_thesis=1&oclc\\_number=648382555](https://www.collectionscanada.gc.ca/obj/thesescanada/vol2/002/NR39811.PDF?is_thesis=1&oclc_number=648382555).
- McLaughlin, B. (1987). *Theories of Second Language Learning*. London: Edward Arnold.
- Merry, S., and Orsmond, P. (2008). Students' attitudes to and usage of academic feedback provided via audio files. *Bioscience Education*, 11. doi:10.3108/beej.11.3
- Min, H., 2006. The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15, 118-141.
- Moon, Y., and Pae, J. (2011). Short-term effects of automated writing feedback and users' evaluation of criterion. *Korean Journal of Applied Linguistics*, 27, 125-150.
- Morch, A. I., Engeness, I., Cheng, V. C., Cheung, W. K., and Wong, K. C. (2017). EssayCritic: Writing to learn with a knowledge-based design critiquing system. *Educational Technology & Society*, 20, 213-223.
- Morgan-Short, K., and Bowden, H. W. (2006). Processing instruction and meaningful output-based instruction: Effects on second language development. *Studies in Second Language Acquisition*, 28, 31-65.
- Mousavi, K., and Gorjian, B. (2018). Using error correction in teaching oral production to Iranian EFL learners. *Journal of Applied Linguistics and Language Learning*, 4(2), 21-29. doi: 10.5923/j.jall.20180402.01.
- Mueller, P. A., and Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, 25(6), 1159-1168. <https://doi.org/10.1177/0956797614524581>.
- Munns, G., and Woodward, H. (2006). Student engagement and student self-assessment: The REAL framework. *Assessment in Education: Principles, Policy & Practice*, 13, 193-213.
- Muñoz, C. (2011). Is input more significant than starting age in foreign language acquisition? *International Review of Applied Linguistics (IRAL)*, 49(2), 113-133.
- Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning*, 50, 617-673.
- Murtiana, R. (2019). An analysis of interlingual and intralingual errors in EFL learners' composition. *Journal of Educational Studies*, 4, 204-216. doi: 10.30983/educative.v4i2.2544.
- Nabei, T., and Swain, M. (2002). Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness*, 11, 43-63.
- Nadasdi, T., and Sinclair, S. (2007). Anything I can do, CPU can do better: A comparison of human and computer grammar correction for L2 writing using BonPatron.com. Unpublished manuscript. Retrieved January 15, 2018 from <https://sites.ualberta.ca/~tnadasdi/Dublin.htm>.

- Nagy, W. E., Herman, P. A., and Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233-253. <https://doi.org/10.2307/747758>.
- Nash, R.A., and Winstone, N.E. (2017). Responsibility-sharing in the giving and receiving of assessment feedback. *Frontiers in Psychology*, 8 (online).
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59, 411-452.
- Nassaji, H. (2015). Qualitative and descriptive research: Data type versus data analysis. *Language Teaching Research*, 19, 129-132. doi: 10.1177/1362168815572747.
- Nassaji, H. (2017). The effectiveness of extensive versus intensive recasts for learning L2 grammar. *The Modern Language Journal*, 101, 353-368. doi: 10.1111/modl.12387.
- Nassaji, H., and Fotos, S. (2011). *Teaching Grammar in Second Language Classrooms: Integrating form-focused instruction in communicative context*. New York: Routledge.
- Neal, M. (2011). *Writing assessment and the revolution in digital texts and technologies*. New York: Teachers' College Press.
- Netten, J. (1991). Towards a more language oriented second language classroom. In L. Malav & G. Duquette (Eds.), *Language, culture and cognition* (pp. 284-304). Clevedon, UK: Multilingual Matters.
- Nguyen, T. T. M., Do, T. T. H., Nguyen, A. T., and Pham, T. T. T. (2015). Teaching email requests in the academic context: a focus on the role of corrective feedback. *Language Awareness*, 24 (2), 1-27. doi: 10.1080/09658416.2015.1010543.
- Nicholls, D. (2003). The Cambridge learner corpus. Error coding and analysis for lexicography and ELT. In D. Archer, P. Rayson, A. Wilson, & T. McEnery (Eds.), *Proceedings of the Corpus Linguistics 2003 Conference* (pp. 572-581). Lancaster University: University Centre for Computer Corpus Research on Language.
- Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. *Assessment & Evaluation in Higher Education*, 35, 501-517. doi: 10.1080/02602931003786559.
- Nicol, D., and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi: 10.1080/03075070600572090.
- Norris, J., and Ortega, L. (2002). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528. doi: 10.1111/0023-8333.00136.
- Nobuyoshi, J., and Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal* 47, 203-210.
- Oladejo, J. (2005). *Automated evaluation in EFL composition: Do we have all the answers?* Paper presented at the conference of Teaching of Languages, Linguistics, and Literature (TELL). Taipei, NKNU.
- Odlin, T., and Alonso Vázquez, C. (2006). Meanings in search of the perfect form: A look at interlanguage verb phrases. *Rivista di Psicolinguistica Applicata* 6(3), 53-63.
- Olesova, L., Richardson, J., Weasenforth, D., and Meloni, C. (2011). Using asynchronous instructional audio feedback in online environments: A mixed methods study. *MERLOT Journal of Online Learning and Teaching*, 7(1), 30-42. Retrieved from [http://jolt.merlot.org/vol7no1/olesova\\_0311.htm](http://jolt.merlot.org/vol7no1/olesova_0311.htm).
- Oller, J., and Ziahosseiny, S. (2006). The contrastive analysis hypothesis and spelling errors. *Language Learning*, 20, 183-189. 10.1111/j.1467-1770.1970.tb00475.x.
- Oomen-Early, J., Bold, M., Wiginton, K. L., Gallien, T. L., and Anderson, N. (2008). Using asynchronous audio communication (AAC) in the online classroom: A comparative study. *MERLOT Journal of Online Learning and Teaching*, 4(3), 267-276. Retrieved from [http://jolt.merlot.org/vol4no3/oomen-early\\_0311.htm](http://jolt.merlot.org/vol4no3/oomen-early_0311.htm).

- merlot.org/vol4no3/oomenearly\_0908.htm.
- Orrella, J. (2006). Feedback on learning achievement: rhetoric and reality. *Teaching in Higher Education*, 11(4), 441-456.
- Ortega, L. (2009). *Understanding second language acquisition*. New York: Hodder Education.
- Otoshi, J. (2005). An analysis of the use of Criterion in a writing classroom in Japan. *The JALT CALL Journal*, 1(1), 30-38.
- O'Mahony, B., Verezub, E., Dalrymple, J., and Bertone, S. (2013). An evaluation of research students' writing support intervention. *Journal of International Education in Business*, 6, 22-34. <https://doi.org/10.1108/18363261311314935>.
- Paninos, D. (2005). *The role of output for learner attention to input in second language acquisition*. Dissertations available from ProQuest. <https://repository.upenn.edu/dissertations/AAI3179787>.
- Panova, I., and Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly*, 36, 573-595. doi: 10.2307/3588241.
- Park, K. (1996). *The article acquisition in English as a foreign language*. Retrieved September 9, 2020 from <https://files.eric.ed.gov/fulltext/ED397647.pdf>
- Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37, 361-383.
- Pastor, M. D., and Selistean, R. T. (2015). *'If I has a lot of money...': learner errors in foreign language writing / 'Si yo tuvieras mucho dinero...': errores del aprendizaje en la escritura en lengua extranjera*. Retrieved August 11, 2020 from <https://mascvuex.unex.es/revistas/index.php/tejuelo/article/view/1805/1117>.
- Perruchet, P., and Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in cognitive sciences*, 10(5), 233-238. <https://doi.org/10.1016/j.tics.2006.03.006>.
- Philp, J. (2003). Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25, 99-126. 10.1017/S0272263103000044.
- Pica, T., Young, R., and Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21, 737-759.
- Pinker, S. (1996). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- Pindo, A. (2016). Error analysis on the use of prepositions in students' writing. *Journal of English Language Teaching*, 5(2), 1-6.
- Pinto, D., and Rex, S. (2006). The acquisition of the Spanish prepositions por and para in a classroom setting. *Hispania*, 89 (3), 611-622. doi:10.2307/20063366.
- Planas-Lladó, A., Soley, L. F., Fraguell-Sansbelló, R. M., Pujolras, G. A., Planella, J. P., Roura-Pascual, N., and Moreno, L. M. (2013). Student perceptions of peer assessment: An interdisciplinary study. *Assessment & Evaluation in Higher Education*, 39(5), 592-610. doi:10.1080/02602938.2013.860077.
- Polio, C. (2010). *Written error correction and theories of second language acquisition*. Symposium on second language writing (May, 2010). Universidad de Murcia.
- Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of Second Language Writing*, 21, 375-389.
- Polio, C., Fleck, N., and Leder, N. (1998). "If only I had more time": ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7, 43-68.
- Posner, M., and Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25-42.



- Posner, M., and Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A. D. Milner, & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 91-112). Cambridge, MA: MIT Press.
- Potter, R., and Fuller, D. (2008). My new English partner? Using the Grammar Checker in writing instruction. *English Journal*, 98(1), 36-41.
- Powers, D. E., Burstein, J. C., Chodorow, M., Fowles, M. E., and Kukich, K. (2002). Stumping E-rater: Challenging the validity of automated essay scoring. *Computers in Human Behavior*, 18(2), 103-134.
- Puckett, K., Judge, S., and Brozo, W. (2009). Integrating content area literacy and assistive technology: A teacher development institute. *Southeastern Teacher Education Journal*, 2(2), 27-38.
- Qassemzadeh, A., and Soleimani, H. (2016). The impact of feedback provision by Grammarly software and teachers on learning passive structures by Iranian EFL learners. *Theory and Practice in Language Studies*, 6, 1884-1894.
- Quinn, P. (2014). *Delayed versus immediate corrective feedback on orally produced passive errors in English* [Unpublished PhD dissertation, University of Toronto]. Retrieved April 5, 2017 from [https://tspace.library.utoronto.ca/bitstream/1807/65728/1/Quinn\\_Paul\\_201406\\_PhD\\_thesis.pdf](https://tspace.library.utoronto.ca/bitstream/1807/65728/1/Quinn_Paul_201406_PhD_thesis.pdf).
- Ramineni, C. (2013). Validating automated essay scoring for online writing placement. *Assessing Writing*, 18(1), 40-61. Retrieved October 29, 2021 from <https://www.learntechlib.org/p/92343/>.
- Ranalli, J., Link, S., and Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1), 8-25.
- Ranta, L., and Lyster, R. (2007). A cognitive approach to improving immersion students' oral language abilities: The awareness-practice-feedback sequence. In DeKeyser, R. (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 141-160). New York: Cambridge University Press.
- Rassaei, E. (2013). Corrective feedback: Learners' perceptions, and second language development. *System*, 41(2), 472-483.
- Reber, A., and Allen, R. (2000). Individual differences in implicit learning: Implications for the evolution of consciousness. In R. Kunzendorf and B. Wallace (Eds.), *Individual differences in conscious experience*, (pp. 227-250). Amsterdam: John Benjamins.
- Reiners, T., Dreher, C., and Dreher, H. (2011). Six key topics for automated assessment utilisation and acceptance. *Informatics in Education*, 10(1), 47-64.
- Reinders, H. W. (2012). Towards a definition of intake in second language acquisition. *Applied Research on English Language*, 1(2), 15-36.
- Reinholz, D. (2016). The assessment cycle: A model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, 41(2), 301-315.
- Révész, A. J. (2007). *Focus on form in task-based language teaching: Recasts, task complexity, and L2 learning*. Unpublished doctoral dissertation, Columbia University
- Rich, C. (2012). The impact of online automated writing evaluation: A case study from Dalian: *Chinese Journal of Applied Linguistics*, 35(1), 63-79. <https://doi.org/10.1515/cjal-2012-0006>.
- Richards, J. C. (1971). A non-contrastive approach to Error Analysis. *English Language Teaching Journal*, 25, 204-219.
- Richards, J. C. (1974). *Error analysis: Perspectives on second language acquisition*. London: Longman.
- Richards, J. C. (1985). *The context of language teaching*. Cambridge: Cambridge University Press.

