

UNIVERSIDAD DEL PAÍS VASCO/EUSKAL HERRIKO UNIBERTSITATEA



# **THE USE OF MODELS AS WRITTEN CORRECTIVE FEEDBACK IN SPONTANEOUS EFL WRITING**

---

**Master's degree in Language Acquisition in Multilingual Settings (LAMS)**

**Faculty of Arts (University of the Basque Country)**

By: Udane Loidi Labandibar

Supervised by: María del Pilar García Mayo

Vitoria-Gasteiz, June, 2016

**Official MA program in**  
***Language Acquisition in Multilingual Settings (LAMS)***  
**Universidad del País Vasco/Euskal Herriko Unibertsitatea**  
***Non-Plagiarism Statement***

This form must be completed, dated and signed and must be included at the beginning of every copy of the MA Thesis you submit for assessment.

<i>Name and surnames:</i>	Udane Loidi Labandibar
<i>MA Thesis title:</i>	The use of models as written corrective feedback in spontaneous EFL writing
<i>Supervisor:</i>	María del Pilar García Mayo


**I HEREBY DECLARE THAT**

- This MA Thesis that I am submitting for assessment is entirely my own work and I have written it completely by myself.
- I have not previously submitted this work or any version of it for assessment in any other programme or institution.
- I have not used any other sources or resources than the ones mentioned.
- I have identified and included the source of all facts, ideas, opinions and viewpoints of others through in-text referencing and the relevant sources are all included in the list of references at the end of my work. Direct quotations from books, journal articles, internet sources or any other source whatsoever are acknowledged and the sources cited are identified in the list of references.

I understand that plagiarism and copying are serious offences. In case of proof that this MA Thesis fails to comply with this declaration, either as negligence or as a deliberate act, I understand that the examiner has the right to exclude me from the assessment act and consequently all research activities conducted for this course will be declared null and the MA Thesis will not be presented for public defence, thus obtaining the lowest qualification.

**Date:** 03 June 2016

**Signature:**



## ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Professor María del Pilar García Mayo for her constant help and advice, her valuable comments and mostly for her patience and understanding with my delays and inevitable pauses.

I am also grateful to Oiartzo Batxilergo Ikastola, IES Usandizaga-Peñaflorida-Amara BHI and all the students who made this project possible for their kind collaboration.

I am equally indebted to Jeremy Hurdus and Naroa Fagoaga, who are the authors of the models used in the study.

My sincere gratitude goes to Dr. Vicente Núñez-Antón for his assistance with the statistical analysis and for sharing his time and knowledge with me.

Amaia Munarriz and Maria-José Ezeizabarrena also deserve my appreciation for giving me the opportunity to learn by their side as well as for trusting my inexperienced work.

My immense gratitude to all my family members, to Julen and to my friends for their love, care and support. My most heartfelt thanks to Atto for all the candles he lit in my life: Azkenengo honetan bat piztuko ez didazun arren, ziur aurreko urtetako guztiek indarra emango didatela.

I am also extremely grateful to all the LAMS team for introducing me in this exciting field.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	i
LIST OF TABLES .....	iv
LIST OF ABBREVIATIONS .....	v
ABSTRACT .....	vi
<b>1. Introduction</b> .....	<b>1</b>
<b>2. Literature review</b> .....	<b>5</b>
2.1 <i>Key concepts</i> .....	5
2.1.1 <i>Output</i> .....	5
2.1.2 <i>Corrective Feedback</i> .....	10
2.1.3 <i>Focus on Form</i> .....	14
2.2 <i>Previous research</i> .....	17
2.3 <i>Research questions and hypotheses</i> .....	26
<b>3. The study</b> .....	<b>29</b>
3.1 <i>Context and participants</i> .....	29
3.2 <i>Tasks and procedures</i> .....	30
3.3 <i>Analysis</i> .....	34
<b>4. Results</b> .....	<b>40</b>
<b>5. Discussion</b> .....	<b>57</b>
<b>6. Limitations and further research</b> .....	<b>65</b>
<b>7. Conclusions and pedagogical implications</b> .....	<b>67</b>
<b>References</b> .....	<b>69</b>
<b>Appendices</b> .....	<b>77</b>
Appendix A: Consent form .....	77
Appendix B: Visual stimuli .....	79
Appendix C: Sheet 1 .....	80
Appendix D: Sheet 2 .....	81
Appendix E: Likert-scale .....	82
Appendix F: The models .....	83
Appendix G: Sheet 3 – Unguided noticing group (UNG).....	84
Appendix H: Sheet 3 – Guided noticing group (GNG).....	85
Appendix I: Sheet 4.....	87
Appendix J: Exit questionnaire .....	88

## LIST OF TABLES

Table 1. Details of the participants in the study. ....	300
Table 2. Data collection time frame for the elementary students. ....	311
Table 3. Data collection time frame for the lower intermediate students.....	311
Table 4. Frequencies and proportions of PFNs. ....	400
Table 5. Frequencies and proportions of PFNs in the elementary treatment groups.....	422
Table 6. Frequencies and proportions of PFNs in the lower intermediate treatment groups.. ....	42
Table 7. Frequencies and proportions of FNs.....	433
Table 8. Frequencies and proportions of FNs in the elementary treatment groups.....	44
Table 9. Frequencies and proportions of FNs in the lower intermediate treatment groups. ....	44
Table 10. Stage 2 FN in relation to Stage 1 PFNs.....	46
Table 11. Solutions to PFNs noticed and incorporated across groups.....	48
Table 12. Solutions to PFNs noticed and incorporated across proficiency groups. ....	49
Table 13. Solutions to PFNs noticed and incorporated within the elementary treatment groups..	49
Table 14. Solutions to PFNs noticed and incorporated within the lower intermediate treatment groups .....	500
Table 15. Level of eagerness, enjoyment and usefulness of models in the elementary and lower intermediate treatment groups .....	511
Table 16. Participants' eagerness to use models in future EFL lessons .....	533
Table 17. Familiarity with models in the four treatment groups. ....	555

## LIST OF ABBREVIATIONS

- CF: Corrective Feedback
- EFL: English as a Foreign Language
- ESL: English as a Second Language
- EC: Error Correction
- FN: Feature Noticed
- FNI: Feature Noticed and Incorporated
- FonF: Focus on Form
- GNG: Guided Noticing Group
- IL: Interlanguage
- LLP: Language Learning Potential
- L1: First Language
- L2: Second Language
- LRE: Language Related Episode
- PFN: Problematic Feature Noticed
- SLA: Second Language Acquisition
- TL: Target Language
- UNG: Unguided Noticing Group

## ABSTRACT

While a number of studies have investigated the noticing function of output, the language learning potential of writing and the way in which written corrective feedback (CF) can promote language acquisition have been underexplored. In a three-stage writing task including output, comparison and delayed revision, the present study investigates what Basque-Spanish teenage learners (n=60) of English as a foreign language (EFL) notice when writing a composition in response to some pictures (Stage 1) and when comparing their texts with two models (Stage 2). It also explores how these noticing and feedback processing affects their subsequent revisions (Stage 3). The findings revealed that participants noticed mainly lexical problems, although the comparison with the models also allowed them to pay attention to content features. Regarding proficiency and guiding effects, it was found that more proficient learners and those who received guidance noticed more features. Finally, a qualitative analysis of the results indicated that learners had quite a negative attitude towards writing and modelling and that those who showed more positive beliefs incorporated more features in subsequent revisions. A number of implications drawn from these results are discussed for research and pedagogy.

**Keywords:** Noticing, composing, written feedback, models, Output Hypothesis, EFL.

## 1. Introduction

Research on learners' output and on the corrective feedback (CF) provided to it has attracted plenty of attention in the field of second language acquisition (SLA). For years, scholars have investigated the language learning potential (LLP) of producing output (Adams, 2003; Hanaoka, 2007; Lapkin, Swain & Smith, 2002; Swain & Lapkin, 1995; Swain & Lapkin, 2002) and its role in drawing learners' attention to form (Schmidt, 2010). According to Swain (1995), although providing learners with comprehensible, accurate and large amounts of input is necessary (Krashen, 1982), this is not sufficient. She contends that pushing learners to produce output allows them to become aware of their limitations in the second language (L2), to test out hypotheses about the features in the target language (TL), to reflect upon the use and meaning of target forms in contextualised environments and to automatize existing knowledge. That is to say, by producing output learners become not only more fluent, but also more accurate in the foreign language (FL).

Nevertheless, output per se may not serve these functions unless learners receive sufficient feedback regarding the appropriateness and accuracy of their message. Several studies (Adams, 2003; Qi & Lapkin, 2001; Santos, López-Serrano & Manchón, 2010; Swain, 1995) have acknowledged that the processing of CF can engage learners in actions such as noticing and focus on form (FonF) (Long, 1991, 1996) and, as Schmidt (1990, 2001) puts it, these are prerequisites for learning to occur. According to this scholar, input does not become intake unless it is noticed (i.e. consciously registered). Qi and Lapkin (2001) went a step further and claimed that it is not only noticing, but noticing with understanding that leads to language learning. Therefore, we may affirm that SLA is largely driven by what learners pay attention to and become aware of. It is in this sense that producing output and processing feedback can develop learners' interlanguage (IL).



However, and despite the central position that **oral** CF studies occupy in L2 acquisition research (see Lyster, Saito & Sato, 2013), the way in which **written** CF can contribute to language acquisition has been understudied until recently. As Santos et al. (2010) point out, the L2 writing research agenda has mainly addressed issues related to the manner in which learners acquire their writing skills, i.e. the learning-to-write dimension of writing (see Byrnes & Manchón, 2014; Manchón, 2011). In contrast, little attention has been devoted to the writing-to-learn dimension, whose focus shifts from studying how people learn to write, to investigating how writing can promote language acquisition (Manchón, 2011). Yet, this scenario is changing and it is increasingly accepted that writing can lead to IL development in ways which may not have been apparent before (Ferris, 2010; Santos et al., 2010; Manchón, 2011). As Sheen (2010:175) summarizes it:

[...] instead of viewing the goal of teaching writing as that of improving the learners' writing skills, practice in writing can be seen as one form of output that, in conjunction with CF, can facilitate interlanguage development. In other words, instruction that incorporates written CF constitutes a technique to draw L2 learners' attention to linguistic forms in their own output and thereby facilitate acquisition.

Nonetheless, as Manchón (2011) and Yang and Zhang (2010) argue, in spite of the current shift towards the writing-to-learn dimension, there is still a need for further research on the LLP of writing in different educational contexts. Thus, with the intention of shedding some light on this issue, the present study investigated the relationship between what Basque-Spanish adolescent learners of English as a Foreign Language (EFL) noticed while composing and processing CF in the form of models and what they incorporated in subsequent revisions.

The study was inspired by Hanaoka (2007) who in a four-stage writing task found out that his Japanese EFL learners were able to notice their limitations while composing a text, to autonomously find solutions to their problems in the models provided and to integrate these solutions in subsequent revisions. Moreover, he discovered that more proficient learners detected more problematic features and incorporated more solutions than less proficient learners. Hence, Hanaoka (2007) reported that text composition and modelling are effective techniques to enhance noticing and, thereby, promote language learning. However, these results were based on the performance of Japanese university students whose L2 proficiency ranged from intermediate to advanced and, as such, they might not be generalizable to other contexts, populations and levels of proficiency. In addition, given that Hanaoka (2007) did not include a control group, the findings should be interpreted with caution. Moreover, the way in which learner-related variables other than proficiency (e.g. attitudes and beliefs) affect learners' noticing remains an open question; especially when it comes to the use of models, since, to the best of our knowledge, there is no study analysing this issue.

Within this context, the present study attempts to investigate what Basque-Spanish adolescent EFL learners whose proficiency ranges from elementary to lower intermediate spontaneously notice and incorporate when engaged in a three-stage writing task: composing (stage 1), comparing through the use of models (stage 2) and rewriting (stage 3). The study intends to add empirical support to the key role noticing, FonF and model texts play in the L2 development of individuals in this specific EFL context and within the writing-to-learn approach. Furthermore, it aims at tentatively exploring what learners' attitudes towards models are and how these can affect their effectiveness as language learning tools. Finally, the study is also an attempt to see whether the use of models as a form of feedback should be encouraged in foreign language settings.

Unlike in Hanaoka's (2007) study, in the present one there is only a delayed revision instead of an immediate and delayed one, since this enabled us to focus on those features that remained longer in the participants' memories. In addition, it includes a control group and two distinct treatment groups in terms of type of noticing (i.e. a guided noticing group (GNG) and an unguided noticing group (UNG)) in order to investigate the impact noticing and noticing with understanding (Qi and Lapkin, 2001) can have on language development. Finally, the participants in the present study were asked to fill in an exit questionnaire to analyse their beliefs towards models and the effect of these in the language learning process.

The rest of the paper is divided into six parts and structured as follows: after this introductory section, the definition and presentation of key concepts (2.1) and an overview of what other researchers have written within this field (2.2) is provided in section 2. Research questions and hypotheses (2.3) are also included in this section. Next, section 3 describes the context and participants (3.1), tasks and procedures (3.2) and the analysis of the data (3.3). Section 4 presents the findings of the study, while section 5 centres on their discussion. Section 6 includes the limitations of the study and further research avenues, and finally, section 7 points out the main conclusions and the pedagogical implications.

## 2. Literature review

This section briefly explains and discusses the key concepts and relevant aspects of the study and provides an account of previous research in the field of writing-to-learn.

### 2.1 Key concepts

Three main concepts will be defined and discussed in this part, namely: output, corrective feedback and focus on form.

#### 2.1.1 Output

Output refers to the language learners produce either in speaking or writing. In the field of first language (L1) and second language (L2) acquisition, it has been often assumed that output is nothing more than a sign of already acquired L2 competence and that it does not serve any significant function in the process of language acquisition other than determining whether learning has been effective (Krashen, 1982). That is to say, output has been a mere synonym of what learners have learned. However, this limited view of output has been questioned by the work by Swain (1995 *et passim*) and her Output Hypothesis, which argued for a more active role of output in the whole process of language learning and teaching.

The hypothesis was essentially formulated as an addition to Krashen's (1982, 1985) Input Hypothesis, which claimed that the only way in which learners could acquire language was by being exposed to comprehensible input. Nevertheless, several years of research on the Canadian immersion programmes led Swain (1995) to conclude that, albeit invaluable, comprehensible input was not sufficient for developing near-native competence. She observed that learners in these programmes failed to acquire a native-like level of accuracy in spite of the continuous exposure to the L2 and argued that one of the main reasons behind this was that they engaged too little in language production. Swain (1995) posits that learners need to be encouraged to deliver

messages which are not only conveyed, but conveyed precisely, coherently and appropriately, because this pushes them to process language more deeply (moving from semantic processing to syntactic processing) and with more mental effort than input. According to the author, this more active deployment of cognitive resources contributes to language learning, thus, suggesting that output is not the result of the acquisition process, but rather a step in it. In what follows, the four main functions of output identified by Swain (1995) will be described.

**a) The fluency function:** According to Swain, language learners need opportunities to use the language in meaningful contexts in order to automatize access to the existing L2 knowledge and, thus, develop a fluent productive performance. This obviously requires output. That is to say, according to Swain producing the language enhances fluency. However, fluency and accuracy are two distinct dimensions of language performance and improvement in one does not necessarily lead to improvement in the other (Robinson, 2011; Skehan, 2009). That is why three other functions were proposed that relate more to accuracy.