- Richards, J. C., and Reppen, R. (2014). Towards a pedagogy of grammar instruction. *RELC Journal*, 45(1), 5-25. <https://doi.org/10.1177/0033688214522622>.
- Richards, J. C., and Rodgers, T. (2001). *Approaches and Methods in Language Teaching*. New York: Cambridge University Press.
- Richards, J. C., Talbot, J., and Platt, H. (1992). *Dictionary of language teaching and applied linguistics*. Essex: Longman.
- Richards, J. C., Tung, P., and Ng, P. (1992). The culture of the English language teacher: a Hong Kong example. *RELC Journal*, 23(1), 81-102.
- Ridha, N. (2012). The effect of EFL learners' mother tongue on their writings in English: An error analysis study. *Journal of the College of Arts*, 60, 22-45.
- Roberts, M. A. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention & awareness in foreign language learning* (pp. 163-182). Honolulu, HI: University of Hawai'i at Manoa.
- Romaine, S. (2003). Variation. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 409-435). Oxford: Blackwell.
- Rouhi, A., and Samaiei, M. (2010). The effects of focused and unfocused indirect feedback on accuracy in EFL writing. *The Social Sciences*, 5(6), 481-485.
- Robb, T., Ross, S., and Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly* 20, 83-93.
- Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, 45, 283-331.
- Robinson, P. (2003). Attention and memory during SLA. In C. Doughty, & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 631-678). Oxford: Blackwell.
- Rolheiser, C., and Ross, J. A. (2000). Student self-evaluation—What do we know? *Orbit*, 30(4), 33-36.
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59, 23-30. doi: 10.1093/elt/cci003.
- Romaine, S. (2003). Variation. In C. Doughty and M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 409-435). Oxford: Blackwell.
- Ross, J. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation*, 11(10), 1-13.
- Rothstein, B. (2007). *Tempus [Tense]*. Heidelberg: Winter.
- Rowe, F. J., and Webb, W. T. (2000). *A guide to the study of the English language*. London: Oxford University Press.
- Rudner, L., Garcia, V., and Welch, C. (2006). An Evaluation of IntelliMetric™ Essay Scoring System. *The Journal of Technology, Learning and Assessment*, 4(4). Retrieved November 20, 2019 from <https://ejournals.bc.edu/index.php/jtla/article/view/1651>.
- Rudner, L., and Liang, T. (2002). Automated Essay Scoring Using Bayes' Theorem. *The Journal of Technology, Learning, and Assessment*, 1(2), 1-22.
- Russell, J., and Spada, N. (2006). The effectiveness of feedback for the acquisition of L2 grammar. In J. D. Norris, & L. Ortega (Eds.) *Synthesizing research on language learning and teaching* (pp. 133-164). Amsterdam: John Benjamins.
- Sachs, R., and Polio, C. (2007). Learners' use of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition*, 29, 67-100.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27-52. [https://doi.org/10.1016/s0010-0277\(98\)00075-4](https://doi.org/10.1016/s0010-0277(98)00075-4).
- Saito, H. (1994). Teachers' practices and students' preferences for feedback on second language

- writing: A case study of adult ESL learners. *TESL Canada Journal*, 11(2), 46-70. <https://doi.org/10.18806/tesl.v11i2.633>.
- Salazar Campillo, P. (2003). An analysis of implicit and explicit feedback on grammatical accuracy. *Miscelánea: A Journal of English and American Studies*, 27, 209-228.
- Santos, M., López-Serrano, S., and Manchón, R. (2010). The differential effects of two types of direct written corrective feedback on noticing and uptake: Reformulation vs. error correction. *International Journal of English Studies*, 10, 131-154.
- Sattayatham, A., and Honsa, S. (2007). Medical students' most frequent errors at Mahidol University, Thailand. *The Asian EFL Journal*, 9(2), 170-194.
- Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language*, 24(1), 139-161. <https://doi.org/10.1017/S030500099600298X>.
- Schachter, J. (1983). A new account of language transfer. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 98-111). Rowley, MA: Newbury House.
- Schachter, J. (1998). Recent research in language learning studies: Promises and problems. *Language Learning*, 48, 557-583.
- Schackne, S. (2002). Second language acquisition research. *Journal of Language and Linguistics*, 1(2), 1-7.
- Scharle, A., and Szabo, A. (2000). *Learner autonomy. A guide to develop learner responsibility*. Cambridge: Cambridge University Press.
- Schmid, L. M., (1999). *The effects of peer response on essay drafts*. Unpublished MA Thesis in TESOL, Department of English, California State University, Sacramento.
- Schmidt, R. (1983). Interaction, acculturation and the acquisition of communicative competence. In N. Wolfson, & E. Judd, (Eds.), *Sociolinguistics and language acquisition* (pp. 137-174). Rowley, MA: Newbury House.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 13, 206-226. Retrieved August 16, 2020 from <http://www.nflrc.hawaii.edu/PDFs/SCHMIDT%20Awareness%20and%20second%20language%20acquisition.pdf>.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.). *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp.1-63). Honolulu, HI: University of Hawaii.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.
- Schmidt, R., and Frota, S. N. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.
- Schulz, R. A. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar teaching and corrective feedback: USA-Colombia. *The Modern Language Journal*, 85, 244-258. <http://dx.doi.org/10.1111/0026-7902.00107>.
- Schunk, D. H. (1996). *Learning Theories: An Educational Perspective*. Englewood Cliffs, N.J.:

- Merrill.
- Schwartz, B. (1993). An alternative account of apparent inaccessibility to UG in L2A. *Newcastle and Durham Working Papers in Linguistics*, 1, 240-250.
- Schwartz, B., and Sprouse, R. (1996). L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research*, 12, 40-72. <http://dx.doi.org/10.1177/026765839601200103>.
- Scovel, T. (2001). *Learning New Languages*. Boston: Heinle and Heinle.
- Selinker, L. (1972). "Interlanguage". *International Review of Applied Linguistics*, 10, 209- 231.
- Semke, H. (1984). The effects of the red pen. *Foreign Language Annals*, 17(3), 195-202.
- Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165-179.
- Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research*, 10, 361-392.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41, 255-283.
- Sheen, Y. (2010). Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition*, 32(2), 203-234. doi:10.1017/S0272263109990507.
- Sheen, Y., Wright, D., and Moldawa, A. (2009). Differential effects of focused and unfocused written correction on the accurate use of grammatical forms by adult ESL learners. *System*, 37, 556-569.
- Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal*, 23(1), 103-110. <https://doi.org/10.1177/003368829202300107>.
- Shermis, M., and Burstein, J. (2003). *Automated Essay Scoring: A cross-disciplinary perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Shi, X. L. (2012). Research on the application of the automatic writing system in college English writing teaching: A case study with Pigai network. *Modern Educational Technology*, 10, 67-71.
- Shintani, N., Ellis, R., and Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103-131. doi: 10.1111/lang.12029.
- Shirai, Y. (1991). *Primacy of aspect in language acquisition: simplified input and prototype*. Unpublished PhD dissertation, University of California, Los Angeles.
- Shirai, Y. (1992). Conditions on transfer: A connectionist approach. *Issues in Applied Linguistics*, 3, 91-120. doi: 10.5070/L431005153.
- Shirai, Y. (2004). A multiple-factor account for the form-meaning connections in the acquisition of tense-aspect morphology. In J. Williams, B. VanPatten, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition*, (pp. 91-112). Mahwah, NJ: Lawrence Erlbaum.
- Shirai, Y., and Andersen, R. (1995). The acquisition of tense-aspect morphology: A prototype account. *Language*, 71, 743-762. doi: 10.2307/415743.
- Simard, D., and Wong, W. (2001). Alertness, orientation, and detection: The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition*, 23(1), 103-124.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sipple, S. (2007). Ideas in practice: Developmental writers' attitudes toward audio and written feedback. *Journal of Developmental Education*, 30(3), 22-31.
- Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skinner, E. A., and Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21-44). New York: Springer. [https://doi.org/10.1007/978-1-4614-2018-7\\_2](https://doi.org/10.1007/978-1-4614-2018-7_2).
- Snape, N. (2005). The use of articles in L2 English by Japanese and Spanish learners. *Essex Graduate*

- Student Papers in Language and Linguistics*, 7, 1-23.
- Snape, N., García-Mayo, M. P., and Gürel, A. (2009). Spanish, Turkish, Japanese and Chinese L2 learners' acquisition of generic reference. In M. Bowles, T. Ionin, S. Montrul, & A. Tremblay (Eds.) *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)* (pp. 1-8). Somerville, MA: Cascadilla Proceedings Project.
- Snape, N., García-Mayo, M. P., and Gürel, A. (2013). L1 transfer in article selection for generic reference by Spanish, Turkish and Japanese L2 learners. *International Journal of Studies*, 13, 1-28.
- Song, M.-J. and Suh, B.-R. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System*, 36, 295-312.
- Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, 30(2), 73-87. doi:10.1017/S0261444800012799.
- Spada, N., Lightbown, P. M., and White, J. (2005). The importance of form/meaning mappings in explicit form-focused instruction. In A. Housen, & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 199-234). Berlin: Mouton de Gruyter.
- Stanley, J., (1992). Coaching student writers to be effective peer editors. *Journal of Second Language Writing*, 1(3), 217-233.
- Stenson, N. (1974). Induced errors. In J. H. Schumann & N. Stenson (Eds.) *New Frontiers in Second Language Learning*. Massachusetts: Newbury House.
- Storch, N. (2010). Critical feedback on written corrective feedback research. *International Journal of English Studies*, 10, 29-46. doi: 10.6018/ijes.10.2.119181.
- Storch, N., and Wigglesworth, G. (2010). Learners processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, 32, 303-334.
- Suh, K.-H. (1992). *A discourse analysis of the English tense-aspect-modality system*. Los Angeles, CA: UCLA dissertation.
- Sundeen, T. (2015). Writing instruction for adolescents in the shadow of the common core state standards. *Journal of Adolescent & Adult Literacy*, 59(2), 197-206. doi: 10.1002/jaal.444.
- Suzanne, N. (2017). An analysis of prepositions errors: The case of in, on and at. *Lingua Didaktika*, 11(1), 13-21.
- Suzuki, M. (2005). Corrective feedback and learner uptake in adult ESL classrooms. *Studies in Applied Linguistics & TESOL*, 10(4), 361-392. doi: 10.7916/salt.v4i2.1592.
- Svalberg, A., and Chuchu, H. (2007). Are English and Malay worlds apart? Typological distance and the learning of tense and aspect concepts. *International Journal of Applied Linguistics*, 8, 27-60. doi: 10.1111/j.1473-4192.1998.tb00120.x.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. M. Gass, & C. Madden (Eds.), *Input in Second Language Acquisition* (pp. 235-256). New York: Newbury House.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review*, 50(1), 158-164.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M., and Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471-483). Mahwah, NJ: Lawrence Erlbaum.
- Swain, M. (2009). We do need methods. In V. Cook, & L. Wei (Eds.), *Contemporary applied*

- linguistics* (pp. 117-136). London: Continuum.
- Tahaineh, Y. (2010). Arab EFL university students' errors in the use of prepositions. *Modern Journal of Applied Linguistics*, 1(6), 76-112.
- Tanimura, M., Takeuchi, K., and Isahara, H. (2004). From learners' corpora to expert knowledge description: Analyzing prepositions in the NICT JLE (Japanese Learner English) Corpus. *IWLel 2004: An Interactive Workshop on Language e-Learning*, 139-147.
- Taylor, R. P. (1980). *The computer in school: Tutor, tool, tutee*. New York: Teachers College Press.
- Terán, V. (2014). *Telicity and the developmental acquisition of the English present perfect by L1 Spanish speakers* (MA Thesis). Retrieved from Southern Illinois University Carbondale, ProQuest Dissertations Publishing (AAT 1567012).
- Terauchi, M. (1994). Keitaiso no syuutoku [The acquisition of morphemes]. In SLA Kenkyuukai (Eds.), *Daini Gengo Syuutoku Kenkyuu ni Motozuku Saisin no Eigo Kyooiku* [The latest English language teaching based on second language acquisition research] (pp. 24-48). Tokyo: Taisyuukan.
- Tetreault, J. R., and Chodorow, M. (2008). The ups and downs of preposition error detection in ESL writing. *COLING*, 865-872.
- Thouësny, S. (2011). *Modeling second language learners' interlanguage and its variability: A computer-based dynamic assessment approach to distinguishing between errors and mistakes*. PhD thesis, Dublin City University. Retrieved February 5, 2021 from <http://doras.dcu.ie/16559/>
- Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10, 335-355. doi: 10.1017/S0142716400008663.
- Thumvichit, A. (2016). A comparative corpus analysis of the use of the present perfect by Thai EFL learners and by native speakers. *Humanities Journal*, 23(1), 239-259.
- Tiene, D., and Luft, P. (2001). Teaching in a technology-rich classroom. *Educational Technology*, 41, 23-31.
- Tomasello, M., and Herron, C. (1988). Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom. *Applied Psycholinguistics*, 9(3), 237-246. <https://doi.org/10.1017/S0142716400007827>.
- Tomlin, R., and Villa, H. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-203.
- Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Optimizing new modes of assessment: In search of qualities and standards* (pp. 55-87). Dordrecht, The Netherlands: Kluwer Academic.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *SLA Research*, 14, 103-135.
- 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, 111-122.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16(4), 255-272.
- Truscott, J. (2009). Arguments and appearances: A response to Chandler. *Journal of Second Language Writing*, 19(1), 59-60.
- Truscott, J., and Hsu, A. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, 17(4), 292-305. 10.1016/j.jslw.2008.05.003.
- Truscott, J., and Sharwood Smith, M. (2011). Input, intake, and consciousness. *Studies in Second Language Acquisition*, 33, 497-528. doi: 10.1017/S0272263111000295.
- Tryzna, M. (2009). Questioning the validity of the Article Choice Parameter and Fluctuation

- Hypothesis: Evidence from L2 English article use by L1 Polish and L1 Mandarin Chinese Speakers. In M. P. García-Mayo, & R. Hawkins (Eds.) *Second language acquisition of articles* (pp. 67-86). Amsterdam: John Benjamins.
- Tuck, J. (2012). Feedback-giving as social practice: Teachers' perspectives on feedback as institutional requirement, work and dialogue. *Teaching in Higher Education*, 17, 209-221. 10.1080/13562517.2011.611870.
- Tuzi, F. (2004). The impact of e-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition*, 21(2), 217-235. doi: 10.1016/S8755-4615(04)00008-8.
- Uno, M. (2014). Lexical aspect in the use of the present perfect by Japanese EFL learners. *International Review of Applied Linguistics in Language Teaching*, 52, 31-57.
- van Leeuwen, C. A., and Gabriel, M. A. (2007). Beginning to write with word processing: Integrating writing process and technology in a primary classroom. *International Reading Association*, 60(5), 420-429.
- Vandergrift, L. (1986). Second language writing and correction: Toward an improved model for composition correction. *The Canadian Modern Language Review*, 42, 3, 658-667.
- VanPatten, B. (2003). *Processing Instruction: Theory, Research and Commentary*. Mahwah, NJ: Lawrence Erlbaum.
- VanPatten, B. (Ed.). (2004). *Processing instruction: Theory, research, and commentary*. Mahwah, N. J.: Lawrence Erlbaum Associates.
- VanPatten, B., and Benati, A. G. (2010). *Key terms in second language acquisition*. London: Continuum.
- VanPatten, B., and Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15, 225-241. <http://dx.doi.org/10.1017/S0272263100011979>.
- VanPatten, W., and Williams, J. (2007). *Theories in second language acquisition: An introduction*. New York: Lawrence Erlbaum Associates.
- 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. <https://doi.org/10.6018/ijes/2010/2/119171>.
- Van Beuningen, C., de Jong, N., and Kuiken, F. (2008). The effect of direct and indirect corrective feedback on L2 learner's written accuracy. *International Journal of Applied Linguistics*, 156, 279-296. doi: 10.2143/ITL.156.0.2034439.
- Van Beuningen, C., de Jong, N., and Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error correction in second language writing. *Language Learning*, 62, 1-41. doi: 10.1111/j.1467-9922.2011.00674.x.
- Van Lier, L. (1996). *Interaction in the Language Curriculum*. London: Longman.
- Vershinskaya, O. (2014). *Reading and writing on paper and on the screen: Russian case*. Report from Russian Academy of Sciences for COST Action FP1104 WG1.
- Vincent, J. (2016). Students' use of paper and pen versus digital media in university environments for writing and reading –a cross-cultural exploration. *Journal of Print Media and Media Technology Research*, 5(2), 97-106.
- Vyatkina, N. (2010). The effectiveness of written corrective feedback in teaching beginning German. *Foreign Language Annals*, 43, 671-689. doi: 10.1111/j.1944-9720.2010.01108.x.
- Walker, J. L. (1973). Opinions of university students about language teaching. *Foreign Language Annals*, 7, 102-105.
- Wang, Q., and Castro, C. D. (2010). Classroom interaction and language output. *English Language Teaching*, 3(2), 175-186.
- Wang, Y., and Liu, Z. Q. (2012). The effects of teacher feedback on accuracy, fluency, complexity and overall quality of English writing. *Foreign Language Education*, 6, 49-53.

- Wang, Y., Shang, H. and Briody, P. (2012). Exploring the impact of using automated writing evaluation in English as a foreign language university students' writing. *Computer Assisted Language Learning*, 26, 1-24. doi: 10.1080/09588221.2012.655300.
- Wardhaugh, R. (1970). The Contrastive Analysis Hypothesis. *TESOL Quarterly*, 4(2), 123-130. <https://doi.org/10.2307/3586182>.
- Ware, P. (2005). Automated writing evaluation as a pedagogical tool for writing assessment. In A. Pandian, G. Chakravarthy, P. Kell, & S. Kaur (Eds.), *Strategies and practices for improving learning literacy* (pp. 174-184). Selangor, Malaysia: Universiti Putra Malaysia Press.
- Ware, P. (2011). Computer-generated feedback on student writing. *TESOL Quarterly*, 45(4), 769-774.
- Warden, C., and Chen, J. (1995). Improving feedback while decreasing teacher burden in ROC ESL Business English writing classes. In P. Bruthiaux, T. Boswood, & B. Du-Babcock (Eds.), *Explorations in English for Professional Communications* (pp. 125-137). Hong Kong: City University of Hong Kong.
- Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.) *Multimedia language teaching* (pp. 3-20). Tokyo: Logos International.
- Warschauer, M. (1999). CALL vs. electronic literacy: Reconceiving technology in the language classroom. Retrieved December 2, 2017, from <http://www.cilt.org.uk/research/resfor2/warsum1.htm>.
- Warschauer, M., and Ware, P. (2006). Automated writing evaluation: defining the classroom research agenda. *Language Teaching Research*, 10(2), 1-24.
- Wacyn-Jones, P. and Allsop, J. (1990). *Test Your Prepositions*. Harlow: Penguin Books. Retrieved January 15, 2021 from [https://tuxdoc.com/download/test-your-peter-wacyn-jones-jake-allso-peter-wacyn-jones-test-your-prepositio\\_pdf](https://tuxdoc.com/download/test-your-peter-wacyn-jones-jake-allso-peter-wacyn-jones-test-your-prepositio_pdf).
- Wei, M. (2015). An empirical study into the effects of feedback on college English writing. *Journal of Tianjin Foreign Studies University*, 22, 43-50.
- White, J. (1996). *An input enhancement study with ESL children: Effects on the acquisition of possessive determiners*. Unpublished doctoral dissertation, McGill University, Montreal.
- White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty, & J. Williams (Eds.), *Focus on form in second language classroom acquisition* (pp. 91-128). New York: Cambridge University Press.
- White, L. (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied Linguistics*, 8, 95-110.
- White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research*, 7(2), 133-161. Retrieved January 29, 2021, from <http://www.jstor.org/stable/43104427>.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- White, L. (2008). Different? Yes. Fundamentally? No. Definiteness effects in the L2 English of Mandarin speakers. In R. Slabakova, J. Rothman, P. Kempchinsky & E. Gavrusseva (Eds.), *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)* (pp. 251-261). Somerville, MA: Cascadilla Proceedings Project.
- Wilkins, D. (1972). *Linguistics in Language Teaching*. London: Edward Arnold.
- Williams, C., Ritchie, and Taj, K. B. (1999). *Handbook of child language acquisition*. San Diego, CA: Academic Press.
- Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook on research in second language teaching and learning* (pp. 673-691). Mahwah, NJ: Lawrence Erlbaum.
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of*



- Second Language Writing*, 21, 321-331. doi: 10.1016/j.jslw.2012.09.007.
- Williams, M., and Burden, R. (1997). *Psychology for Language Teachers: A Social Constructivist Approach*. Cambridge: Cambridge University Press.
- Winstone, N. E., and Carless, D. (2019). *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. London: Routledge. doi: 10.4324/9781351115940.
- Winstone, N. E., Nash, R. A., Parker, M., and Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: a systematic review and a taxonomy of recipience processes. *Educational Psychologist* 52, 17-37. doi: 10.1080/00461520.2016.1207538.
- Wolfe, E., and Manalo, J. (2004). Composition medium comparability in a direct writing assessment of non-native English speakers. *Language Learning & Technology*, 8(1), 53-65.
- Wong, W. (2005). *Input enhancement: From theory and research to the classroom*. New York: McGraw-Hill.
- Wood, K. A., Moskovitz, C., and Valiga, T. M. (2011). Audio feedback for student writing in online nursing courses: Exploring student and instructor reactions. *Journal of Nursing Education*, 50(9), 540-543. doi:10.3928/01484834-20110616-04.
- Woroniecka, I. (1998). *A nonnative student's reactions to instructors' feedback on his papers: A case study of an undergraduate history student* (Master's thesis). University of Toronto. Department of Curriculum, Teaching, and Learning: Ontario. Retrieved May 13th, 2019 from [http://www.nlcnc.ca/obj/s4/f2/dsk1/tape11/PQDD\\_0004/MQ40680.pdf](http://www.nlcnc.ca/obj/s4/f2/dsk1/tape11/PQDD_0004/MQ40680.pdf)
- Yamaoka, T. (1995). A prototype analysis of the learning of 'on' by Japanese learners of English and the potentiality of prototype contrastive analysis (Part I). *Hyogo University of Teacher Educational Journal*, 15(2), 51-59.
- Yamaoka, T. (1995). A prototype analysis of the learning of 'on' by Japanese learners of English and the potentiality of prototype contrastive Analysis (Part II). *Hyogo University of Teacher Educational Journal*, 16(2), 43-49.
- Yang, N. D. (2004). Using MyAccess in EFL writing. *Proceedings of 2004 International Conference and Workshop on TEFL & Applied Linguistics*, (pp. 550-564). Taipei, Ming Chuan University.
- Yang, X. Q., and Dai, Y. C. (2015). Practice research on self-regulated writing model of college English teaching based on Pigai network. *Computer-Assisted Foreign Language Education*, 2, 17-23.
- Yang, Y., and Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32, 235-263. doi: 10.1017/S0272263109990519.
- Yeh, Y. L., and Yu, Y. T. (2004). Computerized feedback and bilingual concordancer for EFL college students' writing. *Proceedings of the 21st International Conference on English Teaching and Learning in the Republic of China* (pp. 35-48). Taichung, Taiwan: Chaoyang University of Technology.
- Ying, S. Y. (1987). *Types of errors in English compositions by Chinese students: A research for interlanguage*. Unpublished master's thesis, Fu Jen Catholic University, Taipei, Taiwan.
- Yildiz, M. (2015). Contrastive analysis of Turkish and English in Turkish EFL learners' spoken discourse. *International Journal of English Studies*, 16, 57-74 doi: 10.6018/ijes/2016/1/212631.
- Ying, H. (1995). What sort of input is needed for intake? *International Review of Applied Linguistics*, 33(3), 175-194. doi: 10.1515/iral.1995.33.3.175.
- Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, 75, 426-439.
- Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 135-175). Amsterdam: John Benjamins. <https://doi.org/10.1075/sibil.10.07you>.