**b) The noticing/triggering function:** In producing the TL learners may notice a gap between what they want to say and what they can say. In other words, when producing output in an L2 (be it orally or in writing) learners are able to recognize autonomously what they know, what they do not know, and what they know only partially. They notice the limitations in their linguistic resources. There are several levels of noticing: learners may realize that the form they produce is different from the TL, that is, they may “notice the gap” (Schmidt & Frota, 1986), or they may “notice the hole”, when they become aware that they cannot convey the message they want accurately in the TL (Swain, 1995). Due to the heightened sense of problematicity, these situations may induce learners to engage in more focused attention with subsequent input; they might try to analyse the way in which the TL expresses the message they just had difficulty in conveying (Izumi, 2002). This involves cognitive processes which “generate linguistic

knowledge that is new for learners or that consolidates their existing knowledge” (Swain, 2005: 474). That is to say, it involves cognitive processes that might trigger IL development. This important role of noticing should be considered on the basis of Schmidt’s (1990, 2001) Noticing Hypothesis, which, will be further explained below, but, essentially, claims that learners must notice input consciously for it to become intake. From this perspective, output and noticing can be factors accelerating the natural L2 acquisition process.

Here is an example taken from Lapkin, Swain and Smith (2002: 492) where a pair of English students in a Canadian immersion programme recognizes the limitations to find the French verb *se recoucher* (1) and modifies the output on the basis of the feedback received (2).

(1) The learners discussed:

M: Elle veut... elle veut... elle redorme [non-existent] ou quelque chose.

*She wants... she wants... she goes back to sleep or something.*

S: Non, ... ah... elle retourne au lit.

*No... ah... she returns to bed.*

M: Oh! Oui, oui, oui.

*Oh! Yes, yes, yes.*

S: Elle... Retourne... au lit.

*She... returns... to bed.*

In the reformulation received, the students noticed the solution provided *se recoucher*. In their think aloud<sup>1</sup> they said:

---

<sup>1</sup> A think aloud is a data collection technique in which participants are asked to say out loud any thoughts which come to their mind while completing a task. It is mainly used to obtain more direct access to the learners’ mental processes.

(2) S: We had *retourne au lit*.

M: *se recouche* which is...

S: It makes more sense.

M: Yes. [repeating to herself] *se recouche*.

S: *retourne au lit* would me more like wake up to go to the bathroom,

M: Yeah. It's like come back.

S: but she was, she just she woke up and she was still in her bed and she went back to sleep.

M: Yeah and she went back to sleep [in a lower voice]. So it makes more sense.

This example is revealing because it shows that the production of output pushed the learners to notice the gap between their expression *retourner au lit* and the TL form *se recoucher* provided in the feedback. Indeed, this noticing led to intake, as both students used the item correctly in the post-test.

**c) The hypothesis-testing function:** It has been argued that output, especially erroneous output, can reveal hypotheses held by learners about how the TL works. That is, output represents the learners' best guess of how something should be said or written in the TL. In this sense, output is regarded as a trial run. Interestingly, learners' language productions may sometimes invoke feedback from the interlocutor, which can help them judge the comprehensibility and linguistic well-formedness of their IL and modify or reprocess it, if necessary. The assumption is that this process of modification contributes to the acquisition of the L2. According to Swain (2005), when modifying their output learners experiment with new language forms and structures, thus expanding their IL to meet communicative needs. Consider the following example (3):

(3) Learner 1: John arrive, arrove, arrive or arrove?

Learner 2: arrove is in past

Learner 1: arrove airport. Or arrived.

Learner 2: arrove, is in past

Learner 1: I mean arrove or arrived.

Learner 2: arroved the airplane

Learner 1: Arrived or arroved?

Learner 2: Arrove

Learner 1: Arrove the airport at 8:30 am

(Adams, 2007: 48-49)

In the conversation between the two students, Learner 1 tries out four different past forms for the verb *arrive*, as (s)he seems not to be sure whether the verb is regular or irregular. From this hypothesis testing, (s)he receives feedback from Learner 2 and learns a new, although incorrect, form in the TL.

**d) The metalinguistic (reflective) function:** Swain (1995) also argues that output processes do not only enable learners to test their hypothesis, but also to reflect on them using language. This metalinguistic reflection, especially when done collaboratively, is thought to deepen learners' awareness of form and form-function relationships in communicative and meaningful contexts. In (4), for instance, two advanced learners reflect on whether *inclined* should be followed by *to* or by *-ing* and, as can be observed, there is metalanguage in their conversation.

(4) Learner 1: men are less incline... it has to be an adjective.. inclined to confess, you are inclined to do something

Learner 2: to confession ..

Learner 1: to confess...



Learner 2: but after a preposition...

Learner 1: to confess ... what?

Learner 2: to is a preposition...

Learner1: yeah...

Learner 2: so it should be followed by ing

Learner 1: inclined to confessing...

Learner 2: yeah

Learner 1: no, because to is part of the second verb .. inclined to confess .. yeah

Learner 2: ok, I trust you

(García Mayo, 2002:329)

To summarise, Swain's (1995, 2005) Output Hypothesis claims that output can foster language acquisition by enabling learners to try out and extend their IL capabilities. In doing so, learners may become conscious of problems in their IL through internal and/or external feedback, which may prompt them to generate new hypotheses by searching existing knowledge or by paying more focused attention to relevant input. Language production is, therefore, seen as an important means possibly influencing what aspects of comprehended input become intake. Within this view, output is no longer a mere product of acquisition, but an active component in the SLA process. However, note that Swain (1995) also added that output on its own does not necessarily serve these functions, and acknowledged the role of CF by stating that students should receive sufficient feedback regarding the extent to which their messages have been successfully conveyed in terms of accuracy, appropriateness and coherence. Thus, the following section will delve into the field of CF.

### *2.1.2 Corrective Feedback*

As Gass and Mackey (2006:7) report, feedback is generally defined as a form of negative evidence or information that a particular utterance is deviant from the TL norms. As such, it has

been considered a central and pedagogically relevant aspect in second language research, teaching and learning. However, the literature has not been unequivocally conclusive about its role in SLA and, therefore, there is an ongoing debate on its potential to improve students' accuracy. While some researchers (Truscott, 1996, 2007) have claimed that feedback can be frustrating, harmful for students' fluency and unhelpful to improve performance in subsequent writings, many others have defended that feedback, especially coming from the teacher, is highly valued by students (Hyland, 1998; 2011) and that it positively affects language acquisition (Bitchener, 2012; Bitchener & Knoch, 2010; Chandler, 2003; Ferris, 2010).

All these claims have been based on several measures of CF effectiveness such as performance in immediate or delayed post-tests, noticing and uptake. Yet, note that the last one has been regarded as a controversial measure of acquisition since the notion has not been well-established in the SLA literature (Sheen, 2004). Some researchers, such as Lyster and Ranta (1997), have defined uptake as a discourse phenomenon uttered by the learner as an immediate reaction to teacher feedback. Thus, some studies (Ohta, 2000) have questioned its relation to the psycholinguistic processes involved in language acquisition. Nonetheless, other scholars, including Santos et al. (2010) and Storch and Wigglesworth (2010), have operationalised uptake as the amount of corrections incorporated by the participants, thereby, making it a potential indicator of CF effectiveness. In this study, the term was defined following the second approach.

Apart from the debate on the effectiveness of CF, it is noteworthy that even among those who advocated for the contribution of written feedback to language development, the question of what form error correction (EC) should take remains without a clear answer (See Ellis, 2009 for a typology of written CF). A considerable number of studies have investigated the effectiveness of direct and indirect feedback strategies (Bitchener, 2008; Bitchener & Knoch, 2010; Chandler, 2003; Coyle & Roca de Larios, 2014), mainly focusing on the former. In spite of the positive

results these works obtained with a limited number of linguistic structures, such as the English article system (Bitchener, 2008), some researchers have criticized the usefulness of EC on the basis of the lack of clarity, precision, and consistency with which it has been used (Ellis, 2009), the frustration and confusion it can cause among students (Hyland, 1998) and, the passive reflection and superficial processing it may involve (Adams, 2003; Qi and Lapkin, 2001). As Qi and Lapkin (2001) argue, EC usually requires minimal processing on the part of the learner, and thus, it does not provide optimal conditions to help students *notice* the gap between their IL and TL. This fact has led some authors to suggest this type of EC may not push learners to test their hypotheses in depth (Adams, 2003). As a consequence, and assuming the tenets of the Output Hypothesis (Swain, 1995) and the Noticing Hypothesis (Schmidt, 1990), researchers (Martínez Esteban & Roca de Larios, 2010; Qi & Lapkin, 2001) have claimed that other forms of feedback which engage L2 learners in deeper processing should be used, at least as partial alternatives. Two of these are reformulations (Hanaoka, 2006b; Qi & Lapkin, 2001; Santos et al., 2010; Yang & Zhang, 2010) and modelling (Coyle & Roca de Larios, 2014; Hanaoka, 2006a; Hanaoka, 2007; Hanaoka & Izumi, 2012) which, contrary to the itemized nature of EC, are textual in orientation.

A reformulation is a procedure which involves rewriting the student's entire text to make the language seem as native-like as possible, yet maintaining the content of the original intact (Ellis, 2009). In contrast, models are good examples written by native speakers or teachers considering learner's age and proficiency as well as the composition's genre and content, but they are not contingent on learner output (Hanaoka, 2006b). The implicit signalling of errors in both feedback forms is supposed to push learners to actively attempt to identify and understand their mistakes, a process that may lead to deeper processing and, ultimately, to language development (Sachs & Polio, 2007). It is assumed that when comparing the reformulated versions or the models with their own writings, learners will notice both similarities and differences between their IL and TL,

thus confirming and/or refuting their hypotheses about the L2. Moreover, these types of feedback provide a wide range of lexical, syntactic and, discourse alternatives, which are of paramount importance for EFL learners whose exposure to native-forms is often limited (Manchón, 2009).

Despite these shared benefits, both feedback types differ in some other respects. First, models serve the dual role of addressing both form and meaning: they do not only provide learners with new vocabulary and expressions which may enrich their repertoire, but also with alternative ideas and contents on which students can draw to develop and express new meanings (Hanaoka, 2006b). Reformulations, in contrast, do not prove beneficial in this sense, as they have to remain as faithful as possible to the original meaning. Second, as Hanaoka (2006a) points out, when producing output learners might resort to reduction strategies; that is, they may avoid including some problematic elements because of their limited linguistic resources. These features may be termed “covert features” and they may result in topic avoidance. While it is quite difficult for reformulations to address these forms, models, being independent from learners’ output, have the potential to provide solutions to them (Hanaoka & Izumi, 2012). Finally, even if reformulation has gained recognition as a technique leading to L2 development (Yang & Zang, 2010), it is time-consuming and, thereby, inapplicable in classroom settings where the student ratios are high and the time allotted to English in EFL contexts limited. In the light of these drawbacks, modelling appears to be a solution.

This technique is not new in language teaching contexts. Yet models have traditionally been regarded as ideal products to be imitated and, so, they have been usually presented at the initial stage of the writing process. Nonetheless, following Hanaoka (2007) and the basic tenets of the Output Hypothesis (Swain, 1995), in this study, models were provided at a later stage as a resource to solve problems that participants themselves had noticed while writing. It was assumed that this would engage learners in more focused attention and deeper processing. In

spite of such potentially useful role argued for models, caution is needed. As Hanaoka and Izumi (2012) stated, learners, especially those with a lower proficiency level, may have difficulties in finding differences between the model (or reformulated version) and their own text, unless some kind of enhancement is provided. In addition, if the language contained in the model is overly complex its usefulness as a scaffolding<sup>2</sup> strategy is diminished (Coyle & Roca de Larios, 2014; Swain & Lapkin, 2002). Therefore, tailoring the text to learners' proficiency level is of paramount importance to encourage noticing. Finally, the interactive and social dimensions of feedback should not be overlooked, since CF can only have an impact if students attend to it. As Hyland (2010) pointed out, active student participation is crucial to exploit the LLP of feedback. Nevertheless, the literature on learners' attitude towards and engagement with models is scarce. That is why the present study has also tried to approach this aspect as a factor possibly influencing the effectiveness of models, although it was not its primary goal.

### *2.1.3 Focus on Form*

Keeping in mind the functions of output and the role of CF in language learning, we should not forget that for this processes to be effective learners need to attend to form. As Schmidt (1990) claims, there is no IL development without conscious attention to form. In fact, the main tenet of the FonF approach is that shifting learners' attention to linguistic structures while engaged in communicative tasks is beneficial for SLA. However, FonF is amenable to a number of interpretations.

The original notion was proposed by Long (1991:46) and it referred to the cognitive processes by which learners address the linguistic problems that 'arise incidentally in lessons whose overriding focus is on meaning or communication'. That is to say, FonF draws students' attention

---

<sup>2</sup> This notion linked to the constructivist approach to language learning (Vigotsky, 1978) refers to a process of interaction where a More Knowledgeable Other – peer, teacher, parent, caretaker- provides a less skilled learner with guided support to help him or her solve a problem, complete a task or build new knowledge.

to linguistic elements such as words, collocations, grammatical structures and pragmatic patterns in meaning-focused communication. Importantly, Long regarded it as a crucially incidental process. Yet, other authors such as Ellis (2001) provided a more inclusive definition of the term, including both planned and unplanned discussion of form (see Ellis (2016) for updated information on this topic). Regardless of this difference, we will refer to FonF as the shift in attention from content to specific linguistic code features which takes place in a lesson.

This processing has been traditionally operationalized in the literature in the form of Language Related Episodes (LREs) (Hanaoka, 2007; Swain & Lapkin, 1995). According to Swain and Lapkin (1995), LREs are any segment of the dialogue where learners question their language use, correct themselves or others, or refer to the language they are producing or to a problem they encountered while writing or speaking. In other words, it refers to parts of learner interaction where students focus explicitly on linguistic items while carrying a communicative task. As will be further explained below, in this study LREs were operationalized in the form of individual note-taking. Nevertheless, they have mostly been studied in situations where learners were orally engaged in different tasks (Azkarai & García Mayo, 2015; Basterrechea & García Mayo, 2013; García Mayo & Azkarai, 2016; Lapkin et al., 2002; Swain & Lapkin, 1995). Excerpt (5) exemplifies an oral LRE where two participants focus on the third person singular present morpheme *-s*:

(5) Learner 1: she... knows...but she... only...sees....them (*uttered while writing*)

Learner 2: see them

Learner 1: *esto está mal, no?* (this is wrong, isn't it?)

Learner 2: she only see them

Learner 1: no sees

Learner 2: she only see from school and the other ones are people she know

Learner 1: She knows

Learner 2: She know

Learner 1: *¡a ver!, tercera persona* (pay attention!, third person)

(Basterrechea and García Mayo, 2013: 35)

Researchers have investigated the influence of different variables such as learners' proficiency and task type on the occurrence, nature and quality of LREs. As for the former, several studies (Hanaoka, 2007; Swain & Lapkin, 1995) have revealed that higher proficiency learners produce more LREs and that they tend to focus more on formal aspects than lower proficiency learners, who mainly pay attention to meaning. In addition, Qi and Lapkin (2001) reported that more advanced learners not only engaged more in noticing but that they understood better the nature of the noticed gaps.