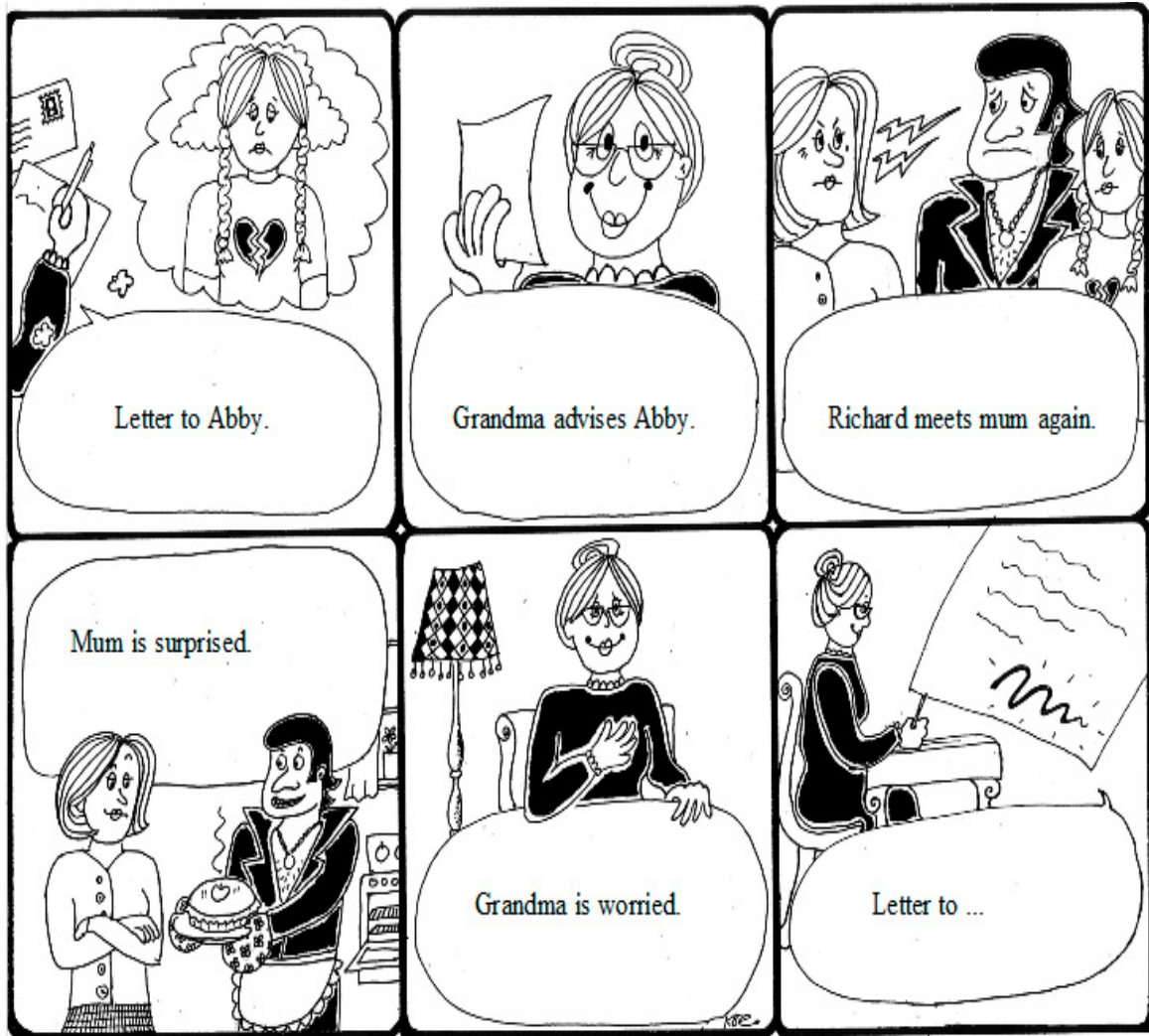
- Yorke, M., and Longden, B. (2004). *Retention and student success in higher education*. Maidenhead, UK: Open University Press.
- Yousefi, M. H. (2011). Interactional feedback, task-based interaction and learner uptake. *Contemporary Online Language Education Journal*, 1(1), 1-19.
- Yu, Y. T., and Yeh, Y. L. (2003). Computerized feedback and bilingual concordancer for EFL college students' writing. *Proceedings of 2003 International Conference on English Teaching and Learning in the Republic of China*, (pp. 35-48). Taipei, Crane.
- Zalbidea, J. (2021). On the scope of output in SLA: ask modality, salience, L2 grammar noticing, and development. *Studies in Second Language Acquisition*, 43(1), 50-82. doi:10.1017/S0272263120000261.
- Zdorenko, T., and Paradis, J. (2007). *The role of the first language in child second language acquisition of articles*. Retrieved February 5, 2021 from <http://www.lingref.com/cpp/galana/2/paper1587.pdf>
- Zdorenko, T., and Paradis, J. (2008). The acquisition of articles in child L2 English: Fluctuation, transfer, or both? *Second Language Research*, 24, 227-250.
- Zdorenko, T., and Paradis, J. (2012). Articles in child L2 English: When L1 and L2 acquisition meet at the interface. *First Language*, 32, 38-62. doi: 10.1177/0142723710396797.
- Zhan, L. (2016). Written teacher feedback: Student perceptions, teacher perceptions, and actual teacher performance. *English Language Teaching*, 9, 73-84. doi: 10.5539/elt.v9n8p73.
- Zhou, Y. S. (2013). A comparative study of teacher feedback, peer feedback and online feedback in English writing. *Foreign Language World*, 3, 87-96.
- Zohrani, K., and Ehsani, F. (2014). The role of implicit and explicit corrective feedback in Persian speaking EFL learners' awareness of and accuracy in English grammar. *Procedia-Social and Behavioral Sciences*, 98, 2018-2024.
- Zurniati, V. (2019). An analysis of the students' preposition errors in writing recount text at English Department in academic year 2013/2014 at IAIN Imam Bonjol Padang. *International Journal of Educational Dynamics*, 1, 55-68. doi: 10.24036/ijeds.v1i1.45.

# Appendices

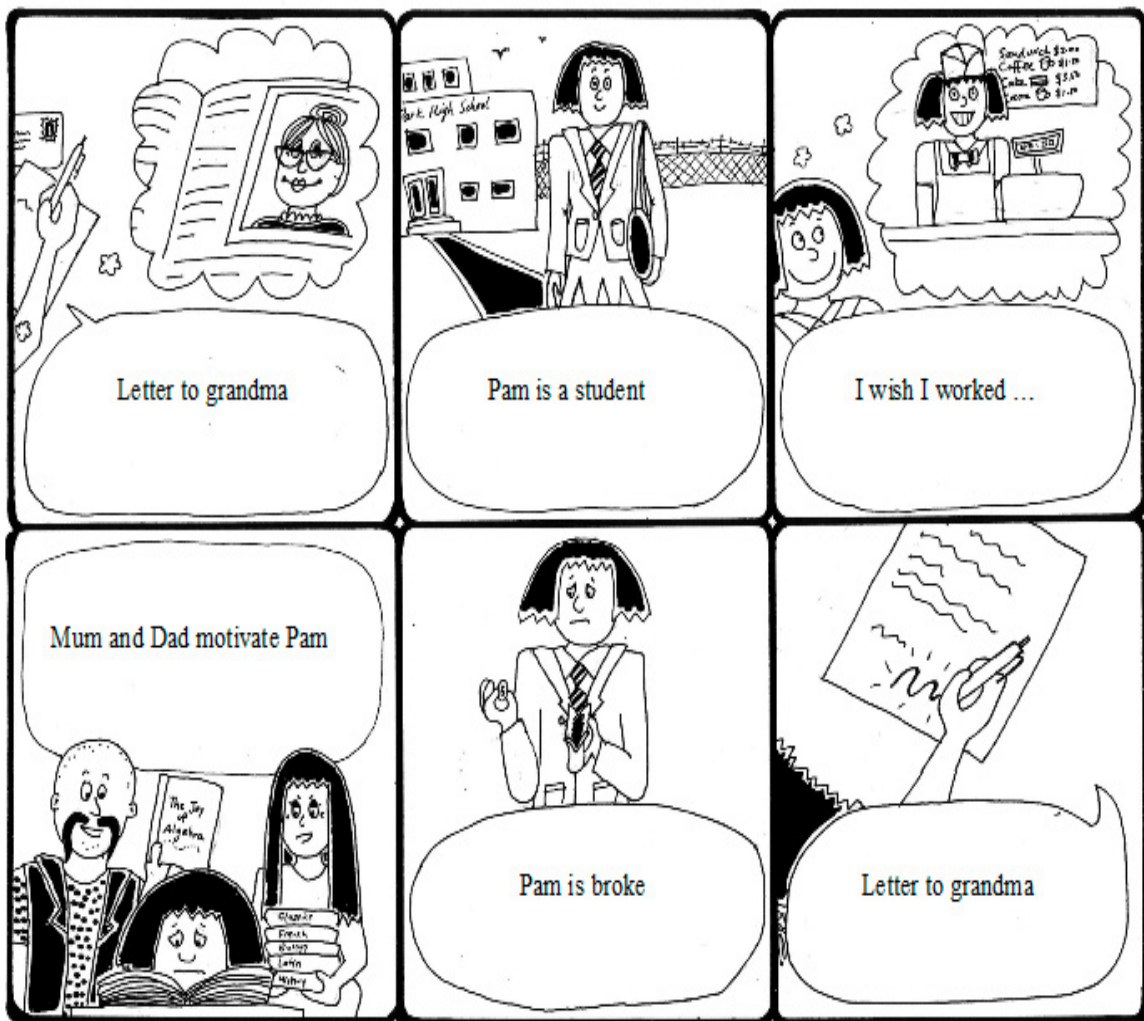
## Appendix 1: Comic strips

Dear Abby in Love (*ABBY*) extracted from [https://sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/picture\\_cues/](https://sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/picture_cues/)





Abby's Part-time job (PAM) extracted from [https://sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/picture\\_cues/](https://sites.educ.ualberta.ca/staff/olenka.bilash/Best%20of%20Bilash/picture_cues/)





## Appendix 2: Questionnaire

### Personal and professional information

1. Student name:
2. Nationality:
3. Age:

If you do not wish to tell your age, are you:

- under 20
- between 21 and 30
- between 31 and 40
- between 41 and 50
- between 51 and 60
- 61 or over?

4. Mother tongue/s:
5. Other spoken language/s spoken (please specify level: poor, average, good, very good)
6. How long have you studied English?
7. Why do you study English?
8. Have you ever been to an English-speaking country?

- a. Where and how long?
- b. What did you do exactly? • Erasmus • Work • Internship • Holiday • Other (please specify):

9. Do you have any previous studies?
10. What did you study or what are you currently studying?

### Language skills and feedback

11. How good are you at writing texts (1 = excellent; 5 = poor)?
12. With what material do you prefer to write your essays?
  - a. Paper and pencil
  - b. The computer
  - c. Both
13. If there were many errors in your writing, what do you think your English teacher should do? You can check more than one!
  - a. Teacher should mark all errors.
  - b. Teacher should mark all major but not the minor ones.

- c. Teacher should mark most of the major errors, but not necessarily all of them.
- d. Teacher should mark only the errors that interfere with communicating your ideas.
- e. Teacher should mark no errors and respond only to the ideas and content.

14. If there are many different errors in your written work, which type(s) of error do you want your English teacher to point out most? Circle one number that best describes each statement.

1=not useful at all (useless) | 2=not useful | 3=doesn't matter | 4=quite useful | 5=very useful

|   |   |   |   |   |   |
|---|---|---|---|---|---|
| A. Teacher points out organisation errors.<br>(Example: paragraph, structure, sentence order) | 1 | 2 | 3 | 4 | 5 |
| B. Teacher points out grammatical errors.<br>(Example: tense, word order, sentence structure) | 1 | 2 | 3 | 4 | 5 |
| C. Teacher points out content/idea errors.<br>(Example: comments on your ideas)               | 1 | 2 | 3 | 4 | 5 |
| D. Teacher points out punctuation errors.   | 1 | 2 | 3 | 4 | 5 |
| E. Teacher points out spelling errors.<br>(Example: word is spelled wrong)                    | 1 | 2 | 3 | 4 | 5 |
| F. Teacher points out vocabulary errors.<br>(Example: wrong choice, wrong meaning)            | 1 | 2 | 3 | 4 | 5 |
| G. Other: _____   |   |   |   |   |   |

15. How would you describe your attitude towards instructor-feedback in the writing course?

- a. Positive
- b. Negative
- c. Neutral
- d. Not Sure

16. Please rate your understanding of your instructor's comments on the following areas:

|  | Poor | Fair | Average | Good | Excellent |
|--|------|------|---------|------|-----------|
| 1) Grammatical issues in your paper    |      |      |         |      |           |
| 2) Organisational issues in your paper |      |      |         |      |           |
| 3) Content issues in your paper        |      |      |         |      |           |



17. Please rate how challenging the technology was in obtaining your comments to your papers. Please circle one number.

| 1                                 | 2                                 | 3                                | 4   | 5                               |
|-----------------------------------|-----------------------------------|----------------------------------|---|---------------------------------|
| I had many technological problems | I had some technological problems | I had few technological problems | I had very few technological problems that were easily resolved | I had no technological problems |

18. How would you describe your attitudes towards feedback BEFORE the use of computers?

19. How would you describe your attitudes towards feedback AFTER the use of computers?

20. What is your opinion about the use of Grammar Checker in the assessment of writing?

Explain in detail how the use of computerized-feedback managed to or failed to help you?