Task type and task modality also affect learners' attention to form. For instance, Azkarai and García Mayo (2015) and García Mayo and Azkarai (2016) found that participants produced LREs more frequently in tasks that require both an oral and written component. Moreover, it seems that text reconstruction or editing tasks elicit more grammatical LREs than free writing tasks. For example, Basterrechea and García Mayo (2013) reported that 51% of the LREs produced during a collaborative dictogloss task (Wajnryb, 1990), where the purpose is to reconstruct a passage as accurately as possible, dealt with grammar. In contrast, Swain and Lapkin (1995), Hanaoka (2007) and Martínez Esteban and Roca de Larios (2010), among others, reported that an overwhelming number of all the LREs produced by learners engaged in picture description tasks were lexical.

Besides, it seems that learners generally focus on form when they start having communication problems, so it is at that particular moment when they are more ready to understand the function

or the meaning of the new form. Furthermore, several studies have shown that when learners face a communicative problem, they are more eager to receive feedback (Martínez Esteban & Roca de Larios, 2010; Sachs & Polio 2007). Therefore, we may hypothesize that most of the CF provided at this particular moment would be assimilated. Consequently, in this study models were provided immediately after the learners had written a text in response to a series of pictures.

Finally, empirical research has shown that the quality of the FonF experienced by a learner while processing feedback can have direct implications on learning outcomes. Data in Qi and Lapkin (2001) and Sachs and Polio (2007), among others, have indicated that noticing without understanding or noticing for no reason does not have the same effect on IL development as does noticing with understanding. It seems that the use of metalanguage and the provision of reasons while noticing show a deeper level of awareness, which in turn, leads to more accurate revisions and more intake (Qi and Lapkin, 2001; Sachs and Polio, 2007). However, it still remains unclear how learners can be encouraged to go beyond the level of simple detection and engage in deeper processing of input. For this reason, in this study two types of noticing were distinguished: a guided noticing group which followed the noticing with understanding method designed by Santos et al. (2010), and an unguided noticing group, which was asked to merely list the features noticed.

## *2.2 Previous research*

This section provides an overview of previous empirical studies which have shed light on the LLP of writing by engaging learners in individual or collaborative writing tasks to explore the functions of language output mentioned above and/or the issues of attention and noticing.

Cumming (1990) was among the first to acknowledge the potential role of composition writing in L2 learning. He analysed the think-aloud protocols of 23 adult francophone English as a



Second Language (ESL) learners engaged in two individual writing tasks. The main aim of the study was to obtain detailed descriptions of the noticing and metalinguistic thinking of the participants, to see how this can be influenced by learner-related and task-related factors, and to infer the way in which learners' thinking was linked to SLA processes. These analyses were regarded by Cumming as a preliminary step before establishing a causal effect between composition writing and language learning. He concluded that "composition writing might function broadly as a psycholinguistic output condition wherein learners analyse and consolidate L2 knowledge that they have previously (but not yet fully) acquired" (Cumming, 1990: 483).

Later, this study was partially replicated by Swain and Lapkin (1995). They also focused on composition writing and think-aloud protocols, but with adolescent learners of French. They investigated three main aspects: First, whether language learners engaged in noticing processes while writing, second, which actions the participants reported implementing to solve the problems encountered, and, third, whether they conducted grammatical and syntactic analyses to overcome their linguistic difficulties. They found that noticing while writing led to modified output and suggested that certain cognitive processes (including grammatical analysis) involved in problem-solving might facilitate L2 learning. In addition, the results showed that learners with a higher proficiency paid more attention to grammar, hence, suggesting that proficiency may be a variable affecting linguistic awareness.

These two pioneering studies provided rich descriptions of noticing while composing and set the path to numerous empirical studies analysing how the act of writing itself and the processing of feedback can promote language development. For example, Qi and Lapkin (2001) conducted a case study where two Mandarin speaking ESL learners engaged in a three-stage writing task (i.e. describing a picture, comparing it with a reformulated version, and revising the original composition). Throughout the process, the participants were asked to verbalize their thoughts by

means of think aloud protocols. The researchers showed that noticing in the composing stage influenced what learners noticed in the feedback processing stage and that the quality of noticing directly influenced the revisions made in the final written product. Qi and Lapkin (2001:294) claimed that noticing without understanding (i.e. perfunctory noticing) did not have the same impact as noticing with understanding (i.e. substantive noticing) on learning in the L2. Therefore, the researchers suggested that L2 writing pedagogy should not only engage in promoting noticing but in finding ways to promote a higher quality noticing. The present study addressed this issue by comparing two different forms of noticing: on the one hand, a guided noticing with which we tried to promote a deeper processing of feedback (Santos et al., 2010) and, on the other, an unguided noticing which consisted in the mere detection of differences between the models and the original composition (See 3.2 tasks and procedure below). In addition to the conclusions mentioned, Qi and Lapkin (2001) stated that models of native-like writing may be more useful than the traditional EC. Moreover, they revealed that writers with a higher level of L2 proficiency noticed more features and understood their nature better than participants with a lower proficiency. Interestingly, this finding has been repeatedly reported in the literature (Hanaoka, 2007; Lapkin et al., 2002).

Lapkin et al. (2002) examined a set of collaborative dialogues which occurred in a multistage composition task where 8 7th grade (aged 12-13) French immersion students worked in dyads of different proficiency. Each pair was asked to write a story, to notice the differences between their text and the reformulated version, to reflect on their noticing in a stimulated recall<sup>3</sup> session and to rewrite the original story individually. The researchers found that although all participants performed better in the post-test, stronger pairs produced richer collaborative dialogues in terms of LREs and that their noticing of the reformulations was more detailed.

---

<sup>3</sup> Stimulated recall is an introspective technique used to explore learners' thoughts about the learning process. Participants are presented with a stimulus (e.g. video/audio tapes or drafts of a composition) and asked to recall thoughts entertained while they were carrying the task.

A closer look at a subset of these data enabled Swain and Lapkin (2002) to investigate the reactions of two French immersion learners to reformulated writing. They showed that participants considered the native-like reformulation too sophisticated and that they would have preferred to receive a reformulated version from someone “at about the same level or a little higher [...] to better understand the words” (Swain & Lapkin, 2002: 299). This illustrates the abovementioned need for tailoring reformulations and models to learners’ proficiency level in order to guarantee successful scaffolding. The researchers also found that task repetition, noticing and participation in stimulated recalls might affect positively target like usage in subsequent output.

With the aim of investigating the effect of each of these factors separately, Adams (2003) replicated and extended Swain and Lapkin’s (2002) study. In her research, Adams randomly assigned 56 L2 Spanish learners to three groups: (i) task repetition, where students repeated the task without additional treatment, (ii) noticing, where participants repeated the task and compared their original output to a native-speaker reformulation, (iii) noticing + stimulated recall, where participants engaged in a stimulated recall session after completing the same comparison stage as the previous group. Analysis of the data showed that there were quantitative between-group differences in the participants’ post-tests, with the noticing and noticing + stimulated recall groups significantly outperforming the task repetition group. In turn, the noticing + stimulated recall incorporated more target-like forms than the noticing group. However, this last difference did not reach statistical significance. In view of these results, Adams (2003) pointed out that noticing surface differences between IL and TL can facilitate learning, and that this learning can be enhanced by means of follow-up activities, such as stimulated recalls. Finally, based on the large within group variations observed, Adams (2003) claimed that individual differences such as motivation may regulate the effectiveness of noticing in facilitating language development. This issue will be further discussed below.

Another piece of empirical research investigating the effectiveness and noticing function of reformulations was Sachs and Polio's (2007) study. They compared written EC and reformulations as two means of improving learners' grammatical accuracy on a three-stage writing task (i.e. composition, comparison and revision). The results demonstrated that participants in the error condition performed significantly better than those in the reformulation condition. Note that this result contrasts with Qi and Lapkin (2001), who hypothesised that reformulations were possibly more useful than EC. Furthermore, on the basis of think aloud protocols, Sachs and Polio (2007) concluded that the noticing of feedback was related to the accuracy of subsequent revisions. However, they also observed that learners who were asked to think aloud while processing the feedback produced less accurate revisions. This suggests that even if verbal protocols might be of help to study learner-internal processes, they should be employed and interpreted with care. This is one of the reasons why the current study opted for note-taking as the form to operationalize LREs.

The studies reviewed so far are of relevance to the current piece of research mainly because of their focus on the LLP of writing and on the cognitive processes active while composing. However, most of them opted for reformulations instead of for models as the form of feedback provided to participants, possibly because of the little research conducted on the latter. Hanaoka (2006b) was among the first to investigate the potentially different roles played by models and reformulations in a small-scale exploratory study. The participants were two Japanese EFL learners with different proficiency levels, who, after having written a composition in response to two sets of pictures, were exposed to both forms of feedback during the comparison stage. The results indicated that models and reformulations play complementary roles as feedback techniques; while models may serve to address alternative forms and to develop the original content, reformulations might promote noticing of linguistic inadequacies. Furthermore, Hanaoka (2006b) found out that the two learners benefited differently from the two types of

feedback; whereas the more advanced learner mainly incorporated features unique to the models, the less proficient participant incorporated features which were present in both types of texts. Nevertheless, these results have to be treated with caution given the size of the sample.

Hanaoka (2006a) also added support to the effectiveness of models as feedback tools by focusing on their specific role of providing solutions to covert problems. These are problematic features which learners omit or partially avoid incorporating in their compositions owing to linguistic restrictions. In this study 37 female Japanese learners of EFL were asked to write a narrative in response to a picture prompt (Stage 1), to compare their output with two model texts (Stage 2) and to revise their original text twice, immediately after the comparison (Stage 3) and two months later (Stage 4). The results showed that the models offered a solution to more than 70% of the covert features and that the participants noticed and incorporated them in the following revisions, which suggests that native-like modelling is a valid pedagogic tool that deserves further exploration.

Using the same data set as in the study above, Hanaoka (2007) examined what Japanese college students of different proficiency levels spontaneously noticed while composing a story in response to a picture prompt (Stage 1) and while comparing their original output to two native-speaker models (Stage 2). He also studied how the nature of such noticing affected subsequent learning in their immediate (Stage 3) and delayed revisions (Stage 4). For the first research question, participants were asked to note down all the problematic features they noticed (henceforth PFNs) in the first stage as well as the solutions noticed in the models (i.e. Features Noticed – henceforth FNs) in the second stage. These notes were later tallied and analysed grouped in four different categories depending on their nature; namely, lexical, grammatical, content and other features (for a more detailed description of each of the categories see section 3.3 Analysis). As for the second research question, Hanaoka (2007) examined the number of FNs

which were incorporated (henceforth FNIs) by the participants in the immediate and delayed revisions considering two variables: First, the nature of the FNIs (i.e. lexical, grammatical, content or other) and second whether the FNIs were related to the FNs noticed in Stage 2 or not. The findings indicated that the participants were able to identify the linguistic features they lacked, to autonomously find solutions in the feedback provided and to integrate them into further revisions. This confirms that output plays a useful role in helping learners notice their linguistic gaps as well as in facilitating subsequent learning of features needed to fill those gaps in. What is more, Hanaoka (2007) found out that the participants overwhelmingly identified lexical features and that, interestingly, those which were related to the problems participants had faced through output were incorporated at a higher rate and retained longer. Therefore, it might be suggested that noticing the hole, not only serves as an important stimulus for noticing the gap, but also facilitates retention of the solutions in short and long-term memory. Finally, Hanaoka (2007) revealed that more proficient learners noticed significantly more features than the less proficient ones and that they attended to form better.

This study was recently replicated by Martínez Esteban and Roca de Larios (2010) who investigated the role of models in individual and collaborative EFL writing. The participants were 17 Spanish secondary-school pupils at a low intermediate proficiency level who completed a three-stage writing task. The results revealed that the students noticed mainly lexical problems at the composing stage and that they could not find sufficient solutions to those problems in the models. According to the authors, the reason behind this phenomenon could be the large linguistic gap between the written texts and the models, which made the noticing exceedingly challenging. For this reason they highlighted the necessity of adapting the models to the participants' proficiency level. In addition, Martínez Esteban and Roca de Larios (2010) found that participants, especially those who wrote collaboratively, were more concerned with the contents of the texts provided and that, indeed, they incorporated a reasonable number of them in

their revisions. This finding, hence, supports the role of models as a form of written feedback.

In the same line, Yang and Zhang (2010) analysed the effectiveness of reformulation and model texts in a three-stage (composing-comparison-revision) collaborative writing task with 10 Chinese EFL university students. The study revealed that participants focused more on the lexical and form level than on the macro level (e.g. organization and development of ideas) in the composing and comparison stages, as was found in Hanaoka (2007). Furthermore, the students appreciated being provided with a native model and a reformulation to broaden their range of language input. Yang and Zhang (2010) also observed an improvement when comparing the original and revised versions, even though some errors made at the pre-test were repeated at the post-test. This may simply mean that the students did not notice all the holes in their IL or were not capable of incorporating them in their revisions. At the end of the study, the researchers concluded that a model text can be an important teaching tool since it can offer students ‘a good sample of native writing not just for a specific sentence but for the whole discourse’ (Yang & Zhang, 2010: 480).

Similarly, Coyle and Roca de Larios (2014) explored the role played by EC and model texts on the noticing and subsequent revisions of written output. The study was carried out with 11- and 12-year-old EFL children who engaged in a three stage collaborative writing task. In line with previous studies, the results indicated that children noticed and incorporated mainly lexical features. Moreover, learners in the EC condition were found to report more noticing of grammar at the comparison stage, which later resulted in higher accuracy in their revisions. The participants in the model’s condition, in contrast, directed most of their attention toward language chunks and content. The researchers concluded that both types of feedback should be combined in the classroom as these were found to promote the noticing of different features: EC may be more useful to direct attention to form, while models can provide lexis and expressions

beyond learner's linguistic repertoires. In addition and contrasting with previous studies, Coyle and Roca de Larios (2014) found that it was the lower proficiency children who noticed a greater number of features, thus, suggesting that the relationship between noticing and L2 proficiency may not be as straightforward as it was thought.

Finally, Cánovas Guirao, Roca de Larios and Coyle (2015) studied the effectiveness of model texts as a form of CF with children aged 10-11 who engaged in a three stage writing task (composition-comparison-revision). The findings indicated that higher proficiency learners noticed and incorporated more features than students with a lower proficiency. Moreover, in line with Hanaoka (2007), participants were found to focus mainly on lexis. Note also that the improvement of both the experimental and control group in Canovas Guirao et al. (2015) led the scholars to acknowledge the potential influence of task repetition. Considering that all the studies mentioned above analysed the performance of experimental groups only, the present work opted to further explore this issue including a control group.

To sum up, the above-mentioned studies have shown that pushing learners to produce output can promote noticing and that providing them with appropriate written feedback can lead to subsequent learning in EFL settings. However, as pointed by Manchón (2011), there is still need to do research on EFL writing contexts and to further explore the extent to which native speaker modelling independent from learner's original output can lead to language learning, since the number of studies on the use of models is rather scarce and, indeed, those which have analysed it frequently lacked a control group. Thus, the aim of this research project is to address this gap.

Furthermore, it has been suggested that students' motivation and attitudes affect the extent and depth of processing of feedback, thereby, enhancing or limiting its LLP (Hyland, 2011; Kormos, 2012; Manchón & Roca de Larios, 2011). As Kormos (2012) points out, noticing gaps in one's



IL and engaging in problem-solving processes are effortful tasks which require high levels of motivation, intrinsic interest in language learning and positive attitudes towards feedback. As a result, it is also necessary to investigate learners' own beliefs towards models to better understand their efficacy as language learning tools.