## APPENDIX 3: Examples of GC and teacher feedback

Example of a complete set of writings of a student in Group 1 (computer feedback).

### *ABBY 1st DRAFT*

Initial text

The screenshot shows a web-based Grammar Checker interface. At the top, there is a header with the text "grammar checker" and a yellow highlighter icon. Below the header, there are several navigation buttons: "+ New text", "Spelling", "Incorrect sequences", "Problem words", "Pairs filter", and "Summary". The main content area is titled "Enter your text:" and contains the following text:

ABBY IN LOVE  
 There was a boy called Richard of whom Abby was in love with. Before summer, Abby knew Richard by sight until she attended at her cousin's birthday where Richard was too. Abby fell in love with Richard the moment she met him. Since this day, they were going out together. During the summer, the relation was going very well, until Abby's mum met Richard by casualty. It was a Saturday morning, when the couple bumb into Abby's mum. The mother didn't like Richard at once as she didn't know anything about her daughter relation. Mum was very annoyed. Richard seemed to her to be a great cocky because of their appearance. Therefore, Abby was not allowed to go out with Richard, and Abby was extremely heartbroken. For the next month Abby felt very miserable. She missed Richard and she became more and more depressed; so, she decided to ask her grandma for advice. The girl described Richard as a loyal person. In a few days, grandma recommended her to provoque a new meeting with Richard and her mother, but Richard would have to go to the Abby's home with a cake made by him. Abby made in this way. The meeting was a success. But, on the other hand, grandma was thoughtful and worried. "If he was not suitable for her granddaughter?" So, asked Abby for a meeting with Richard.

On the right side of the interface, there is a yellow "Instructions" button. Below it, the text reads: "Find out what Grammar Checker does and how to use it:" followed by a link: "o Guide [\(English\)](#) [\(Spanish\)](#)".

## Spelling

+ New text
• Spelling
≡ Incorrect sequences
--? Problem words
•• Pairs filter
☰ Summary

### Result of the spelling filter:

**ABBY** IN LOVE There was a boy called Richard of whom Abby was in love with. Before summer, Abby knew Richard by sight until she attended at her cousin's birthday where Richard was too. **Abby** fell in love with Richard the moment she met him. Since this day, they were going out together. During the summer, the relation was going very well, until Abby's mum met Richard by casualty. It was a Saturday morning, when the couple **bumb** into Abby's mum. The mother didn't like Richard at once as she didn't know anything about her daughter relation. Mum was very annoyed. Richard seemed to her to be a great cocky because of their appearance. Therefore, Abby was not allowed to go out with Richard, and Abby was extremely heartbroken. For the next month Abby felt very miserable. She missed Richard and she became more and more depressed; so, she decided to ask her grandma for advice. The girl described Richard as a loyal person. In a few days, grandma recommended her to **provoque** a new meeting with Richard and her mother, but Richard would have to go to the Abby's home with a cake made by him. **Abby** made in this way. The meeting was a success. But, on the other hand, grandma was thoughtful and worried. If he was not suitable for her granddaughter? So, asked Abby for a meeting with Richard.

✎ Modify

**How does it work?**

The **spelling filter** analyzes the frequency of each word you used. Then it tells you if some of your words are infrequent or very infrequent.

Legend:

- Yellow:** spelling mistakes with pedagogic feedback.
- Purple:** words that are not amongst the 90, 000 most frequent English words.

Click on them to get more info.

## Incorrect Sequences

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
·· Pairs filter
☰ Summary

### Incorrect sequences filter

ABBY IN LOVE There was a boy called Richard of whom Abby was in love with. Before summer, Abby knew Richard by sight until she attended at her cousin's birthday where Richard was too. Abby fell in love with Richard the moment she met him. Since this day, they were going out together. During the summer, the relation was going very well, until Abby's mum met Richard by casualty. It was a Saturday morning, when the couple bump into Abby's mum. The mother didn't like Richard at once as she didn't know anything about her daughter relation. Mum was very annoyed. Richard seemed to her to be a great cocky because of their appearance. Therefore, Abby was not allowed to go out with Richard, and Abby was extremely heartbroken. For **the next month** Abby felt very miserable. she missed Richard and she became more and more depressed; so, she decided to ask her grandma for advice. The girl described Richard as a loyal person. In a few days, grandma recommended her to provoke a new meeting with Richard and her mother, but Richard would have to go to the Abby's home with a cake made by him. Abby made in this way. the meeting was a success. But, on the other hand, grandma was thoughtful and worried. If he was not suitable for her granddaughter? So, asked Abby for a meeting with Richard.

✎ Modify

**How does it work?**

The **incorrect sequences filter** compares your text with one million of the most frequent incorrect sequences in English.

This filter will give you personalized feedback to help you improve your text.

## Problem words

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
·· Pairs filter
☰ Summary

### Problem words filter

ABBY IN LOVE There was a boy called Richard of whom Abby was in love with. Before summer, Abby knew Richard by sight until she attended at her cousin's birthday where Richard was too. Abby fell in love with Richard the moment she met him. Since this day, they were going out together. During the summer, the relation was going very well, until Abby's mum met Richard by casualty. It was a Saturday morning, when the couple bump into Abby's mum. The mother didn't like Richard at once as she didn't know anything about her daughter relation. Mum was very annoyed. Richard seemed to her to be a great cocky because of their appearance. Therefore, Abby was not allowed to go out with Richard, and Abby was extremely heartbroken. For next month Abby felt very miserable. She missed Richard and she became more and more depressed; so, she decided to ask her grandma for advice. The girl described Richard as a loyal person. In a few days, grandma recommended her to provoke a new meeting with Richard and her mother, but Richard would have to go to the Abby's home with a cake made by him. Abby made in this way. The meeting was a success. But, on the other hand, grandma was thoughtful and worried. If he was not suitable for her granddaughter? So, asked Abby for a meeting with Richard.

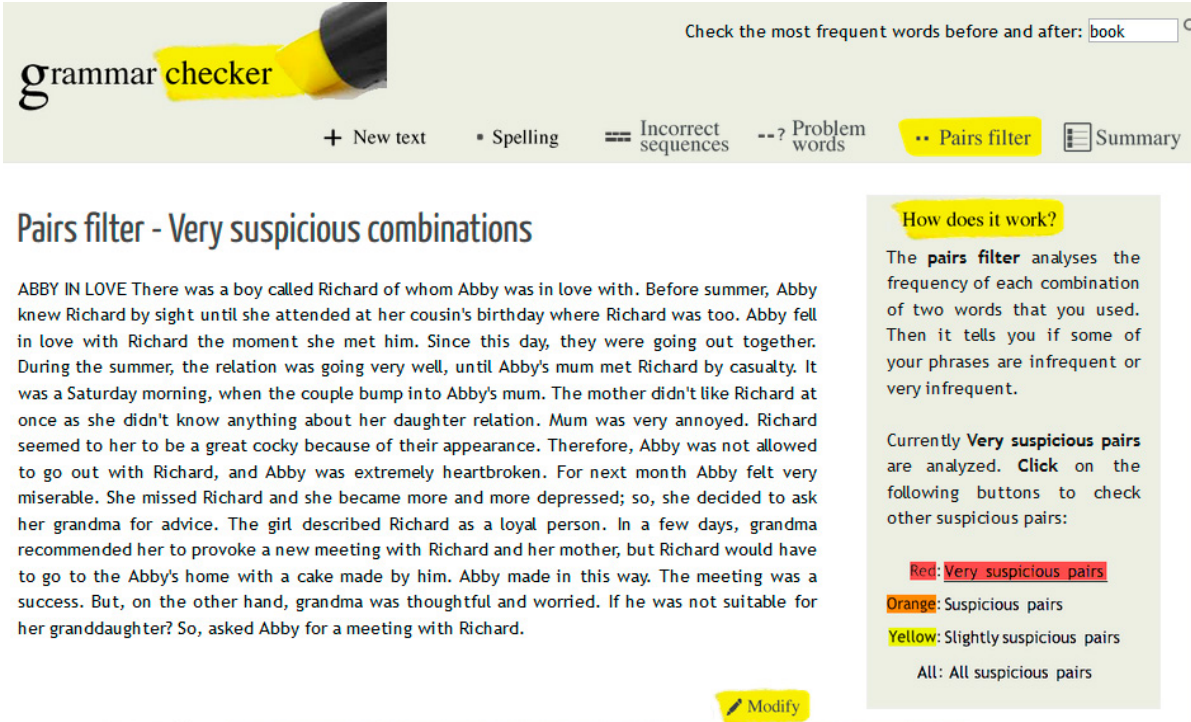
✎ Modify

**How does it work?**

The **problem words filter** shows at the bottom of the screen the words associated with error - words that many students of English use incorrectly.

Grammar Checker does not know if you have used them correctly or incorrectly. If you are not certain, click on the word and Grammar Checker will help you decide.

## Pairs filter and final version



Grammar checker

Check the most frequent words before and after:

+ New text • Spelling == Incorrect sequences --? Problem words **.. Pairs filter** ☰ Summary

### Pairs filter - Very suspicious combinations

ABBY IN LOVE There was a boy called Richard of whom Abby was in love with. Before summer, Abby knew Richard by sight until she attended at her cousin's birthday where Richard was too. Abby fell in love with Richard the moment she met him. Since this day, they were going out together. During the summer, the relation was going very well, until Abby's mum met Richard by casualty. It was a Saturday morning, when the couple bump into Abby's mum. The mother didn't like Richard at once as she didn't know anything about her daughter relation. Mum was very annoyed. Richard seemed to her to be a great cocky because of their appearance. Therefore, Abby was not allowed to go out with Richard, and Abby was extremely heartbroken. For next month Abby felt very miserable. She missed Richard and she became more and more depressed; so, she decided to ask her grandma for advice. The girl described Richard as a loyal person. In a few days, grandma recommended her to provoke a new meeting with Richard and her mother, but Richard would have to go to the Abby's home with a cake made by him. Abby made in this way. The meeting was a success. But, on the other hand, grandma was thoughtful and worried. If he was not suitable for her granddaughter? So, asked Abby for a meeting with Richard.

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

**Red:** **Very suspicious pairs**

**Orange:** Suspicious pairs

**Yellow:** Slightly suspicious pairs

**All:** All suspicious pairs

Modify

*ABBY 2nd DRAFT*

Initial text

The screenshot shows a web-based Grammar Checker interface. At the top, there is a navigation bar with several options: a yellow button labeled '+ New text', a 'Spelling' option with a square icon, an 'Incorrect sequences' option with a double equals icon, a 'Problem words' option with a double dash and question mark icon, a 'Pairs filter' option with a double dot icon, and a 'Summary' option with a list icon. Below the navigation bar, the main area is titled 'Enter your text:'. It contains three paragraphs of text. The first paragraph describes Abby's love for Richard and her mother's disapproval. The second paragraph describes the advice Abby's grandmother gave her. The third paragraph describes the successful outcome of following the advice. To the right of the text area is a sidebar with a yellow header 'Instructions'. Below the header, it says 'Find out what Grammar Checker does and how to use it:' followed by a link 'Guide (English) (Spanish)'. At the bottom right of the text area, there is a yellow button labeled 'Start' with a right-pointing triangle icon.

+ New text   Spelling   Incorrect sequences   Problem words   Pairs filter   Summary

### Enter your text:

Once upon a time there was a girl who was in love with a boy with a bad taste, Richard. As she was too young, she hadn't said to her mother anything about her relation in the long time she was loving Richard because she knew her mother wouldn't like him. But as their relation was going more and more well, she decided to introduce Richard at the family and, of course, the mother didn't like Richard since the moment she saw him. The worst thing was that Abby was not allowed to follow with the relation. So, Abby decided to write a letter to grandma telling her about the situation that made her feel very sad.

The answer of the grandma didn't make Abby wait. Grandma gave some advice to the girl that couldn't go wrong. The piece of advice consisted in preparing a get-together with her mother again but, instead of carrying flowers to the mum, Richard should cook a cake which receipt that was attached in the answer letter. So, Abby proceed following the piece of advice and the meeting was made. It was a successful get-together, the delicious cake surprised the mother.