In this sense, although several studies have analysed students' attitudes towards the learning potential of writing in general (Manchón & Roca de Larios, 2011) and their engagement with certain types of feedback such as reformulations (Storch & Wigglesworth, 2010) and EC (Hyland, 1998; Hyland, 2011), to our knowledge the only steps taken in analysing learners' beliefs on the use of models as a form of feedback are Hanaoka's (2007) measurement of the participants' willingness to receive the models and Yang and Zhang's (2010) short interviews with the participants to elicit comments on reformulation and modelling. However, these scholars did not focus on investigating how the learners' attitudes could affect their performance in subsequent revisions. Therefore, this study decided to address this gap by including a background questionnaire (Appendix J) on the use and perceived effectiveness of models.

### *2.3 Research questions and hypotheses*

On the basis of the literature review briefly presented above, the present study aimed at exploring (i) the aspects of language learners notice while engaged in spontaneous FonF in L2 writing and (ii) the effectiveness of native speaker models as a feedback tool, taking into consideration the participants' level of proficiency, depth of noticing and attitude. In order to do that, the following research questions were entertained:

- 1) What aspects of language do L2 learners of different proficiency levels notice while writing a composition on their own?

- 2) What aspects of language do L2 learners of different proficiency levels notice as they compare their texts to two models? Are there any differences between the students who received some guidance on noticing and those who did not?
- 3) What are the effects of Stage 1 (Composing) and Stage 2 (Comparison) noticing on subsequent revisions?
- 4) What is the participants' attitude towards models as a form of feedback? How does this affect their noticing and performance in subsequent revisions?

In order to answer these questions we formulated the following hypotheses:

- 1) Hypothesis 1: Participants will mainly notice lexical items and there will be between group differences in the amount of PFNs produced, with more proficient learners reporting more PFNs (Hanaoka, 2007; Martínez Esteban & Roca de Larios, 2010; Yang & Zhang, 2010).
- 2) Hypothesis 2: Participants will notice the gap when comparing their output to the models and, although they will mainly pay attention to lexical items, content FNs will also be frequent (Coyle & Roca de Larios, 2010; Hanaoka, 2007; Yang & Zhang, 2010). Moreover, we expect the GNG to notice more FNs than the UNG (Santos et al., 2010; Qi & Lapkin, 2001).
- 3) Hypothesis 3: Noticing and models will lead to the incorporation of noticed elements in subsequent revisions (Hanaoka, 2007; Martínez Esteban & Roca de Larios, 2010; Yang & Zhang, 2010).
- 4) Hypothesis 4: Participants will have a positive attitude towards models and this will impact on their role as an effective form of feedback (Hanaoka, 2006a; Hanaoka, 2007; Yang & Zhang, 2010).

The present piece of research was conceived as a partial replication of Hanaoka (2007) and goes in line with most of the research briefly reviewed above. As stated at the beginning of this section, the main focus will be on (i) what the participants notice while composing and comparing their outputs with two models and (ii) what they incorporate in subsequent rewritings. The participants will engage in a three-stage individual writing task (composing-comparison-delayed revision) and their LREs will be operationalized in the form of note-taking. Importantly, a unique revision will be used instead of two revisions (immediate and delayed revisions), as in Hanaoka (2007). Moreover, the effect of other variables, such as, guidance on noticing and affective factors will also be analysed. Finally, unlike the original, this work will be carried out with secondary and high school students and will include a control group. All these methodological aspects will be described in more detail in Section 3 below.

### **3. The study**

#### *3.1 Context and participants*

The present study was carried out in the eastern part of Gipuzkoa (Basque Country, Spain), where Basque and Spanish share official status. Two schools participated in it: a public secondary school (ages 12-16) and a semi-private high school (ages 16-18), both of which had Basque as their main language of instruction. The project was conducted with the approval of the headmaster, the teacher board and the parents of all the students involved (see Appendix A for the consent form).

The participants were 60 Basque-Spanish bilinguals (31 females and 29 males) who received 3-5 hours of English per week at school. Their age ranged from 13 to 17 (mean: 14,78 s.d: 1,53) and they were all either in 2<sup>nd</sup> year of secondary school or 1<sup>st</sup> year of high-school. The sample studied was a subset of a larger sample of 173 students, out of which 163 completed the three research stages. For the selection of the group analysed, the total number of hours of English received per week and the amount of exposure received by the time of the study were not considered. Instead, it was the students' proficiency that was taken into account. In order to determine their level, all the students took a proficiency test (Oxford Placement Test, Versions 1 or 2), according to which five levels were established: beginner (N=14), elementary (N=75), lower intermediate (N=58), upper intermediate (N=15) and advanced (N=1). Given that the elementary and lower intermediate levels were the most common, and that the other levels may have not been statistically representative, 30 participants in each of these two groups were chosen semi-randomly: All participants in the elementary group were coming from the secondary school while those in the lower intermediate studied were in the high school. In addition, gender issues were also taken into consideration, thus, trying to include the same number of males and females in the elementary and lower intermediate groups. Moreover, at each proficiency level three more groups were distinguished: (i) a control group, which completed the composing and rewriting

stages, but did not receive the models, (ii) a guided noticing group (GNG), which completed the three stages and used a guided form of noticing in the comparison stage, and (iii) an unguided noticing group (UNG), which completed the three stages, but used a non-guided form of noticing in the comparison stage. For this division, the schools' own class arrangements were maintained. Each class was randomly assigned to one of the groups by the researcher. Table 1 below summarises details of the participants:

	<b>Elementary</b>	<b>Lower Intermediate</b>
<b>Control group</b>	N=10 (5♂, 5♀)	N=10 (5♂, 5♀)
<b>Guided noticing group</b>	N=10 (5♂, 5♀)	N=10 (4♂, 6♀)
<b>Unguided noticing group</b>	N=10 (5♂, 5♀)	N=10 (5♂, 5♀)

*Table 1. Details of the participants in the study.*

### *3.2 Tasks and procedures*

In this study a modified version of Hanaoka's (2007) writing task was used to prompt participants to write a text individually in response to visual stimuli. The task was expected to provide learners with the opportunity (i) to notice their linguistic limitations when composing a text, (ii) to notice gaps between their IL and the TL by comparing their written output with two models and (iii) to revise and rewrite their original draft based on what they noticed in the comparison stage.

The data collection was carried out over a period of three weeks for the lower intermediate students and over four days for the elementary students due to a school's time constraints. The task consisted of 3 stages: the first and second stages were carried out in the same day and the third one either a week or three days later. The following tables (Table 2 and Table 3) summarise the procedure followed with the elementary and lower intermediate students, respectively.

Day 1	Day 4
<ul style="list-style-type: none"> <li>• Stage 1: Composing (pre-test) (20 mins)</li> <li>• Likert scale (2 mins)</li> <li>• Stage 2: Comparison stage (20 mins)</li> <li>• Background questionnaire (10 mins)</li> </ul>	<ul style="list-style-type: none"> <li>• Stage 3: Rewriting (post-test) (15-20 mins)</li> <li>• Exit questionnaire (10 mins)</li> <li>• Placement test (40 mins)</li> </ul>

Table 2. Data collection time frame for the elementary students.

Week 1	Week 2	Week 3
<ul style="list-style-type: none"> <li>• Placement test (40 mins)</li> <li>• Background questionnaire (10 mins)</li> </ul>	<ul style="list-style-type: none"> <li>• Stage 1: Composing (pre-test) (20 mins)</li> <li>• Likert scale (2 mins)</li> <li>• Stage 2: Comparison (20 mins)</li> </ul>	<ul style="list-style-type: none"> <li>• Stage 3: Rewriting (post-test) (15-20 mins)</li> <li>• Exit questionnaire (10 mins)</li> </ul>

Table 3. Data collection time frame for the lower intermediate students.

In the Stage 1 writing task, or composing stage (pre-test), the participants received the visual prompt (Appendix B), Sheet 1 to write their text (Appendix C) and Sheet 2 (Appendix D) to note down whatever problem they had while writing on Sheet 1. The visual stimuli consisted of six colour pictures related to the topic of games and activities that children often practice in their spare time, thus, they were helpful to control the propositional content of the compositions. All images were taken from the website *Flickr* to control for copyright issues. The pictures did not include any verbal hint so that the participants could produce the TL from what they saw. Although the learners were given oral instructions in Basque at the beginning of the experiment, directions were also written at the top of each sheet in the same language, following Hanaoka's (2007) idea about using the learners' L1. Moreover, as in Hanaoka (2007), Sheet 2 provided specific examples of note-taking in Basque: 'I don't know how to say X in English', 'I wrote X, but I'm not sure if this is correct'. 'What is the past tense of X?' and 'I'm not sure whether the

picture is describing X or Y'. No word limits were imposed and the students could take notes in Sheet 2 in the language of their choice: Basque, Spanish or English. A time limit of 20 minutes was imposed because of class time restrictions. Yet, all the participants were able to finish the task within the time allotted.

At the end of Stage 1, the participants were told that they were going to receive two model texts describing the same pictures to review their drafts. However, before that, they had to indicate how eager they were to receive the models in a 5-point Likert scale (1 being 'not at all' and 5 'very much') (Appendix E). This task took them about 2 minutes. Then, Sheet 2 was collected. The students kept their original drafts (Sheet 1) and the visual prompt for Stage 2.

At this second stage, which immediately followed Stage 1, the students received the two models (Appendix F) and Sheet 3 (Appendix G for the UNG and Appendix H for the GNG). Two models were used for two main reasons: One, to avoid the participants copying from a single model and, two, to increase the chances of providing solutions to the problems that the students could have faced in Stage 1. One of the models was written by a native speaker of English and the other one by a highly-proficient speaker of the language. For ease of reference, they were named A and B, respectively. On Sheet 3, the participants were asked to write down all the differences they could notice between their original texts and the models. However, the procedure that the UNG and the GNG followed for that differed. As in Hanaoka (2007), the former merely listed the differences found following the specific examples provided at the top of Sheet 3 (Appendix G): 'I couldn't say X but A/B puts Y', 'I have expressed this idea as X, and A/B puts it as Y', 'I thought the past tense of the verb X was Y, but A/B writes Z'. These examples were slightly different from Hanaoka's (2007). In contrast, following Santos et al. (2010), the GNG had to fill in a noticing table while establishing the comparison between their texts and the models (Appendix H). This form of noticing was chosen because, as mentioned above, previous research has shown that

there may be a relationship between depth of processing and learning outcome (Qi & Lapkin, 2001): It seems that noticing with understanding has more positive effects than noticing without understanding. For this reason, Santos et al.'s (2010) noticing procedure was used with the GNG. This process consisted in filling in a table where the subjects (i) annotated their errors and its corresponding solutions, (ii) decided what type of error it was (lexical, grammatical or discourse) and (iii) stated whether they accepted or refused the solution and why. This group was also provided with the specific examples given above. Before the task, oral instructions were given in Basque and students were allowed to use Basque, Spanish or English to write on Sheet 3. Again a time limit of 20 minutes was imposed for time restrictions. At the end of the stage, Sheet 1, Sheet 3 and the picture prompt were collected.

Stage 3, or the rewriting task, took place either one week or three days later. Each participant was given a copy of his/her original text and was asked to rewrite it on Sheet 4 (Appendix I), trying to incorporate the differences noticed between the models and their drafts. The task lasted between 15 and 20 minutes approximately. The learners had not been informed of this task in advance, with the purpose of avoiding their memorization of the models; we wanted to know what the students really kept from the noticing stage.

Unlike in Hanaoka (2007), at the end of the three stages the students were asked to complete an exit questionnaire (Appendix J) where they had to give their insights on the tasks performed. The survey was based on Marsden (2006) and Marsden and Torgerson (2012), and its aim was to have some qualitative information about the learners' attitudes towards composing in response to a picture prompt and towards models as a form of CF. This was to assess the possibility of learners' performance being influenced by affective factors. However, the questionnaire was used as a secondary tool in the study.



Apart from the three stages in the task and the exit questionnaire, the participants also took a placement test to determine their proficiency level and completed a background questionnaire which gave us information about their linguistic profile. As can be seen in tables 2 and 3, the elementary students filled in the background questionnaire the first day of the actual experiment and took the placement test in the last one, whereas in the case of the lower intermediate students both tasks were performed a week before the first stage.

To summarise, in the current study a modified version of Hanaoka's (2007) writing task was used to prompt students to write a text individually in response to a visual stimuli. As in Hanaoka (2007), the participants were divided into two groups with different English proficiency levels so that the effect of that variable in noticing and subsequent uptake could be assessed. Moreover, noticing was measured by means of note-taking in this study too and two models were used as a form of CF. However, there are some differences between Hanaoka's (2007) research and the present study. First, while Hanaoka used an immediate and a delayed post-test in his experiment only a delayed post-test was used in ours, due to a school's time constraints. Second, the present study used two types of noticing in Stage 3: one of the groups, the UNG, followed Hanaoka's (2007) method of listing the differences noticed between the participants' original drafts and the models, while the second one, the GNG, used Santos et al.'s (2010) noticing with understanding method, where the subjects had to fill in a noticing table. Finally, a qualitative part was introduced in our study with the use of an exit questionnaire at the end of the experiment to assess the influence that affective factors could have on the students' performance.

### *3.3 Analysis*

The primary data source in the study consisted of (i) 60 original compositions written by the participants, (ii) 60 sets of notes made during Stage 1, (iii) 40 sets of notes made during the comparison stage, and (iv) 60 sets of revised compositions. Moreover, 40 answers to the exit

questionnaire, the information gathered in the background questionnaire regarding motivation and the answers to the Likert-scale measuring eagerness to receive the models were also used. All this data were analysed both quantitatively and qualitatively.

For research questions 1 and 2, noticing was operationalized in the form of note-taking (Coyle & Roca de Larios, 2014; Hanaoka, 2007; Hanaoka & Izumi, 2012; Izumi, 2002), that is written LREs or self-reports. As acknowledged by the original work, this technique has some important advantages and drawbacks. The weakest feature is probably the extent to which participants report all they notice, given that the act of writing is physically and cognitively demanding as well as time-consuming. Indeed, as Hanaoka and Izumi (2012) noted, this technique may specially fail to capture vaguely conceptualized ideas, fleeting thoughts, and hard to verbalize notions. However, note-taking has some advantages too, and it was used as the main data collection tool in this study for several reasons. First, for practical issues, the recording of think-aloud protocols or the use of similar alternative methods was not a viable option in noisy classrooms. Second, being an introspective method, note-taking avoids subjects retrieving different information in the time of the verbal report and the actual experimental task (Bowles & Leow, 2005; Sachs & Polio, 2007). That is to say, note-taking can indicate in a more precise way the learners' focus of attention while on task because of the lack of influence of factors such as memory loss. Moreover, as pointed out by Hanaoka (2007), descriptive notes might provide evidence of the learners' type of awareness; for instance, whether a feature is new or already familiar to the participant.