Finally, as Abby didn't write to the grandma to inform about the success, the old woman was worry and thought it would be a good idea invited the couple a weekend at her home.

Start ▶

### Instructions

Find out what Grammar Checker does and how to use it:

- o Guide ([English](#)) ([Spanish](#))

## Spelling

Grammar checker

Check the most frequent words before and after:

+ New text • Spelling == Incorrect sequences --? Problem words .. Pairs filter ☰ Summary

### Result of the spelling filter:

Once upon a time there was a girl who was in love with a boy with a bad taste, Richard. As she was too young, she hadn't said to her mother anything about her relation in the long time she was loving Richard because she knew her mother wouldn't like him. But as their relation was going more and more well, she decided to introduce Richard at the family and, of course, the mother didn't like Richard since the moment she saw him. The worst thing was that Abby was not allowed to follow with the relation. So, Abby decided to write a letter to grandma telling her about the situation that made her feel very sad. The answer of the grandma didn't make Abby wait. Grandma gave some advice to the girl that couldn't go wrong. The piece of advice consisted in preparing a get-together with her mother again but, instead of carrying flowers to the mum, Richard should cook a cake which receipt that was attached in the answer letter. So, Abby proceed following the piece of advice and the meeting was made. It was a successful get-together, the delicious cake surprised the mother. Finally, as Abby didn't write to the grandma to inform about the success, the old woman was worry and thought it would be a good idea invited the couple a weekend at her home.

**How does it work?**

The **spelling filter** analyzes the frequency of each word you used. Then it tells you if some of your words are infrequent or very infrequent.

Legend:

**Yellow:** spelling mistakes with pedagogic feedback.

**Purple:** words that are not amongst the 90, 000 most frequent English words.

Click on them to get more info.

Modify

## Incorrect Sequences

+ New text
▪ Spelling
Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Incorrect sequences filter

Once upon a time there was a girl who was in love with a boy with a bad taste, Richard. As she was too young, she hadn't said to her mother anything about her relation in the long time she was loving Richard because she knew her mother wouldn't like him. But as their relation was going more and more well, she decided to introduce Richard at the family and, of course, the mother didn't like Richard since the moment she saw him. The worst thing was that Abby was not allowed to follow with the relation. So, Abby decided to write a letter to grandma telling her about the situation that made her feel very sad. The answer of the grandma didn't make Abby wait. Grandma gave some advice to the girl that couldn't go wrong. The piece of advice consisted in preparing a get-together with her mother again but, instead of carrying flowers to the mum, Richard should cook a cake which receipt that was attached in the answer letter. So, Abby proceed following the piece of advice and the meeting was made. It was a successful get-together, the delicious cake surprised the mother. Finally, as Abby didn't write to the grandma to **inform about** the success, the old woman was worry and thought it would be a good idea invited the couple a weekend at her home.

✎ Modify

#### How does it work?

The **incorrect sequences filter** compares your text with one million of the most frequent incorrect sequences in English.

This filter will give you personalized feedback to help you improve your text.

## Problem Words

+ New text
▪ Spelling
=== Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Problem words filter

Once upon a time there was a girl who was in love with a boy with a bad taste, Richard. As she was too young, she hadn't said to her mother anything about her relation in the long time she was loving Richard because she knew her mother wouldn't like him. But as their relation was going more and more well, she decided to introduce Richard at the family and, of course, the mother didn't like Richard since the moment she saw him. The worst thing was that Abby was not allowed to follow with the relation. So, Abby decided to write a letter to grandma telling her about the situation that made her feel very sad. The answer of the grandma didn't make Abby wait. Grandma gave some advice to the girl that couldn't go wrong. The piece of advice consisted in preparing a get-together with her mother again but, instead of carrying flowers to the mum, Richard should cook a cake which receipt that was attached in the answer letter. So, Abby proceed following the piece of advice and the meeting was made. It was a successful get-together, the delicious cake surprised the mother. Finally, as Abby didn't write to the grandma to tell about the success, the old woman was worry and thought it would be a good idea invited the couple a weekend at her home.

✎ Modify

#### How does it work?

The **problem words filter** shows at the bottom of the screen the words associated with error - words that many students of English use incorrectly.

Grammar Checker does not know if you have used them correctly or incorrectly. If you are not certain, click on the word and Grammar Checker will help you decide.



## Pairs Filter

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Pairs filter - Very suspicious combinations

Once upon a time there was a girl who was in love with a boy with a bad taste, Richard. As she was too young, she hadn't said to her mother anything about her relation in the long time she was loving Richard because she knew her mother wouldn't like him. But as their relation was going more and more well, she decided to introduce Richard at the family and, of course, the mother didn't like Richard since the moment she saw him. The worst thing was that Abby was not allowed to follow with the relation. So, Abby decided to write a letter to grandma telling her about the situation that made her feel very sad. The **answer of** the grandma didn't make Abby wait. Grandma gave some advice to the girl that couldn't go wrong. The piece of advice consisted in preparing a get-together with her mother again but, instead of carrying flowers to **the mum**, Richard should cook a cake which receipt that was attached in the answer letter. So, Abby proceed following the piece of advice and the meeting was made. It was a successful get-together, the delicious cake surprised the mother. Finally, as Abby didn't write to the grandma to tell about the success, the old woman **was worry** and thought it would be a good idea invited the **couple a** weekend at her home.

✎ Modify

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

Red: Very suspicious pairs  
Orange: Suspicious pairs  
Yellow: Slightly suspicious pairs  
All: All suspicious pairs

## Final version

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Pairs filter - Very suspicious combinations

Once upon a time there was a girl who was in love with a boy with a bad taste, Richard. As she was too young, she hadn't said to her mother anything about her relation in the long time she was loving Richard because she knew her mother wouldn't like him. But as their relation was going more and more well, she decided to introduce Richard at the family and, of course, the mother didn't like Richard since the moment she saw him. The worst thing was that Abby was not allowed to follow with the relation. So, Abby decided to write a letter to grandma telling her about the situation that made her feel very sad. The answer to the grandma didn't make Abby wait. Grandma gave some advice to the girl that couldn't go wrong. The piece of advice consisted in preparing a get-together with her mother again but, instead of carrying flowers to mum, Richard should cook a cake which receipt that was attached in the answer letter. So, Abby proceed following the piece of advice and the meeting was made. It was a successful get-together, the delicious cake surprised the mother. Finally, as Abby didn't write to the grandma to tell about the success, the old woman was worried and thought it would be a good idea invited the couple to dinner at her home.

✎ Modify

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

Red: Very suspicious pairs  
Orange: Suspicious pairs  
Yellow: Slightly suspicious pairs  
All: All suspicious pairs

## PAM 1st DRAFT

### Initial text

+ New text
• Spelling
=== Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Enter your text:

It was the day of her 18th birthday when Pam decided to write a letter to her grandma with the aim of telling her the problem she was worried about and asking her for advice. The fact was that Pam had been looking forward to that day because she needed to be an adult to be able to make her choices in her life. She was excited with the idea to have a coffeehouse and to work as the cashier but, that idea didn't like to her parents who wanted her to study a degree. For that reason, Pam was sad, she had wanted to have a coffeehouse long time ago. So, Pam wrote a letter to her grandma and the piece of advice given by her was an excellent idea. Grandma advised Pam that she could do both, studying and working in her own business. Moreover, grandma could lend her an amount of money to begin the business. When Pam raised the issue with her parents, they became upset but, she was determined, her grandma had encouraged her a lot. In fact, one year later, when the inauguration day of the Pam's coffeehouse arrived, grandma wrote a note to Pam telling she was going to give a speech in the great event.

Start ▶

Instructions

Find out what Grammar Checker does and how to use it:

- Guide [\(English\)](#) [\(Spanish\)](#)

### Spelling

+ New text
• Spelling
=== Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Result of the spelling filter:

It was the day of her 18th birthday when Pam decided to write a letter to her grandma with the aim of telling her the problem she was worried about and asking her for advice. The fact was that Pam had been looking forward to that day because she needed to be an adult to be able to make her choices in her life. She was excited with the idea to have a coffeehouse and to work as the cashier but, that idea didn't like to her parents who wanted her to study a degree. For that reason, Pam was sad, she had wanted to have a coffeehouse long time ago. So, Pam wrote a letter to her grandma and the piece of advice given by her was an excellent idea. Grandma advised Pam that she could do both, studying and working in her own business. Moreover, grandma could lend her an amount of money to begin the business. When Pam raised the issue with her parents, they became upset but, she was determined, her grandma had encouraged her a lot. In fact, one year later, when the inauguration day of the Pam's coffeehouse arrived, grandma wrote a note to Pam telling she was going to give a speech in the great event.

Modify

How does it work?

The **spelling filter** analyzes the frequency of each word you used. Then it tells you if some of your words are infrequent or very infrequent.

Legend:

**Yellow:** spelling mistakes with pedagogic feedback.

**Purple:** words that are not amongst the 90,000 most frequent English words.

Click on them to get more info.

## Incorrect Sequences

+ New text   ▪ Spelling   Incorrect sequences   --? Problem words   .. Pairs filter   ☰ Summary

### Incorrect sequences filter

It was the day of her 18th birthday when Pam decided to write a letter to her grandma with the aim of telling her the problem she was worried about and asking her for advice. The fact was that Pam had been looking forward to that day because she needed to be an adult to be able to make her choices in her life. She was excited with the idea to have a coffeehouse and to work as the cashier but, that idea didn't like to her parents who wanted her to study a degree. For that reason, Pam was sad, she had wanted to have a coffeehouse long time ago. So, Pam wrote a letter to her grandma and the piece of advice given by her was an excellent idea. Grandma advised Pam that she could do both, studying and working in her own business. Moreover, grandma could lend her an amount of money to begin the business. When Pam raised the issue with her parents, they became upset but, she was determined, her grandma had encouraged her a lot. In fact, one year later, when the inauguration day of the Pam's coffeehouse arrived, grandma wrote a note to Pam telling she was going to give a speech in the great event.

[Modify](#)

**How does it work?**

The **incorrect sequences filter** compares your text with one million of the most frequent incorrect sequences in English.

This filter will give you personalized feedback to help you improve your text.

## Problem Words

+ New text   ▪ Spelling   Incorrect sequences   Problem words   .. Pairs filter   ☰ Summary

### Problem words filter

It was the day of her 18th birthday when Pam decided to write a letter to her grandma with the aim of telling her the problem she was worried about and asking her for advice. The fact was that Pam had been looking forward to that day because she needed to be an adult to be able to make her choices in her life. She was excited with the idea to have a coffeehouse and to work as the cashier but, that idea didn't like to her parents who wanted her to study a degree. For that reason, Pam was sad, she had wanted to have a coffeehouse long time ago. So, Pam wrote a letter to her grandma and the piece of advice given by her was an excellent idea. Grandma advised Pam that she could do both, studying and working in her own business. Moreover, grandma could lend her an amount of money to begin the business. When Pam raised the issue with her parents, they became upset but, she was determined, her grandma had encouraged her a lot. In fact, one year later, when the inauguration day of the Pam's coffeehouse arrived, grandma wrote a note to Pam telling she was going to give a speech in the great event.

[Modify](#)

**How does it work?**

The **problem words filter** shows at the bottom of the screen the words associated with error - words that many students of English use incorrectly.

Grammar Checker does not know if you have used them correctly or incorrectly. If you are not certain, click on the word and Grammar Checker will help you decide.

## Pairs Filter

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Pairs filter - Very suspicious combinations

It was the day of her 18th birthday when Pam decided to write a letter to her grandma with the aim of telling **her the** problem she was worried about and asking her for advice. The fact was that Pam had been looking forward to that day because she needed to be an adult to be able to make her choices in her life. She was excited with the idea to have a coffeehouse and to work as the cashier but, that idea didn't like to her parents who wanted her to study a degree. For that reason, Pam was sad, she had wanted to have a coffeehouse long time ago. So, Pam wrote a letter to her grandma and the piece of advice given by her was an excellent idea. Grandma advised Pam that she could do both, studying and working in her own business. Moreover, grandma could lend her an amount of money to begin the business. When Pam raised the issue with her parents, they became upset but, she was determined, her grandma had encouraged her a lot. In fact, one year later, when the inauguration day of the Pam's coffeehouse arrived, grandma wrote a note to Pam telling she was going to give a speech in the great event.