In order to analyse the aspects of language the participants noticed, students' written LREs were classified into three categories: Problematic features noticed (PFNs) while composing in Stage 1, Features noticed (FN) from the comparison with the models in Stage 2 and Features noticed and incorporated (FNI) in Stage 3. Moreover, PFNs and FNs were further classified into four

categories: Lexis, grammar, content and discourse, and other. As in the original work, for the identification of lexical and grammar features William's (2001) classification of LREs was used. The examples to be presented below illustrate how the codification of PFNs and FNs was done. All the notes that follow are presented as they were written by the participants. Note that the students could use the language of their choice to take notes.

### *Lexis*

The lexical category includes items in which the students refer to their inability to find or spell a lexical item in English in Stage 1 (6-7) or features where the students acknowledge a previously unknown word in Stage 2 (8).

- (6) Ez dakit nola esaten den “columpio” ingelesez (Stage 1 PFN).

*'I don't know how to write “columpio” in English.'*

- (7) “Blonde” idatzi dut ile horia jartzeko baina ez nago ziur zuzena ote den (Stage 1 PFN).

*'I have written “blonde” to refer to blond hair, but I'm not sure if this is correct.'*

- (8) Nik “Hula-hop” idatzi dut eta “hula-hoops” idazten da (Stage 2 FN).

*'I have written “hula-hop”, but it should be “hula-hoops”.'*

### *Grammar*

In contrast, the grammar category includes features that focus on tenses (9), choice of preposition (10) and word order, among others.

- (9) Ez dakit “have changed” aditz denbora ondo erabili dudan (Stage 1 PFN).

*'I don't know whether I have used the correct tense in “have changed”.'*

- (10) “On the picture” idatzi dut, baina ereduian “in the picture” agertzen da (Stage 2 FN).

*'I wrote “on the picture”, but in the models they use “in the picture”.'*

### *Content and discourse*

Under this category fall all the features related to the way ideas are expressed and to the content that the students did not include in their compositions and noticed in the models.

- (11) Pilateseko pilota bat dagoela idatzi dut, baina ez dakit zuzena ote den (Stage 1 PFN).

*I have written that there is an exercise ball, but I don't know if this is correct'.*

- (12) Nik ez dut fondoa deskribatzen eta, A/B erduek bai (Stage 2 FN).

*I haven't described the background, whereas A/B models do'.*

- (13) Nik “in the first picture”, “in the second picture”... idatzi dut, baina Ak it describes it without introduction (Stage 2 FN).

*I have written in the first picture, in the second picture, while A describes the pictures without an introduction'.*

The participant in example (11) faced problems when describing a picture, because he was unable to distinguish whether the ball in the picture was an exercise ball or not. In contrast, (12) is an illustration of the noticing of some pieces of information (i.e. the background of the pictures) that were included in the models, but not in the learner’s composition. Finally (13) was classified under content and expression, because the participant compares the way in which the ideas had been expressed and organised in his/her original draft and model A.

### *Other*

The analysis of the data revealed that some features were difficult to classify in any of the categories previously mentioned. For instance, in (14) the participant gives his/her evaluation of one of the models, without referring to any specific aspect. Similarly, in (15) and (16) the learners compare and evaluate their drafts and the models in general terms, without giving specific details.

(14) B testua ondo dagoela iruditzen zait (Stage 2 FN).

*'I think that model B is well written.'*

(15) B testua nirearen antzekoagoa da (Stage 2 FN).

*'My text is more similar to B.'*

(16) Emandako testuak nireak baino egokiagoak dira (Stage 2 FN).

*'The models provided are more appropriate than my composition.'*

For research question 3, following Hanaoka (2007), PFNs were classified as either *solvable* or *unsolvable* in order to obtain a more accurate picture of the learners' incorporation of problematic features. A solvable PFN is a feature for which at least one of the models provides a solution. For example, many participants noted that they needed a lexical item for the Spanish word *columpio* (see example (6) above). Both model A and B provide a solution for it: *swing*. Therefore, PFNs involving *columpio* were classified as solvable. On the other hand, an unsolvable PFN refers to a feature for which neither model A nor B offers a solution. For instance, one of the participants noted that she did not know how to spell *bosgarren* (*fifth*) in English. However, none of the models used this word. Consequently, this PFN was classified as unsolvable. As mentioned above, classifying PFNs as solvable or unsolvable was important to analyse more accurately the degree of PFN incorporation by learners, since not excluding the unsolvable PFNs could have distorted the picture. It is noteworthy that in this study, following Hanaoka (2007), a feature attempted in the revision stage was counted as incorporated even if it contained some minor mistakes. For instance, if a learner noticed the lexical item *swing*, but misspelled it in the revision as *swiing*, it was still counted as an incorporated feature. In contrast, we did not count as incorporation those features which the students incorporated or corrected in their revisions, but did not explicitly discuss in their notes. We speculate that this might have happened because the learners were not able to explain the problematic features or because noting down all the problematic features was physically demanding for them.

Furthermore, to analyse the relationship between proficiency (elementary vs. lower intermediate) and depth of noticing (guided noticing vs. unguided noticing) and number of features, the total frequencies of PFNs, FNs and FNIs were tallied separately for each of the four categories, for each of the proficiency levels and for each of the treatment groups. The features noticed by the participants were also analysed qualitatively in order to trace them across the stages and in the final composition. For interrater reliability, ten per cent of the whole sample of PFNs and FNs was coded by a second researcher and agreement was found to be 100%.

Finally, for research question 4 two sources of data were considered to analyse the participants' attitude towards models as a form of feedback: First, a five-point Likert scale which was completed after Stage 1 and where learners were asked to report their eagerness to receive the models; and second, an exit questionnaire (Appendix J) which aimed at gathering information on learners' appreciation of the feedback tool in terms of usefulness and entertainment. As mentioned above, this tool was based on Marsden (2006) and Marsden and Torgerson (2012) and included Likert-scale questions, yes-no questions and open questions. The answers were both quantitatively and qualitatively analysed.

#### 4. Results

***Research question 1: What aspects of language do L2 learners of different proficiency levels notice while writing a composition on their own?***

As shown in Table 4, the learners noted a total of 187 PFNs, that is, an average of 3.11 PFNs per participant, a value slightly lower than in Hanaoka's (2007) work, where each participant noted an average of 3.5 PFNs. Moreover, those features were overwhelmingly lexical (89.84% of all the PFNs attested), which suggests that the learners' attention at this stage was primarily focused on finding the right words to express their intended ideas. Interestingly, similar results were found by Hanaoka (2007), although in this study the proportion of lexical PFNs reached 92.4%.

	All participants (N=60)				E Group (N=30)				LI Group (N=30)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Lexis</i>	168	89.84	2.8	1.69	77	97.47	2.56	1.36	91	84.26	3.03	1.97
<i>Grammar</i>	14	7.49	0.23	0.53	2	2.53	0.07	0.25	12	11.11	0.4	0.67
<i>Content</i>	5	2.67	0.08	0.28	0	0	0	0	5	4.63	0.17	0.38
<i>Other</i>	0	0	0.0	0	0	0	0	0	0	0	0	0
<b>Total</b>	187	100	3.11	1.47	79	100	2.63	1.3	108	100	3.6	1.63

*Table 4. Frequencies and proportions of PFNs (E=Elementary; LI= Lower intermediate).*

Looking at between group differences, a t-test revealed that there were significant differences between the two proficiency level's total number of PFNs ( $t(49)=2.07, p=0.04$ ). In other words, lower intermediate students noted significantly more PFNs than students in the elementary group. In addition, a MANOVA test was conducted to analyse the effect of the independent variable English proficiency on the groups' means for each of the categories (i.e lexis, grammar and content<sup>4</sup>). In fact, a significant difference was found on the multivariate test (Wilk Lambada's test statistics=0.79  $F(3,56)=4.85 p=0.00$ ). However, the more complete picture

<sup>4</sup> The category *Other* was excluded from the analysis because no PFNs of this type were noticed by the participants.

offered by the univariate tests revealed that although there were significant differences in the *grammar* ( $F(1,58)=6.42$   $p=0.01$ ) and *content* ( $F(1,58)=5.80$   $p=0.02$ ) aspects, this was not the case in the *lexis* ( $F(1,58)=1.14$   $p=0.29$ ) category. These findings and the ones reported in Table 4 above suggest that although both elementary and lower intermediate participants were mainly concerned with finding the appropriate vocabulary to express their ideas, the students with a higher proficiency also paid attention to other aspects, such as grammatical accuracy and expression of contents. This could be the reason for them to note significantly more problematic features.

Data were further analysed to check for significant differences between the various treatment groups (i.e. GNG, UNG and control group) in each proficiency level. Tables 5 and 6 show the results in the elementary and lower intermediate groups, respectively. First, one-way ANOVAs were conducted to analyse whether there were significant differences between the three treatment groups in the amount of PFNs noticed. This test was conducted in both proficiency groups and no significant differences were found in the elementary ( $F(2,27)=1.43$ ,  $p=0.26$ ) or in the lower intermediate levels ( $F(2,27)=0.26$ ,  $p=0.77$ ). Second, MANOVA tests<sup>5</sup> were used to check for significant differences in the type of PFNs noticed by the different treatment groups in each proficiency level. Multivariate and univariate tests<sup>6</sup> did not render significant differences in the elementary (Wilk Lambada's test statistics=0.73  $F(2,27)=1.73$   $p=0.2$ ) or in the lower intermediate levels (Wilk Lambada's test statistics=0.73  $F(6,50)=1,08$   $p=0.39$ ). This indicates that, at least initially, the learners assigned to each of the treatment groups (i.e. GNG, UNG and control) were similar in their focus of attention.

---

<sup>5</sup> The *Other* category was not included in the MANOVAs because no PFNs of this category were noticed by the participants in any of the groups. Similarly, the category Content was excluded from the MANOVA analyzing the elementary groups for the same reason.

<sup>6</sup> The univariate tests in the elementary group did not show any significant difference in the lexical ( $F(2,27)=1,73$   $p=0.26$ ) and grammar ( $F(2,27)=2.25$   $p=0.125$ ) categories. Similarly, no significant differences were found in the lower intermediate group regarding lexis ( $F(2,27)=0.43$ ,  $p=0,66$ ), grammar ( $F(2, 27)=1.60$ ,  $p=0.22$ ) and content ( $F(2,27)=1,72$ ,  $p=0.2$ ).



	E GNG (N=10)				E UNG (N=10)				E Control (N=10)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Lexis</i>	32	100	3.2	1.62	23	92	2.3	0.95	22	100	2.2	1.32
<i>Grammar</i>	0	0	0	0	2	8	0.2	0.42	0	0	0	0
<i>Content</i>	0	0	0	0	0	0	0	0	0	0	0	0
<i>Other</i>	0	0	0	0	0	0	0	0	0	0	0	0
<b>Total</b>	32	100	3.2	1.62	25	100	2.5	1.1	22	100	2.2	1.32

Table 5. Frequencies and proportions of PFNs in the elementary treatment groups (E GNG=Elementary guided noticing group; E UNG=Elementary unguided noticing group; E Control=Elementary control group).

	LI GNG (N=10)				LI UNG (N=10)				LI Control (N=10)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Lexis</i>	35	87.5	3.5	1.95	29	87.87	2.9	1.8	27	77.14	2.7	2.26
<i>Grammar</i>	5	12.5	0.5	0.71	1	3.03	0.1	0.32	6	17.14	0.6	0.84
<i>Content</i>	0	0	0	0	3	9.1	0.3	0.48	2	5.72	0.2	0.42
<i>Other</i>	0	0	0	0	0	0	0	0	0	0	0	0
<b>Total</b>	40	100	4	1.78	33	100	3.3	1.52	35	100	3.5	1.6

Table 6. Frequencies and proportions of PFNs in the lower intermediate treatment groups. (LI GNG= Lower intermediate guided noticing group; LI UNG=Lower intermediate unguided noticing group; LI Control= lower intermediate control group).

**Research question 2: What aspects of language do L2 learners of different proficiency levels notice as they compare their texts to two models? Are there any differences between the students who received some guidance on noticing and those who did not?**

Table 7 explains the frequencies, means and proportions of the FNs in Stage 2 Comparison. Note that the control groups did not complete this stage. Therefore, even if the raw number of FNs in Stage 2 (n= 158) is lower than in Stage 1 (n=187) the mean number of the features noted by each participant is higher in the former (mean= 3.95 FNs) than in the latter (mean= 3.11 PFNs). This finding is in line with Hanaoka (2007). Yet, the mean number in this study is slightly lower than

in the original, where the participants noted 4.4 FNs on average. Interestingly, consistent with Stage 1 noticing, the largest proportion of FNs was lexical (67.72%). However, one notable difference was that almost 24% of the FNs were related to the content and expression of the story. Some students noted that the models' interpretations of the images differed from their own and some others noticed different ways to express their ideas and structure their writings. For instance, a participant said that he only focused on the action that was taking place in the picture and disregarded other aspects which were described in the models, such as the background or the children's possible feelings.

	All participants (N=40)				E group (N=20)				LI group (N=20)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Lexis</i>	107	67.72	2.68	1.46	46	62.16	2.3	1.08	61	72.62	3.05	1.57
<i>Grammar</i>	4	2.53	0.1	0.33	2	2.70	0.1	0.31	2	2.38	0.1	0.31
<i>Content</i>	37	23.42	0.92	1.03	20	27.03	1	1.34	17	20.24	0.85	0.74
<i>Other</i>	10	6.33	0.25	0.50	6	8.11	0.3	0.66	4	4.76	0.2	0.41
<b>Total</b>	158	100	3.95	1.37	74	100	3.7	1.26	84	100	4.2	1.49

Table 7. Frequencies and proportions of FNs (E=Elementary; LI= Lower intermediate).

Concerning between group differences, a t-test was conducted to see the effect that the variable proficiency had on the amount of FNs noticed by the participants. The test showed that although lower intermediate students noticed more FNs, this difference was not statistically significant ( $t(37)=0.75$ ,  $p=0.46$ ). Furthermore, a MANOVA revealed that the students' proficiency did not have a significant effect on the nature of the aspects (i.e. lexis, grammar, content or other) they focused their attention on (Wilk Lambda's test statistics=0.89  $F(4,35)=1.02$   $p=0.41$ ). As shown in Table 7 above, both groups were primarily concerned with the *lexis* and *content* categories.

As with the first research question, data were analysed to check for significant differences between the treatment groups in each proficiency level. In particular, we were interested in

exploring the effect that guidance could have had on the depth and processing of learners' noticing. First, t-tests were conducted to determine whether the means of the FNs noticed by the GNG and UNG groups differed significantly. In the elementary group (see Table 8 for frequencies and proportions), the test showed that although the GNG noticed more FNs, this difference did not reach significance ( $t(17)=1.56, p=0.13$ ). Similarly, in the lower intermediate group (see Table 9 for frequencies and proportions), the t-test did not reveal significant differences ( $t(17)=2.03, p=0.06$ ), even if the participants in the GNG group noticed more FNs in this proficiency level too.

	E GNG (N=10)				E UNG (N=10)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Lexis</i>	26	59,1	2,6	1,17	20	66,67	2	0,94
<i>Grammar</i>	1	2,27	0,1	0,32	1	3,33	0,1	0,32
<i>Content</i>	16	36,36	1,6	1,58	4	13,33	0,4	0,7
<i>Other</i>	1	2,27	0,1	0,32	5	16,67	0,5	0,85
<b>Total</b>	44	100	4,4	1,45	30	100	3	1,03

Table 8. Frequencies and proportions of FNs in the elementary treatment groups. (E GNG=

Elementary guided noticing group; E UNG=Elementary unguided noticing group).