✎ Modify

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

Red: **Very suspicious pairs**

Orange: Suspicious pairs

Yellow: Slightly suspicious pairs

All: All suspicious pairs

## Final version

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Pairs filter - Very suspicious combinations

It was the day of her 18th birthday when Pam decided to write a letter to her grandma with the aim of telling the problem she was worried about and asking her for advice. The fact was that Pam had been looking forward to that day because she needed to be an adult to be able to make her choices in her life. She was excited with the idea to have a coffeehouse and to work as the cashier but, that idea didn't like to her parents who wanted her to study a degree. For that reason, Pam was sad, she had wanted to have a coffeehouse long time ago. So, Pam wrote a letter to her grandma and the piece of advice given by her was an excellent idea. Grandma advised Pam that she could do both, studying and working in her own business. Moreover, grandma could lend her an amount of money to begin the business. When Pam raised the issue with her parents, they became upset but, she was determined, her grandma had encouraged her a lot. In fact, one year later, when the inauguration day of the Pam's coffeehouse arrived, grandma wrote a note to Pam telling she was going to give a speech in the great event.

✎ Modify

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

Red: **Very suspicious pairs**

Orange: Suspicious pairs

Yellow: Slightly suspicious pairs

All: All suspicious pairs

## PAM 2nd DRAFT

### Initial text

+ New text   • Spelling   ≡ Incorrect sequences   --? Problem words   •• Pairs filter   ☰ Summary

**Enter your text:**

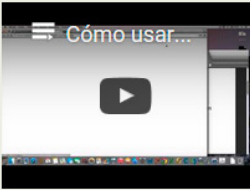
PAM  
 It was the day of her 18th birthday when Pam awoke as fresh as a daisy. So, she decided what she wanted to do in her near future. It was a beautiful and a warm day, the sun was shining as wanting announce the extraordinary change Pam was about to carry out. Pam was a nice girl and a superb student. But, she dreamt about working as a cashier in her own coffeehouse. Evidently, this idea didn't like her parents. They wished she studied any degree before setting up a business. However, today she was determined. Pam wrote a short letter to her tender and wiser grandma telling her what was concerning Pam lately. Grandma was a fabulous woman, she had a lot of experience in life. Immediately, she understood the situation, wrote two letters, one for Pam and the other one for her son and her daughter in law. In a wonderful way, grandma persuaded her son to let Pam to work in holidays and, at the same time, Pam was convinced by her grandma to study a lot to become a great businesswoman. Finally, everyone was pleased with the end of the conflict. Pam and her parents were grateful to the grandmother, and the elderly woman rest in peace.]

Start ▶

**Instructions**

Find out what Grammar Checker does and how to use it:

- o Guide ([English](#)) ([Spanish](#))



### Spelling

+ New text   • Spelling   ≡ Incorrect sequences   --? Problem words   •• Pairs filter   ☰ Summary

**Result of the spelling filter:**

PAM It was the day of her 18th birthday when Pam awoke as fresh as a daisy. So, she decided what she wanted to do in her near future. It was a beautiful and a warm day, the sun was shining as wanting announce the extraordinary change Pam was about to carry out. Pam was a nice girl and a superb student. But, she dreamt about working as a cashier in her own coffeehouse. Evidently, this idea didn't like her parents. They wished she studied any degree before setting up a business. However, today she was determined. Pam wrote a short letter to her tender and wiser grandma telling her what was concerning Pam lately. Grandma was a fabulous woman, she had a lot of experience in life. Immediately, she understood the situation, wrote two letters, one for Pam and the other one for her son and her daughter in law. In a wonderful way, grandma persuaded her son to let Pam to work in holidays and, at the same time, Pam was convinced by her grandma to study a lot to become a great businesswoman. Finally, everyone was pleased with the end of the conflict. Pam and her parents were grateful to the grandmother, and the elderly woman rest in peace.

Modify

**How does it work?**

The **spelling filter** analyzes the frequency of each word you used. Then it tells you if some of your words are infrequent or very infrequent.

Legend:

- Yellow:** spelling mistakes with pedagogic feedback.
- Purple:** words that are not amongst the 90, 000 most frequent English words.

Click on them to get more info.

## Incorrect Sequences

+ New text
▪ Spelling
Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Incorrect sequences filter

PAM It was the day of her 18th birthday when Pam awoke as fresh as a daisy. So, she decided what she wanted to do in her near future. It was a beautiful and a warm day, the sun was shining as **wanting announce** the extraordinary change Pam was about to carry out. Pam was a nice girl and a superb student. But, she dreamt about working as a cashier in her own coffee house. Evidently, this idea didn't like her parents. They wished she studied any degree before setting up a business. However, today she was determined. Pam wrote a short letter to her tender and wiser grandma telling her what was concerning Pam lately. Grandma was a fabulous woman, she had a lot of experience in life. Immediately, she understood the situation, wrote two letters, one for Pam and the other one for her son and her daughter in law. In a wonderful way, grandma persuaded her son to let Pam to work in holidays and, at the same time, Pam was convinced by her grandma to study a lot to become a great businesswoman. Finally, everyone was pleased with the end of the conflict. Pam and her parents were grateful to the grandmother, and the elderly woman rest in peace.

✎ Modify

**How does it work?**

The **incorrect sequences filter** compares your text with one million of the most frequent incorrect sequences in English.

This filter will give you personalized feedback to help you improve your text.

## Problem Words

+ New text
▪ Spelling
Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Problem words filter

PAM It was the day of her 18th birthday when Pam awoke as fresh as a daisy. So, she decided what she wanted to do in her near future. It was a beautiful and a warm day, the sun was shining as announcing the extraordinary change Pam was about to carry out. Pam was a nice girl and a superb student. But, she dreamt about working as a cashier in her own coffee house. Evidently, this idea didn't like her parents. They wished she studied any degree before setting up a business. However, today she was determined. Pam wrote a short letter to her tender and wiser grandma telling her what was concerning Pam lately. Grandma was a fabulous woman, she had a lot of experience in life. Immediately, she understood the situation, wrote two letters, one for Pam and the other one for her son and her daughter in law. In a wonderful way, grandma persuaded her son to let Pam to work in holidays and, at the same time, Pam was convinced by her grandma to study a lot to become a great businesswoman. Finally, everyone was pleased with the end of the conflict. Pam and her parents were grateful to the grandmother, and the elderly woman rest in peace.

✎ Modify

**How does it work?**

The **problem words filter** shows at the bottom of the screen the words associated with error - words that many students of English use incorrectly.

Grammar Checker does not know if you have used them correctly or incorrectly. If you are not certain, click on the word and Grammar Checker will help you decide.

## Pairs Filter

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Pairs filter - Very suspicious combinations

PAM It was the day of her 18th birthday when Pam awoke as fresh as a daisy. So, she decided what she wanted to do in her near future. It was a beautiful and a warm day, the sun was shining as announcing the extraordinary change Pam was about to carry out. Pam was a nice girl and a superb student. But, she dreamt about working as a cashier in her own coffee house. Evidently, this idea didn't like her parents. They wished she studied any degree before setting up a business. However, today she was determined. Pam wrote a short letter to her tender and wiser grandma telling her what **was concerning** Pam lately. Grandma was a fabulous woman, she had a lot of experience in life. Immediately, she understood the situation, wrote two letters, one for Pam and the other one for her son and her daughter in law. In a wonderful way, grandma persuaded her son to let Pam to work in holidays and, at the same time, Pam was convinced by her grandma to study a lot to become a great businesswoman. Finally, everyone was pleased with the end of the conflict. Pam and her parents were grateful to the grandmother, and the elderly woman rest in peace.

✎ Modify

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

Red: Very suspicious pairs

Orange: Suspicious pairs

Yellow: Slightly suspicious pairs

All: All suspicious pairs

## Final version

+ New text
▪ Spelling
≡ Incorrect sequences
--? Problem words
.. Pairs filter
☰ Summary

### Pairs filter - Very suspicious combinations

PAM It was the day of her 18th birthday when Pam awoke as fresh as a daisy. So, she decided what she wanted to do in her near future. It was a beautiful and a warm day, the sun was shining as announcing the extraordinary change Pam was about to carry out. Pam was a nice girl and a superb student. But, she dreamt about working as a cashier in her own coffee house. Evidently, this idea didn't like her parents. They wished she studied any degree before setting up a business. However, today she was determined. Pam wrote a short letter to her tender and wiser grandma telling her what concerned Pam lately. Grandma was a fabulous woman, she had a lot of experience in life. Immediately, she understood the situation, wrote two letters, one for Pam and the other one for her son and her daughter in law. In a wonderful way, grandma persuaded her son to let Pam to work in holidays and, at the same time, Pam was convinced by her grandma to study a lot to become a great businesswoman. Finally, everyone was pleased with the end of the conflict. Pam and her parents were grateful to the grandmother, and the elderly woman rest in peace.

✎ Modify

**How does it work?**

The **pairs filter** analyses the frequency of each combination of two words that you used. Then it tells you if some of your phrases are infrequent or very infrequent.

Currently **Very suspicious pairs** are analyzed. **Click** on the following buttons to check other suspicious pairs:

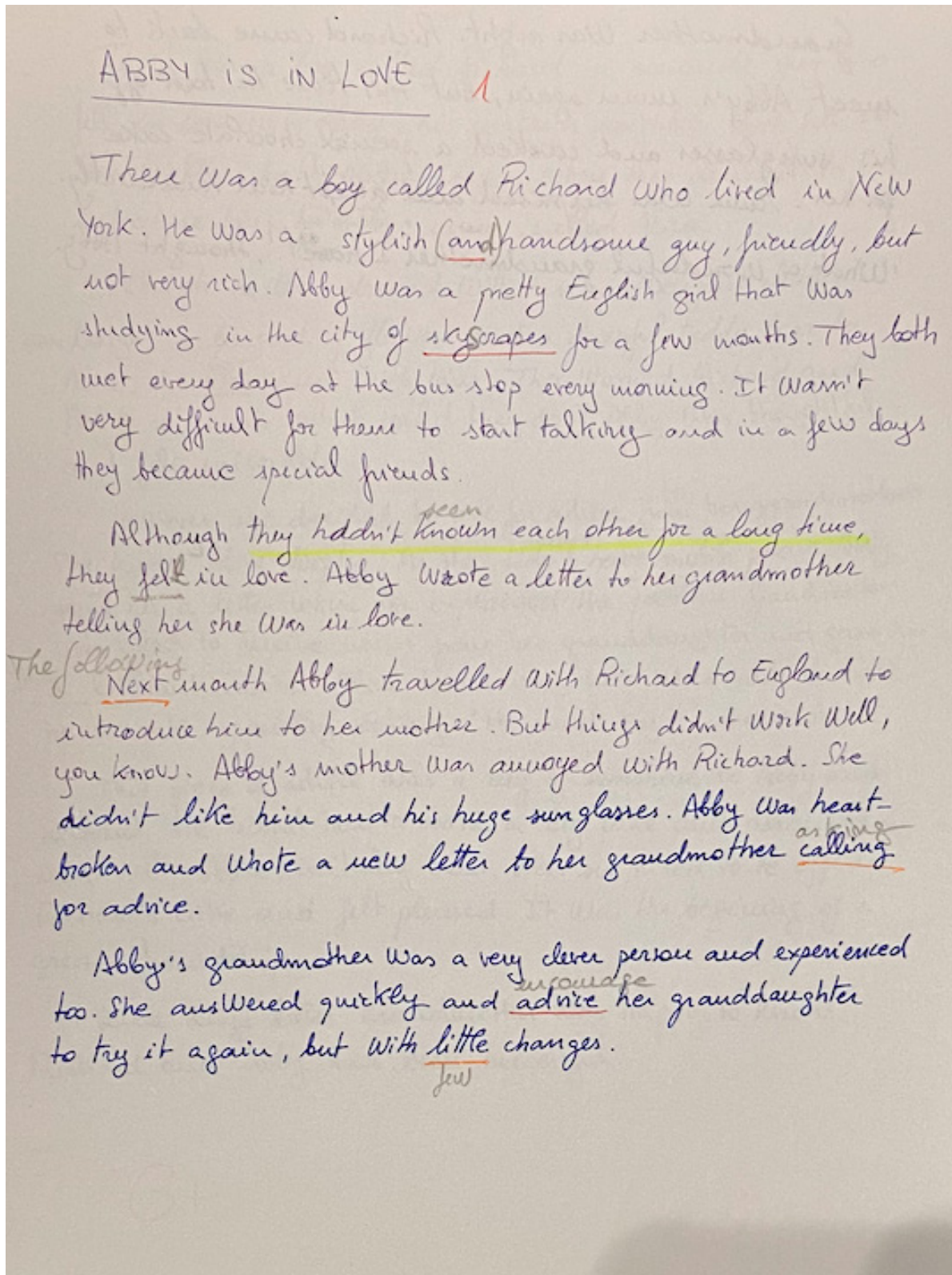
Red: Very suspicious pairs

Orange: Suspicious pairs

Yellow: Slightly suspicious pairs

All: All suspicious pairs

Example of a complete set of writings of a student in Group 2 (teacher feedback).