	LI GNG (N=10)				LI UNG (N=10)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>N</i>	%	<i>mean</i>	<i>s.d</i>
<i>Lexis</i>	41	80,39	4,1	1,2	20	60,6	2	1,15
<i>Grammar</i>	2	3,92	0,2	0,42	0	0	0	0
<i>Content</i>	7	13,73	0,7	0,95	10	30,3	1	0,47
<i>Other</i>	1	1,96	0,1	0,32	3	9,1	0,3	0,48
<b>Total</b>	51	100	5,1	1,83	33	100	3,3	1

Table 9. Frequencies and proportions of FNs in the lower intermediate treatment groups.

(LI GNG= Lower intermediate guided noticing group; LI UNG=Lower intermediate unguided noticing group).

Second, MANOVAs were carried out to check if the treatment groups significantly differed in their focus of attention. In the elementary group, the multivariate test showed no significant treatment effect (Wilk Lambda's test statistics=0.89  $F(4,15)=0.61$   $p=0.09$ ). Nevertheless, the univariate tests revealed that although the two groups performed similarly in the *lexis* ( $F(1,18)=1.59$   $p=0.22$ ), *grammar* ( $F(1,18)=0$   $p=1$ ) and *other* ( $F(1,18)=1.95$   $p=0.18$ ) categories, the GNG group focused significantly more on *content* ( $F(1,18)=4.84$   $p=0.04$ ). Regarding the lower intermediate students, on the other hand, the multivariate test determined that treatment had a significant effect on the learners' focus of noticing (Wilk Lambda's test statistics=0.21  $F(4,15)=14.23$   $p=0.00$ ). Interestingly, the univariate tests in this case showed that participants in the two treatment groups differed significantly in the *lexis* category ( $F(1,18)=15.94$   $p=0.00$ ), but not in the *content* ( $F(1,18)=2.25$   $p=0.15$ ), *grammar* ( $F(1,18)=0.80$   $p=0.35$ ) or *other* ( $F(1,18)=1.20$   $p=0.28$ ) categories.

In sum, these results suggest that, consistent with findings in Stage 1, elementary and lower intermediate participants primarily focused on lexical aspects while comparing their writings with models. However, it is noteworthy that in this second step learners also paid attention to the contents expressed by the models and to the alternative ways in which they did so. Moreover, regarding the differences between the two treatment groups, it seems that there is a tendency for guided note taking to promote more noticing, although this difference does not reach significance in the study. In addition, guided note taking appears to be particularly useful in the lower proficiency levels to help students direct their focus of attention to other aspects than lexis.

Table 10 below displays the relationship of the 158 FNs to the stage 1 PFNs. Sixty-five (65) FNs (about 41%) were related to stage 1 PFNs. These FNs mainly involved the noticing of words that the participants wanted to use in their original compositions but did not know in English or that actually used but misspelled. For instance, many students could not access the English words

*swing, sofa or headphones*, for which the two models offered a solution. Similarly, some participants who faced problems writing the word *hula-hoop* noted its correct spelling. On the other hand, the 93 FNs (about 59%) which were not related to the Stage 1 PFNs represented the noticing of new content and alternative ways of expression which had not been consciously searched for when composing. As an illustration, a participant noted that she ‘only described the main characters in the pictures, while the two models described everything’.

	All participants (N=40)				E Group (N=20)				LI Group (N=20)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Rel</i>	65	41,14	1,6	0,99	30	40.54	1,5	1,15	35	41,67	1,75	0,79
<i>Non-rel</i>	93	58,86	2,3	1,76	44	59.46	2,2	1,44	49	58,33	2,45	1,9
<b>Total</b>	158	100	3,95	1,45	74	100	3,7	1,44	84	100	4,2	1,48

Table 10. Stage 2 FN in relation to Stage 1 PFNs (E=Elementary; LI= Lower intermediate; Rel= Related to Stage 1 PFNs, Non-rel=Non-related to Stage 1 PFNs).

Moreover, as Hanaoka (2007) pointed out, some of the non-related FNs involved delayed noticing of problems with their original output. For example, a student who originally thought the correct preposition to use was ‘on the picture’ noted that she should have used ‘in the picture’. She had not focused on this issue as problematic in Stage 1. Finally, the notes of some of the students gave us the impression that some of the non-related FNs were in fact related to certain problematic features which had been noticed in Stage 1 but not explicitly discussed by the students, perhaps because they were not able to explain them or because they found writing a physically demanding task. For instance, one of them wrote: ‘I did not know how to say *columpio* in English, and text B says *swing*’. However, he did not note down this as a PFN in Stage 1 composing. Therefore, this FN was classified as non-related to Stage 1 PFN, but it seems that it could have been so.

*Research question 3: What are the effects of Stage 1 and Stage 2 noticing on subsequent revisions?*

In order to answer this research question accurately, the 187 PFNs noticed by the students in Stage 1 (See table 4) were first classified as either solvable or unsolvable (see analysis section), following Hanaoka (2007). Out of those 187 PFNs, 140 (74.87%) were solvable and 47 (25.13%) unsolvable, a value slightly lower than in Hanaoka (2007), where 29% of the total PFNs were unsolvable. In line with previous results, an overwhelming number of the solvable PFNs in this study were lexical (n=128), while those in the grammar (n=7) and content categories (n=5) together did not make up 10% of the total. Moreover, during the Stage 1 writing task, participants had already solved by themselves 23 of the 140 solvable PFNs (16.42%) by using exactly the same features as the models: 9 (39.13%) were solved by elementary students and 14 (60.87%) by lower-intermediate participants. Following Hanaoka (2007), these already-solved PFNs were excluded from the analysis. The remaining 117 solvable PFNs were then analysed to examine the frequencies at which students noticed solutions to them in Stage 2, and the frequencies at which they incorporated those solutions in Stage 3. Taking into account that the control groups did not participate in Stage 2 comparison, the results of the treatment groups and the control groups will be reported separately, so as to obtain a more accurate picture of the effect of noticing in subsequent revisions.

As Table 11 shows, the students in the treatment groups noted 63.33% of the solutions available from the models and incorporated 71.93% of them in the revision task. These values were lower than in Hanaoka (2007), where the participants noted 65.4% of the solutions and incorporated 92% of them. However, it should be noted that the values in Hanaoka (2007) correspond to an immediate revision, while those reported in this study correspond to a delayed one. Moreover, in this study 12 participants (7 elementary and 5 lower intermediate) incorporated 15 additional

lexical items (10 attributed to elementary students) in their revisions without noting them in the comparison Stage, pointing at incomplete note-taking by the participants.

Students in the control groups (see Table 11), on the other hand, incorporated 4 out of the 27 solvable PFNs (14.81%). Interestingly, all of them were lexical and incorporated by lower intermediate students. Indeed, three out of the four were attributed to the same participant, which might indicate that this student looked for alternative forms of feedback outside the experiment. A two-sample z-test was conducted to check for significant differences between the treatment and control groups on the incorporation of solvable PFNs. The test showed that, at the  $\alpha= 0.05$  significance level, the participants in the treatment groups incorporated significantly more PFNs than those in the control groups ( $z=2.88, p=0.00$ ), thus, confirming the effectiveness of models as a form of feedback and discarding the influence of task repetition.

	Treatment groups (N=40)				Control groups (N=20)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>N</i>	%	<i>mean</i>	<i>s.d</i>
<i>Solvable PFNs</i>	90	100	2.25	1.19	27	100	1.35	1.14
<i>Stage 2 Noticing</i>	57	63.33	1.43	0.9	--	--	--	--
<i>Stage 3 Inc.</i>	41	45.6	1.03	0.8	4	14.81	0.2	0.7

*Table 11. Solutions to PFNs noticed and incorporated across groups.*

In addition, data from the treatment groups was further analysed to determine whether proficiency had a significant effect on (i) stage 2 noticing and (ii) Stage 3 incorporation of solvable PFNs (see Table 12 for frequencies and proportions in each group). Two two-sample z-tests between proportions were performed, which showed that although the students in the lower intermediate group noticed and incorporated more solvable PFNs, this difference was not significant either in stage 2 noticing ( $z=-1.88, p=0.06$ ) or in stage 3 incorporation ( $z=-1.10, p=0.27$ ) at the  $\alpha= 0.05$  significance level.

	E group (N=20)				LI group (N=20)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Solvable PFNs</i>	43	100	2.15	1.14	47	100	2.35	1.27
<i>Stage 2 Noticing</i>	23	53.49	1.15	0.88	34	72.34	1.7	0.86
<i>Stage 3 Inc.</i>	17	39.53	0.85	0.81	24	51.06	1.2	0.77

Table 12. Solutions to PFNs noticed and incorporated across proficiency groups

(E=Elementary; LI=Lower intermediate).

Moreover, the performance of the treatment groups in each proficiency level were analysed to determine the effect quality of noticing could have had on the noticing and incorporation of solvable PFNs. Table 13 below shows the results of the GNG and UNG groups in the elementary level. The z-test conducted showed that the UNG noticed ( $z=-2.44$ ,  $p=0.02$ ) and incorporated ( $z=-3.45$ ,  $p=0.01$ ) significantly more solutions to PFNs than the GNG. However, note that the difference is only significant if we compare the proportion of features incorporated from the total PFNs noticed in Stage 1. If, alternatively, we look at the proportion of features which were noted in Stage 2 and incorporated in Stage 3 the difference is not significant ( $z=-0.91$ ,  $p=0.37$ ). This indicates that the GNG found fewer solutions to their PFNs than the UNG, but the incorporation of the solutions noted was similar in both groups.

	E GNG(N=10)				E UNG (N=10)			
	<i>n</i>	%	<i>mean</i>	<i>s.d</i>	<i>n</i>	%	<i>mean</i>	<i>s.d</i>
<i>Solvable PFNs</i>	26	100	2.6	1.17	17	100	1.7	0.95
<i>Stage 2 Noticing</i>	10	38.46	1	0.94	13	76.47	1.3	0.82
<i>Stage 3 Inc.</i>	6	23.07	0.6	0.7	11	78,6	1.1	0.88

Table 13. Solutions to PFNs noticed and incorporated in the elementary treatment groups.

(E GNG=Elementary guided noticing group; E UNG=Elementary unguided noticing group).



Table 14, in contrast, shows the noticing and incorporation of solvable PFNs in the lower intermediate treatment groups. Here the z-test indicated that although the GNG noticed and incorporated more features, the differences were not significant either if we compare the proportion of PFNs incorporated (for noticing  $z=0.42$ ,  $p=0.67$  and for incorporation  $z=1.03$ ,  $p=0.30$ ) or the proportion of PFNs noticed and incorporated ( $z=0.98$ ,  $p=0.32$ ).

	LI GNG(N=10)				LI UNG (N=10)			
	<i>n</i>	<i>%</i>	<i>mean</i>	<i>s.d</i>	<i>n</i>	<i>%</i>	<i>mean</i>	<i>s.d</i>
<i>Solvable PFNs</i>	24	100	2.4	1.07	23	100	2.3	1.49
<i>Stage 2 Noticing</i>	18	75	1.8	0.63	16	69.57	1.6	1.07
<i>Stage 3 Inc.</i>	14	58.33	1.4	0.7	10	43.47	1	0.97

*Table 14. Solutions to PFNs noticed and incorporated within the lower intermediate treatment groups. (LI GNG= Lower intermediate guided noticing group; LI UNG=Lower intermediate unguided noticing group).*

***Research question 4: What is the participants' attitude towards models as a form of feedback?***

***How does this affect their noticing and performance in subsequent revisions?***

Table 15 below shows the mean scores of the participants' answers to three five-point Likert-scales measuring (i) the learners' eagerness to receive the models, (ii) their level of enjoyment while using the models and (iii) their appreciation of the usefulness of models as a form of feedback. As the figures show, the participants in this study were not very motivated (mean: 2.42, s.d: 1.03) to read the models and to compare them with their own writings. This contrasts with the findings in Hanaoka (2007), where the mean score in the same 1 to 5 scale was 4.3 (s.d: 0.86).

However, a one-way ANOVA revealed that there were significant between group differences in the students' willingness to receive the models ( $F=5,10$ ;  $d.f.=3$ ;  $p=0.00$ ). More specifically, the t-

tests conducted post-hoc showed that the elementary UNG was significantly more eager than the other three groups, namely the elementary GNG ( $t(17)=4,57$ ;  $p=0.00$ ), the lower intermediate UNG ( $t(18)=2,52$ ;  $p=0.02$ ) and the lower intermediate GNG ( $t(17)=2,36$ ;  $p=0.03$ ) to read the models and to use them as a form of feedback. No significant differences were found among the rest of the groups<sup>7</sup>.

	All participants (N=40)		E group (N=20)				LI group (N=20)			
	<i>mean</i>	<i>s.d</i>	GNG (N=10)		UNG (N=10)		GNG (N=10)		UNG (N=10)	
	<i>mean</i>	<i>s.d</i>	<i>mean</i>	<i>s.d</i>	<i>mean</i>	<i>s.d</i>	<i>Mean</i>	<i>s.d</i>	<i>mean</i>	<i>s.d</i>
<i>Eagerness</i>	2.45	1.03	1.8	0.63	3.3	0.82	2.3	1.06	2.3	0.95
<i>Enjoyment</i>	2.7	0.97	2.5	1.17	2.9	0.88	2.7	0.82	2.7	1.06
<i>Usefulness</i>	3.46	1.07	3.2	1.13	3.5	1.27	3.9	0.57	3.4	1.26

*Table 15. Level of eagerness, enjoyment and usefulness of models in the elementary and lower intermediate treatment groups (E= Elementary group; LI= Lower intermediate group; GNG= Guided noticing group; UNG=Unguided noticing group)*

As for the usefulness and enjoyment of modelling, the students' answers to the exit questionnaire revealed that although they found the models quite useful to improve their English (mean: 3.46, s.d: 1.07), they did not really enjoy studying them (mean: 2.7, s.d:0.97). Moreover, the two one-way ANOVAs which were carried out showed that there were no significant between-group differences in terms of usefulness ( $F=0.72$ ; d.f.=3;  $p=0.55$ ) and enjoyment ( $F=2.7$ ; d.f.=3;  $p=0.8$ ). A qualitative analysis of the responses was valuable to obtain a more detailed picture of these beliefs and attitudes. Regarding the former, many participants reported that model texts offered them the opportunity to learn new vocabulary, forms of expression and ideas on content, as the following answers illustrate:

<sup>7</sup> No significant differences were found between (i) the guided elementary group and the unguided lower intermediate group ( $t(16)=-1,39$ ;  $p=0.18$ ), (ii) the guided elementary group and the guided lower intermediate group ( $t(15)=-1.28$ ;  $p=0.22$ ) and (iii) the guided and unguided lower intermediate groups ( $t(18)=0$ ;  $p=1$ ).

- (17) Erabilgarria iruditu zait espresio eta hitz berriak ikasteko.

*I found it useful to learn new forms of expression and new words.*

- (18) Erabilgarriak dira pentsatzen duguna beste modu batean adierazteko.

*They are useful to express our thoughts in another way.*

- (19) Oso erabilgarriak dira ideia berriak ematen dizkizulako.

*They are very useful because they offer you new ideas.*

However, few other participants stated that the English level of the models was too high and that this made it difficult for them to correct their own writings and learn new aspects of the L2 (20-22). Furthermore, another participant pointed out that receiving individual advice is, in his opinion, more useful than having to compare one's own composition with other texts (22), possibly hinting at EC as a more effective feedback tool.