Grandmother was right. Richard came back to meet Abby's mum again, but this time he took off his sunglasses and cooked a special chocolate cake for her. Mum was surprised and <sup>loved</sup> ~~accept~~ <sup>accepted</sup> him immediately. 'What a wonderful grandmother I have!', thought Abby.

## ABBY IS IN LOVE (AGAIN)

2

There was a boy called Richard, a romantic guy who fell in love with Abby, his upstairs neighbor. Both fell in love quickly, but problems began when they decided to introduce him to Abby's mum, called Alice.

Although he was dressed <sup>up</sup>with his most elegant clothes, ~~she~~ bought a bouquet of flowers and a cheerful teddy bear to Alice, she <sup>got</sup>was upset with him. This worried Richard and Abby so much. What could they do? Abby was thoughtful and felt miserable.

However she decided to ask for advice <sup>to</sup>from her grandmother. She knew Alice deeply. As she <sup>hadn't</sup>didn't have mobile phone Abby said her a letter where she explained the problem. Grandmother was glad to receive news from her granddaughter and gave her <sup>some</sup>a brilliant advice: Alice loved men who were able to make homework, especially cooking. 'He must try it', she wrote.

This piece of advice was a ray of sunshine to Abby and Richard. He learnt how to make a big cake and arranged a new visit. Alice, initially cross, was surprised to be offered Richard's cake and felt pleased. It was the beginning of a great friendship.

Some days later grandmother was happy to know Richard and Abby had been successful.

1

It was the day of her 18<sup>th</sup> birthday when Pam wrote a letter to grandma <sup>where</sup> in which she remembered what <sup>had happened</sup> happened some months ago. Pam was a cheerful high school student who dreamed <sup>about working</sup> to work as a saleswoman in a <sup>coffee</sup> coffee-shop, a challenging job, she thought, to leave the nest.

However, Pam's Mum and Dad had <sup>other</sup> another plans for her. They tried to motivate her to study in order to reach <sup>good prospects</sup> future better jobs, with good prospects. But Pam only thought <sup>about/of</sup> in money and, as she was <sup>broken</sup> broken, she decided to write a letter to grandma, who <sup>always</sup> always was prepared to give some advice.

Surely she was. Grandma replied recommending Pam to continue with her studies until she <sup>would wish</sup> had finished them. So she <sup>were</sup> did and Mum and Dad <sup>again</sup> returned to be happy and satisfied. Unfortunately, when Pam finished her high school studies told her parents she didn't want to continue studying <sup>any</sup> more. She only thought about making money.

Mum and Dad were <sup>frightened</sup> horrified. As they knew the <sup>close</sup> close relationship between Grandma and Pam, <sup>that</sup> wrote her a letter asking for some help. Fastly grandma got <sup>quicker</sup> in touch with Pam and encouraged her to combine work and studies. Finally Pam began to work in the evenings and <sup>to</sup> study at the university in the mornings.

It was the day of her 18<sup>th</sup> birthday when Pam began to dream about how her ideal adulthood <sup>could</sup> will be. She imagined herself working as a saleswoman and earning enough money to travel, go out with friends and also fly the nest. 'How wonderful it might be', she whispered when she was deeply thoughtful.

But dreams are one thing, reality <sup>other</sup> another. Pam's parents were determined to make her study. 'You must study if you want to get a good job in order to have earn a high salary and <sup>have</sup> a comfortable life', they repeatedly argued.

Grandma was aware of everything that was wrong with Pam. Not long afterwards she got in touch with Pam's parents and reminded them their young time, when they combined studies with work. She hit the target.

Out of <sup>the</sup> blue Pam's parents changed their minds and offered Pam an agreement. She ~~started~~ study hard and could work every evening. So, she <sup>could</sup> can earn money without leaving <sup>her</sup> studies. 'Thanks, Grandma', Pam shouted and her astonished parents understood everything...

Eventually everybody was happy. Grandma knew she still was influential, Pam's parents realized that her daughter had grown up and Pam began to have an independent life.

Example of a complete set of writings of a student in Group 3 (Control Group). Words/phrases highlighted in yellow correspond to the three targeted grammatical errors not corrected during the study but marked once it was finished to analyse them.

**ABBY IN LOVE** 1

There was a boy called Richard, that one day, ~~he~~ decided to write a letter to his grandma, in this letter Richard told her that he had met a beautiful girl and he had fallen in love with Abby and she was also in love.

Surprisingly, Richard <sup>to</sup> went Abby's home and she introduced him to her parents.

Unfortunately, Abby's mother was annoyed with them because she didn't like Richard, considering that he was really older than her. For this reason, ~~Abby~~ was heartbroken.

Richard was desperate, so he wrote again a letter to his grandma telling what happened to Abby.

As soon as grandma read the letter, she decided to send a letter to Abby advising her about the relationship.

Fortunately, Abby followed the advice of the grandma and she talked to her mother to explain her feeling for Richard.

After all this,  
Immediately after, they met again and Abby's mother knew him better and she let her prejudices aside. Finally, they had a good relationship.

Meanwhile, grandma was worried, so she wrote a letter to the couple to ask how had finished their love story.

## ABBY IN LOVE II

2

### PAM'S STORY

who

There was a boy called Richard really fell in love with Abby the moment he saw her. His girlfriend was a beautiful and caring girl.

Suddenly, Abby's grandma wrote a letter to her granddaughter to know how she was, but at that time Abby was really sad, because her mother and Richard had argued and her mother told her that she had to end the relationship with Richard, for this reason she was heartbroken.

After that, Abby wrote a letter to grandma to tell her what had happened with her mum and Richard and how she felt.

Consequently, Abby's grandma wrote a letter to Abby giving advice on what Abby could do about it.

From Abby's grandma point of view, Richard and her mother had to meet again, so they could talk about their problems and they could try to solve them. <sup>that</sup>

After a few days, grandma was worried, so ~~that~~ she decided to write another letter to Abby to know what had happened with her difficult relationship with Richard.

## PAM'S STORY

1

It was the day of her 18<sup>th</sup> birthday when Pam found in a box an ancient letter. Immediately, Pam read it and she discovered that it was a letter that someone wrote to her grandma because she wanted to open a coffee shop forty years ago.

At the time Pam was a student, however she had a dream, the same as her grandma, it was open her own coffee shop.

<sup>At that point</sup> ~~the~~ Actually, Pam had several problems, she was not a good student because she was a bit lazy, although her parents tried to motivate her, besides he was in the red.

Consequently, she was so depressed that she decided to write a letter to her grandma to give her some advice.

After that, grandma told Pam that she should continue her studies and when she finished them it was the perfect moment to realize her dream. <sup>offer</sup>

Finally, Pam studied really hard and she got good grades, also she tightened her belt and especially with the help of her parents, she could open her own coffee shop. <sup>so that</sup>

## PAM'S STORY II 2

It was the day of her 18<sup>th</sup> birthday when Pam left her huge High School, it was a sunny day in June and Pam felt really overjoyed to be adult.

Just at that moment she thought one of the most important people in her life, he wanted to speak with her lovely grandma to explain that she was over the moon, however there was an enormous problem, her grandma lived in a tiny village far, so she decided to write a letter to comment how she felt.

As soon as she got home, she realized that the next day she had to make a hard exam, it was the last examination at High School, and she needed to study. At that moment she seemed downhearted. "I can do it!", thought Pam, and she was studying all endless night, so she was exhausted.

Later that day she got went at High School to make the exam and she received good grades, Pam was delighted. Suddenly she realized that she wanted to be a business owner, for this reason she had to continue studying and go at the University. Immediately after she had a serious talk with her parents to explain them her wonderful decision. "Great decision!", screamed her parents. Pam had never seen her parents so elated.

After all that had happened Pam wrote a letter to her charming grandma announced the good news.

In the end, Pam has opened her own restaurant. She is Feeling so pleased for her courageous decision.

## APPENDIX 4: Tailor-made test

## Group 1

1. There was a boy called Richard of whom Abby was in love with. ✓
2. She attended at her cousin's birthday where Richard was too. <sup>was</sup> did / attended ✓
3. Since that day, they were going out together. <sup>went</sup> had been ✓
4. The mother didn't like Richard at once. <sup>at</sup> the beginning of the relation. ✓
5. For next month, Abby felt very miserable.
6. Richard would have to go to the Abby's home with a cake. <sup>should</sup>
7. Richard tried (to make her feel) comfortable and confident. <sup>she felt</sup> long with 'Richard' ✓
8. As she was too young, she hadn't said to her mother anything about her relation in the long time she was living Richard.
9. She decided to introduce Richard to the family. ✓
10. The answer to the grandma didn't make Abby wait. <sup>about</sup>
11. Abby decided to write a letter to grandma telling her about the situation that made her feel sad. <sup>had</sup> <sup>feeling</sup>
12. The receipt was attached in the answer letter.
13. Abby proceeded following the piece of advice. <sup>to follow with</sup>
14. Grandma started another debate but with no success. <sup>had already</sup>
15. When the inauguration day of the Pam's coffehouse arrived, grandma wrote a note. ✓
16. In the end she managed to finish her studies and set up her own business. ✓
17. Pam wrote a short letter to her tender and wise grandma telling her what concerned Pam lately. <sup>had</sup>
18. Grandma persuaded her son to let Pam to work in holidays. ✓
19. Pam and her parents were grateful to the grandmother. <sup>for her help</sup> <sup>grandma</sup> after that, ✓

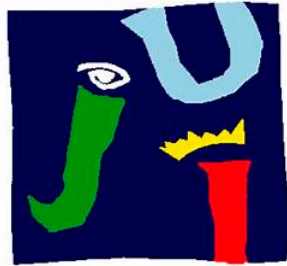


## Group 2

1. Although he bought two gifts for her, my mum felt annoyed with him. ✓
2. I wrote a letter ~~for~~ <sup>to</sup> her. ✓
3. I told her that she was right because things ~~were~~ <sup>had been</sup> much better this time ✓  
(2<sup>nd</sup> try with her mum).
4. Richard tried ~~to~~ <sup>that she felt comfortable and confident</sup> make her feel comfortable and confident. ✓
5. It was love at ~~the~~ first sight. ✓
6. We had been getting together ~~during~~ <sup>for</sup> a month. ✓
7. Grandma started another debate but ~~with no success~~ <sup>there wasn't</sup> ~~unsuccessfully~~. ✓
8. Pam didn't want to ~~work in~~ ~~a~~ full-time job as she preferred to carry on ~~studying~~ <sup>her studies</sup>. ✓
9. In the end she managed to finish her studies and ~~set~~ <sup>take</sup> up her own business. ✓
10. They stared ~~at~~ <sup>at</sup> her with signs of concern and anxiety. ✓

## Group 3

1. He wrote again <sup>a piece of</sup> (a letter) to his grandma telling what happened to Abby. ✓
2. Abby followed the advice of the grandma and she talked to her. ✓
3. Richard tried to make her feel (comfortable and confident.) <sup>better</sup> ✓
4. Abby was really sad because her mother and Richard had argued (and her mother told her that she had to <sup>finish</sup> end the relationship. ✓
5. Grandma started another (debate) but (with no) success. <sup>without</sup> ✓
6. In the end she managed <sup>could</sup> to finish her studies and set up her own business. ✓
7. She thought (one of the most important people in her life; she wanted to <sup>to</sup> speak) talk with her lovely grandma. ✓
8. She had to continue studying and (go) <sup>going to</sup> at the university. ✓
9. In the end, Pam ~~has~~ <sup>could open</sup> opened her own restaurant. ✓



UNIVERSITAT  
JAUME I

Corrective Feedback in the EFL Classroom:  
*Grammar Checker* vs. Teacher's Feedback.