- (20) Ereduak askoz hobeto daudenez, frustratuta sentitu naiz eta horrela zaila da gauzak ikastea.

*As the models are better written, I've felt frustrated and this way is difficult to learn something.*

- (21) Zaila egin zait ereduak ulertu eta gaizki neukana identifikatzea

*I found it difficult to understand the models and to identify my errors.*

- (22) Zure testua aholkuak emanez hobetu behar da, ez beste eredu batzuekin konparatuz.

*You have to improve your text by receiving advice, not by comparing it with other models.*

As for the question on whether they enjoyed studying the models, most participants' reported that this was a boring and long activity (23). Indeed, even those who valued it positively, referred to the usefulness of models rather than to the interest the comparison enkindled in them (24-25). Therefore, we may conclude that the activity was not intrinsically motivating for the students.

(23) Ez zait gehiegi gustatu ariketa luzea eta asperrgarria zelako.

*I didn't like the activity because it was long and boring.*

(24) Egindako akatsak ikusteko aukera izan dudalako gustatu zait ariketa.

*I liked the activity because it gave me the opportunity to identify my errors.*

(25) Ingelesa hobetzeko lagungarria da.

*It is useful to improve your English level.*

In line with the attitudes reported so far, only 37.5% of the participants answered affirmatively when asked whether they would like their EFL teacher to use this feedback method in the future (see table 16):

	All participants (N=40)		E group (N=20)				LI group (N=20)			
			GNG (N=10)		UNG (N=10)		GNG (N=10)		UNG (N=10)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>N</i>	%	<i>n</i>	%
<i>Yes</i>	15	37.5	3	30	3	30	3	30	6	60
<i>No</i>	25	62.5	7	70	7	70	7	70	4	40

*Table 16. Participants' eagerness to use models in future EFL lessons. (E= Elementary group; LI= Lower intermediate group; GNG= Guided noticing group; UNG=Unguided noticing group)*

Among the responses gathered, it is interesting that although some mentioned the value or level of entertainment of models as the cause behind wanting to repeat the activity or not, many participants made reference to their lack of interest in writing (26-27), to their lack of motivation to learn English (28-29), to their self-efficacy beliefs (30-32) and to the role of writing in L2 development (33-34).

(26) Idaztea asperrgarria delako.

*Because writing is boring.*

(27) Ez zaidalako ingelesez idaztea gustatzen.

*Because I don't like writing in English.*

(28) Ingelesa ez delako dibertigarria.

*Because English is not fun.*

(29) Ingelesa ez zait gustatzen.

*I don't like English.*

(30) Ez dut irakasleak erabiltzerik nahi, nik zein gaizki idazten dudan erakusten didalako.

*I don't want the teacher to use it (models as a form of feedback), because it reveals how badly I write.*

(31) Oso gaizki moldatzen naizelako ingelesez.

*Because I am bad at English.*

(32) Ingeles maila bazu samarra dudalako.

*Because my English level is poor.*

(33) Idazten ikastea ingelesa ikasteko erabilgarria delako.

*Because learning how to write is useful to learn English.*

(34) Idazten ikasten da, baina hitz egitea garrantzitsuagoa da.

*You can learn how to write, but speaking is more important.*

All this data indicates that learners' attitudes towards a given form of feedback are influenced by many factors other than the nature of the technique itself. Among them, the study analyzed the use of models as a feedback tool in the participants' EFL lessons, since several pieces of research have suggested that learners' beliefs derive from the techniques previously used by their teachers (Kormos, 2012). Table 17 below shows that for most of the learners in the four groups the use of models was totally new and, as (35) illustrates, this may also have had an influence in the learners' attitude towards modelling.

	All participants (N=40)		E group (N=20)				LI group (N=20)			
	<i>n</i>	%	GNG (N=10)		UNG (N=10)		GNG (N=10)		UNG (N=10)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
<i>Used</i>	1	2.5	0	0	0	0	1	10	0	0
<i>Sometimes</i>	10	25	4	40	4	40	2	20	0	0
<i>Not used</i>	29	72.5	6	60	6	60	7	70	10	100

Table 17. Familiarity with models in the four treatment groups. (*E*= Elementary group; *LI*= Lower intermediate group; *GNG*= Guided noticing group; *UNG*=Unguided noticing group)

(35) Ez zait asko gustatu ez zaidalako arrunta egin.

*I didn't like it much, because it was not common for me.*

Finally, two multiple regression tests were conducted in order to analyse the effect the variables hinting at the learners' attitude towards models (i.e. eagerness to receive the models, the usefulness and enjoyment of using this technique and wanting the teacher to incorporate

modelling in her lessons or not) had on the number of FNs noticed in Stage 2 and FNIs incorporated in Stage 3. As for the former, the test showed that there was not any significant correlation (F-statistics=0.504; d.f.=4;  $p=0.72$ ) between the variables mentioned and the noticing of features in the models. In contrast, the second multiple regression test showed that the model could explain 25% of the variation in the results and that the correlation between the variables and the number of noticed features incorporated in the revision was significant (F-statistics=2.95; d.f.=4;  $p=0.03$ ). Taking a closer look to each of the covariates, we observe that the level of enjoyment and the attitude learners showed towards incorporating modelling in future teaching practices has a significant effect on the number of features incorporated (for enjoyment:  $t=2.12$ ;  $p=0.04$  and for incorporating this method:  $t=2.20$ ;  $p=0.03$ ). Thus, those participants who rated models higher in terms of entertainment and those who said they would like their teachers to use this feedback method were found to have incorporated significantly more features in their revisions. These results seem to support the idea that individual variables affect the effectiveness of models as a form of feedback and that showing positive attitudes may facilitate language development.

## 5. Discussion

This study was set up to investigate the LLP of writing by analysing the participants' noticing processes and subsequent uptake of feedback from models in a written picture description task. In what follows the main sets of findings that emerged from the data will be discussed in the light of previous research.

### 1. Learner noticing in the composing and comparison stages

The data suggested that (i) learners are able and, in fact, do notice 'holes' and 'gaps' (Swain, 1995) in their IL when producing written output and processing feedback (Qi & Lapkin, 2001; Williams, 2001) and that (ii) most of this noticing is lexically driven. In the composing stage, the problems were often formulated in terms of lack of knowledge (e.g. 'ez dakit nola esaten den kolumpio ingelesez – *I don't know how to say swing in English*) or of potential solutions to the problems encountered (e.g. hoola hops idatzi dut baina ez nago ziur zuzena ote den – *I've written hoola hops, but I'm not sure whether this is correct*), while in the comparison stage they were expressed in terms of realisation (e.g. ez nekien kolumpio esaten ingelesez eta An swing jartzen du – *I didn't know how to say swing in English and A says swing*).

The learners' concern with lexis fully confirms the data reported by Hanaoka (2007) and other studies investigating attentional processes in free writing tasks (Cánovas Guirao et al., 2015; Coyle & Roca de Larios, 2014; Hanaoka, 2006a; Hanaoka, 2006b; Hanaoka & Izumi, 2012; Martínez Esteban & Roca de Larios, 2010; Yang & Zhang, 2010). As Williams (2001: 338) claimed 'learners focus, above all things, on words'.

Such a lexical bias could be an artefact of various factors. First, it may be the consequence of learners operating with limited processing capacities. According to VanPatten's (1996) model of input processing, learners are not capable of paying attention to all the information in the input,



and hence, focus primarily on those parts that are immediately relevant to the message. Thus, they search for content words first and focus on grammatical forms only if their resources are not fully depleted. Alternatively, the focus on lexis could be the result of the perceived need for vocabulary in the writing stage. The ‘priming’ effect of output (Izumi, 2003) may have pushed the participants to address the holes they noticed while composing by looking for relevant word forms in the two models. Nevertheless, it should be noted that during the comparison stage learners also noticed new ‘gaps’ in their original output and that, indeed, increased attention was paid to the content of the texts and their expression. These results are in line with Hanaoka (2007) and Martínez Esteban and Roca de Larios (2010) and indicate a useful role for TL models to promote noticing and to diversify participants’ concerns, especially by drawing it towards the content of the texts and its expression.

## *2. From noticing to incorporation in subsequent revisions*

As for the effect of noticing in Stage 1 and 2 on subsequent revisions in Stage 3, the results indicated that the participants in the treatment groups noticed about two-thirds (63.33%) of the solutions available, which points at a strong relationship between Stage 1 and Stage 2 noticing. These findings are in line with Coyle and Roca de Larios (2014), Hanaoka (2007) and Qi and Lapkin (2001), who observed that what learners noticed while composing, greatly influenced what they paid attention to in the comparison task. From a psycholinguistic perspective this can be explained with the Zeigarnik effect (Zeigarnik, 1999), which claims that incomplete tasks create anxiety and dissonance in the human mind and thus tend to be retained in memory better. The noticing of holes through output and the failure to reach a solution may also trigger this psychological effect and push the learners to look for a relevant form that might help them solve the problem (Hanaoka, 2007; Qi & Lapkin, 2001).

Similarly, this effect may also predict the incorporation and retention of solutions in subsequent revisions. In this study after noticing nearly two-thirds of the solutions available, the participants who had access to the models incorporated almost 72% of the relevant forms even though the models and the noticing sheet were taken away. Similar results were also found by Hanaoka (2007). This seems to indicate that apart from engaging learners in the noticing of holes and in the search for solutions, output serves to encourage learners to incorporate the solutions identified.

Another important finding was that while the learners in the control groups provided a solution to 14% of the problems identified in Stage 1, those in the treatment groups incorporated a relevant form in 45% of the cases. This contrasts with the results found by Cánovas Guirao et al. (2015) who claimed that task repetition may have been responsible for the improved performance in their research. Alternatively, our study lends support to the effective role models play as feedback tools. Moreover, it should be noted that the learners who were provided with TL models incorporated in addition other features they did not identify as problematic while producing output. This may indicate that models, as opposed to other EC forms which must remain more faithful to the original, are useful to provide learners' with alternative lexis, forms of expression and content. However, the possibility of incomplete note-taking also needs to be acknowledged. Taking into account that writing may be regarded as a physically demanding task, it might be the case that learners did not explicitly discussed all the problematic features noticed in Stage 1 and 2.

### *3. Proficiency effects in noticing and incorporation*

The data in this study showed that the more proficient group noticed significantly more PFNs than the less proficient group in the composing stage and that their focus of attention was more

varied. Lower intermediate students noticed significantly more grammatical and content features than the elementary students, who almost uniquely focused on lexis. Similarly, in the Stage 2 comparison task, higher-proficiency learners noticed more features than lower-proficiency learners from the two native-speaker models. However, the difference was not significant in this case and there were no differences in the nature of the FNs. Moreover, although the more proficient group noticed and incorporated a higher percentage of solutions to their PFNs, the difference was not significant. Consequently, these results add only partial support to Hanaoka (2007), Qi and Lapkin (2001) and Swain and Lapkin (1995) who, among others, showed that more advanced learners produced more LREs and understood their nature better. Nevertheless, it seems to suggest that spontaneous FonF tasks using models may be more beneficial for students with a higher L2 proficiency.

#### *4. The effect of guidance in noticing and incorporation*

The results of the study showed that the GNG groups both at the elementary and lower intermediate levels noticed more features in the comparison stage than the UNG groups. Although the difference did not reach significance, the data seems to indicate there is a tendency for guidance to promote noticing. However, further research should be carried out in this respect.

The results also showed that guided note taking was particularly useful in the lower proficiency levels to help learners direct their attention to aspects other than lexis, since the participants in the GNG group noticed significantly more content features than those in the UNG group, who mostly focused on lexis. This finding seems to support Yang and Zhang's (2010) claim that learners with a lower level of L2 proficiency have to be guided and trained in their noticing and processing of feedback through awareness-raising activities.

Nevertheless, the effect of guidance in the noticing and incorporation of solutions to Stage 1 PFNs is less straightforward. On the one hand, data from the lower intermediate group showed that, even if the difference did not reach significance, the GNG group noticed and incorporated more features than the UNG group. On the other hand, the results in the elementary group indicated that at this level, learners who did not receive any guidance noticed and incorporated significantly more solutions to Stage 1 PFNs than those who received guidance - note that this group was significantly more motivated than the rest to receive the models. As a consequence, it may seem that the study lends partial support to, or even contradicts, Qi and Lapkin (2001), Sachs and Polio (2007) and other scholars who have claimed that the quality of noticing of the learners has a direct impact on their learning outcome and that noticing with comprehension (i.e. substantive noticing) is associated with greater intake. However, the findings should be interpreted with caution, since they might have been influenced by other factors.

Firstly, as Hanaoka (2006a) and Sachs and Polio (2007) acknowledged, it is possible that even noticing with understanding needs some rehearsal in memory or positive assessment from the part of the learner in order to be incorporated in the long-term memory. This would highlight the importance of designing extended practices and activities to reinforce the linguistic forms noticed, and thereby, facilitate their acquisition.

Secondly, it may be the case that Santos et al.'s (2010) method to encourage a deeper noticing and processing of feedback was not effective enough to represent awareness at the level of understanding or what Qi and Lapkin (2001) named as substantive noticing. Learners in the GNG group might have completed the task by simply identifying all the differences between their output and the models provided without engaging in a deeper level of understanding. That is to say, it is possible that even if the form in which the participants in the GNG and UNG groups noted down the features noticed differed, the quality of noticing at both groups was

similar. However, it should be pointed out that Santos et al.'s (2010) noticing technique offers valuable information regarding the possible reasons for not incorporating a noticed feature. As mentioned above, noticing may have to be accompanied by positive assessment from the learner (Hanaoka, 2006a) and following Hanaoka's (2007) note taking method it is difficult to know if a participant accepts the alternative in the model. In contrast, Santos et al.'s (2010) chart asks learners to assess the features noticed overtly (See appendix H), thus, providing data which may clarify whether the non-incorporation of a noticed feature is a deliberate decision or not.

Thirdly, the analysis of the students' attitudes towards modelling showed that the elementary UNG was significantly more motivated than the rest of the groups to receive the models. This might have resulted in a deeper noticing and processing of the feedback, which could explain why this group noticed and incorporated significantly more features than the elementary GNG, even if the treatment it received was not specifically designed to enhance a substantive noticing. This finding seems to add support to the important role individual variables play in the extent and depth of processing of feedback, as well as in its LLP.

Finally, the way in which the incorporation of features was analysed might have also influenced the results showing the effect of guidance. Following Hanaoka (2007), the present study considered only the incorporation of the solutions to the PFNs noticed in Stage 1 and disregarded the incorporation of other alternatives provided by the models and, indeed, noticed by the participants in Stage 2. This might have mostly influenced the lower proficiency learners in the GNG group who noticed significantly more content features than the UNG group at the comparison stage. Interestingly, most of these features were not related to Stage 1 PFNs and, therefore, their incorporation was not tallied. It is possible that the broadening in the focus of attention in this group resulted in the incorporation of fewer solutions to stage 1 PFNs, but on a greater overall development of their IL which was not reflected in the analysis carried out.

### *5. Learners' attitude towards models*

The results of the study showed that participants had a rather negative attitude towards modelling and writing in general. The Likert-scale measuring their willingness to compare the writings with the models revealed their reluctance to receive this type of feedback, which contrasts with the high motivation levels showed by the participants in Hanaoka (2007). This difference could be due to the age of the learners in the two studies. While Hanaoka (2007) analysed the performance of 37 undergraduate students, this study looked at adolescents in secondary and high school.

The answers to the exit questionnaires also showed that although the students found the models useful to learn new vocabulary and new ways of expression, most of them did not enjoy the activity because they considered it long and boring. In this sense, operationalising noticing in the form of note-taking might have added an extra physical demand which negatively affected learners' attitude towards modelling. Because of all these reasons more than half of the participants (62.5%) reported that they would not like their teachers to incorporate this type of feedback in their teaching practices.

However, the qualitative analysis of the responses revealed that this negative attitude was influenced by many other factors which were not intrinsic to the technique; namely the students' lack of interest in writing, their lack of motivation to learn English, their low self-efficacy beliefs, the secondary role they think writing plays in L2 development and the feedback conventions shaped by previous language instructions - note that the large majority of the students reported not having used this technique before-. These findings are in line with Kormos (2012) and Storch and Wigglesworth (2010), which maintained that learners' expectations and beliefs are often shaped

by teacher practice and motivational factors such as personal goals and intrinsic interest in language learning.

Furthermore, the quantitative analysis carried out indicated that those students who showed a more positive attitude towards modelling incorporated more features in subsequent revisions. These results add support to Hyland (2011), Kormos (2012) and Storch and Wigglesworth (2010) which claimed that writing, noticing gaps and engaging in problem-solving tasks are laborious activities which require extensive engagement and high levels of motivation for learning to occur. Therefore, we may conclude that affective factors influence not only the attention paid to feedback by learners or their involvement in text revision activities but also the likelihood of retaining these corrections. Thus, it seems that while students who hold negative beliefs are likely to complete the writing task without engaging in further cognitive processes, learners with a positive attitude are prone to reach higher levels of uptake and to further progress in their L2 acquisition process. Nonetheless, these results should be interpreted with caution because the analysis did not show any significant effect of the variables mentioned on the number of features noticed by the participants' in Stage 2. This might be due to the fact that the measuring of the participants' attitudes was exploratory in nature and, hence, the tool needed further refinement.

## **6. Limitations and further research**

The present study has some important shortcomings which should be mentioned and taken into account for future research. First, with regard to the methodology used, the fact that learners incorporated more features than they reported noticing, points at incomplete note-taking. Therefore, we may conclude that the technique was physically demanding and that consequently it did not probably offer us a completely valid and reliable coverage of participants' thoughts and feelings while noticing. As a result, it is necessary that future studies triangulate the findings employing different introspective methods such as think-aloud protocols. Second, as mentioned in the discussion session, the study mainly focused on the incorporation of solutions to the PFNs noticed in Stage 1, and disregarded the incorporation of other features which were not regarded as problematic at first but which may reflect an overall development of the students' IL or writing skills. As a consequence, we believe that tallying the features incorporated but not related to Stage 1 PFNs would provide us with a more complete insight of the LLP of models as a form of feedback.

Third, although the analysis of the participants' attitudes towards models was exploratory, we acknowledge that the techniques used had some important limitations such as the small number of items included in the questionnaire and the use of few data sources (i.e. the exit questionnaire and the Liker-scale measuring eagerness). Hence, we argue that one of the most important aspects for the future research agenda is to qualitatively investigate learners' beliefs, attitudes and motivation using more sophisticated questionnaires, interviews, observations or diaries; since these individual differences seem to have a great influence on the relationship between noticing, feedback technique and language learning.

Fourth, considering the results of the study, it remains an open question to see to what extent the guidance offered to the participants in the study following Santos et al. (2010), can be considered



a technique leading to a higher quality noticing. It may be the case that learners need more assistance than simply offering them a chart to complete in the comparison stage to reach the substantive noticing level. Taking into account that models were a new feedback tool for most of them and that they involved considerable changes at all textual levels, future studies analysing the influence of guidance should include more detailed instructions to identify the differences and their effects, guided classroom discussions or awareness-raising exercises among others. In relation to this, given the difficulties that some participants, especially the younger ones experienced when understanding the language in the models, it is important that future pieces of research closely match the texts to learners' proficiency level.

Fifth, although the post-test included in the study was delayed, the present piece of research offers a single picture of the feedback-treatment-incorporation process. Hence, the results may be regarded as evidence of uptake rather than acquisition. It would be interesting for future studies to make use of longitudinal designs so as to investigate the long term effects of modelling as well as the evolution in learners' beliefs and attitudes as they become more familiar with the feedback technique.

Finally, the participants in this study worked individually, and thus, the positive effect pair-talk may have on awareness rising and noticing was not considered. Future lines of research could compare individual and collaborative work to analyse the influence hypotheses sharing and testing could have on the incorporation of features in subsequent revisions.

## **7. Conclusions and pedagogical implications**

The present study investigated the role of models as a form of written feedback in a three-stage writing task carried out by 60 EFL learners aged between 13 and 17. Although the findings were in line with most empirical studies carried out with adult learners, including Hanaoka (2007), new insights into the attitudes and beliefs of teenagers towards this feedback technique were also provided. The results evidenced the important noticing function of output, as the participants were able to autonomously notice their language holes, most of which were lexical, to find solutions in the models provided and to incorporate them in subsequent revisions, regardless of their proficiency and type of guidance received. Yet, note that the more proficient learners were found to notice and incorporate more features and that guidance positively affected the number of features noticed in the comparison stage. These results point at the usefulness of providing models to learners immediately after their composing attempt, so that they can use them as feedback tools. It seems that when learners are struggling with the linguistic forms to express their intended meaning, the IL system becomes more ready to notice and incorporate new features. Models, as opposed to other CF techniques, offer the possibility of presenting these new features at that exact time.

Nevertheless, learners showed in general a rather negative attitude towards writing and modelling, partially influenced by their lack of familiarity with the technique and the high physical demand involved in the task. Therefore, if we aim at generalizing the application of this feedback technique to EFL teaching, a number of pedagogical recommendations should be made. First, it is crucial to assess and adapt the models' language level to the learners' proficiency, so that students can understand the texts without feeling frustrated and thus find solutions to their problems and maximally benefit from the feedback. Second, considering that improving the quality of noticing is crucial to enhance the incorporation and retention of features in the long term memory, it would be ideal to provide learners with clear instructions and extended practice

on how to use the models and on how to allocate their attentional resources. For instance, teachers could design follow up pair, group or class discussions to clarify problematic issues. In addition, this collaboration could help reduce the workload involved in writing tasks, thereby prompting a more positive attitude towards writing and modelling. Yet, we believe that further research should be carried out to analyse to which extent this kind of spontaneous writing activities catalysing noticing and FonF can lead to effective L2 learning.

## References

- Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research*, 7, 347-376.
- Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 29-52). Oxford: Oxford University Press.
- Azkarai, A., & García Mayo, M. P. (2015). Task-modality and L1 use in EFL oral interaction. *Language Teaching Research*, 19(5), 550-571.
- Basterrechea, M., & García Mayo, M. P. (2013). Language-related episodes during collaborative tasks: A comparison of CLIL and EFL learners. In K. McDonough, & A. Mackey (Eds.), *Second language interaction in diverse educational contexts* (pp. 25-43). Amsterdam: John Benjamins.
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17, 102-118.
- Bitchener, J. (2012). A reflection on 'the language learning potential' of written CF. *Journal of Second Language Writing*, 21, 348-363.
- Bitchener, J., & Knoch, U. (2010). The contribution of written corrective feedback to language development: A ten month investigation. *Applied Linguistics*, 31, 193-214.
- Byrnes, H., & Manchón, R. (Eds.) (2014) *Task-based language learning – Insights from and for L2 writing*. Amsterdam: John Benjamins.

- Bowles, M., & Leow, R. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation . *Studies in Second Language Acquisition*, 27, 415-440.
- Cánovas Guirao, J., Roca de Larios, J., & Coyle, Y. (2015). The use of models as a written feedback technique with young EFL learners. *System*, 52, 63-77.
- 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.
- Coyle, Y., & Roca de Larios, J. (2014). Exploring the role played by error correction and models on children's reported noticing and output production in a L2 writing task . *Studies in Second Language Acquisition*, (36), 451-485.
- Cumming, A. (1990). Metalinguistic and ideational thinking in second language composing. *Written Communication*, , 482-511.
- Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning*, 51, 1-46.
- Ellis, R. (2009). A typology of written corrective feedback types. *ELT Journal*, 63, 97-107.
- Ellis, R. (2016). Focus on form. A critical review. *Language Teaching Research* DOI: 10.1177/1362168816628627
- 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.

- García Mayo, M. P. (2002). Interaction in advanced EFL pedagogy: A comparison of form-focused activities. *International Journal of Educational Research*, 27, 323-341.
- García Mayo, M. P., & Azkarai, A. (2016). EFL task-based interaction: Does task modality impact on language-related episodes? In M. Sato, & S. Ballinger (Eds.), *Peer interaction and second language learning: Research agenda and pedagogical implications* (pp. 241-266). Amsterdam: John Benjamins.
- Gass, S., & Mackey, A. (2006). Input, interaction and output. *AILA Review*, 19, 3-17.
- Hanaoka, O. (2006a). Exploring the role of models in promoting noticing in L2 writing. *Jacket Bulletin*, 42, 1-13.
- Hanaoka, O. (2006b). Noticing from models and reformulations: A case study of two Japanese EFL learners. *Sophia Linguistica: Working Papers in Linguistics*, 54, 167-192.
- Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11(4), 73-93.
- Hanaoka, O., & Izumi, S. (2012). Noticing and uptake: Addressing pre-articulated covert problems in L2 writing. *Journal of Second Language Writing*, 21, 332-347.
- Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, 7, 255-286.
- Hyland, F. (2010). Future directions in feedback on second language writing: Overview and research agenda. *International Journal of English Studies*, 10, 171-182.

- Hyland, F. (2011). The language learning potential of form-focused feedback. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 159-179). Amsterdam: John Benjamins.
- Izumi, S. (2002). Output, input enhancement and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24, 541-577.
- Izumi, S. (2003). Comprehension and Production Processes in Second Language Learning: In search of the Psycholinguistic Rationale of the Output Hypothesis. *Applied Linguistics*, 24, 168-196.
- Kormos, J. (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing*, 21, 390-403.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. London: Pergamon.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York NY: Longman.
- Lapkin, S., Swain, M., & Smith, M. (2002). Reformulation and the learning of french pronominal verbs in a canadian french immersion context. *The Modern Language Journal*, 86, 485-507.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, D. Coste, R. Ginsberg & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspectives* (pp. 39-52). Amsterdam: John Benjamins.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W.C. Ritchie & T.K. Bhatia (Eds.), *Handbook of language acquisition: Vol. 2. Second language acquisition* (pp. 413-468). New York: Academic Press.

- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 22, 471-497.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching* 46, 1-40.
- Manchón, R. M. (Ed.) (2009). *Writing in foreign language contexts. learning, teaching and research*. Clevedon: Multilingual Matters.
- Manchón, R. M. (Ed.) (2011). *Learning-to-write and writing-to-learn in an additional language*. Amsterdam: John Benjamins.
- Manchón, R. M., & Roca de Larios, J. (2011). Writing to learn in FL contexts: Exploring learners' perceptions of the language learning potential of L2 writing. In R. M. Manchón (Ed.), *Learning-to-write and writing-to-learn in an additional language* (pp. 181-207). Amsterdam: John Benjamins.
- Marsden, E. (2006). Exploring input processing in the classroom: An experimental comparison of processing instruction and enriched input. *Language Learning*, 56(3), 507-566.
- Marsden, E., & Torgerson, C. (2012). Single group, pre-post test research designs: Some methodological concerns. *Oxford Review of Education*. 38(5), 583-616.
- Martínez Esteban, N., & Roca de Larios, J. (2010). The use of models as a form of written feedback to secondary school pupils of English. *International Journal of English Studies*, 10, 143-170.
- Ohta, A.S. (2000). Re-thinking recasts: a learner-centered examination of corrective feedback in the Japanese language classroom. In Hall, J.K. and Verplaeste, I. (Eds.), *The construction of second and foreign language learning through classroom interaction*, Mahwah, NJ: Erlbaum, 47-71-



- Qi, D. S., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing, 10*, 277-303.
- Robinson, P. (Ed.) (2011) *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (2011). Amsterdam: John Benjamins.
- Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on a L2 writing revision task. *Studies in Second Language Acquisition, 29*, 67-100.
- Santos, M., López-Serrano, S., & Manchón, R. M. (2010). The differential effect of two types of direct written corrective feedback on noticing and uptake: Reformulation vs. error correction. *International Journal of English Studies, 10*(1), 131-154.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W.M. Chan, S. Chi, K.N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSic 2010*, Singapore, December 2-4 (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.
- Schmidt, R. (2001). The centrality of attention in SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R., & 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.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8*, 263-300.

- Sheen, Y. (2010). The role of oral and written corrective feedback in SLA. *Studies in Second Language Acquisition*, 32, 169-179.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510-532.
- Storch, N., & Wigglesworth, G. (2010). Learners' processing, uptake, and retention of corrective feedback on writing: Case studies. *Studies in Second Language Acquisition*, 32, 303-334.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidlhofer (Eds.), *Principles and practice in applied linguistics* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook on research in second language learning and teaching* (pp. 471-483). Mahwah NJ: Lawrence Erlbaum.
- Swain, M., & 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., & Lapkin, S. (2002). Talking it through: Two french immersion learners' response to reformulation. *International Journal of Educational Research*, 37, 285-304.
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16, 255-272.
- VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, New Jersey: Ablex.

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.
- Williams, J. (2001). Learner-generated attention to form. In R. Ellis (Ed.), *Form-focused instruction and second language learning* (pp. 303-346). Malden, MA: Blackwell.
- Yang, L., & Zhang, L. (2010). Exploring the role of reformulations and a model text in EFL students' writing performance. *Language Teaching Research*, 14, 464-484.
- Zeigarnik, B. (1999). On finished and unfinished tasks. In W. D. Ellis (Ed.), *A source book of gestalt psychology* (pp. 300-314). London: Routledge.

## Appendices

### Appendix A: Consent form

#### **Idazlanek ingelesaren jabekuntzan izan dezaketen eragina aztertzeke ikerketan parte hartzeke baimen-orria**

Zuen seme-alabek ingeleseko ikasgelan egingo dituzten idatziak aztertzeke baimena eskatzeko zuzentzen gara zuengana.

Ikerlanaren xede nagusia ingelesa hizkuntza atzerritar gisa ikasten duten Batxilergoko ikasleek idazlan bat idazterakoan eta hau eredu batekin alderatzerakoan arreta hizkuntzaren zein ataletan jartzen duten zehaztea da. Bigarren helburua arretaren sakontasunak eta feedback metodo berritzaileagoek ingelesaren garapenean izan dezaketen eragina aztertzea da, honek hizkuntza atzerritarren jabekuntzarako onuragarriak suerta litezken ildo pedagogikoak sortzen lagunduko baitigu.

Zuen seme-alaben datuak era anonimoan aztertuko dira uneoro, haien izen edo irudirik ez baita agertuko ez ikerketa honen emaitzak lortzean ez eta argitaratzean ere. Bildutako datuak ikerketarako bakarrik erabiliko dira eta, beti ere, hizkuntza atzerritarren irakaskuntzan aurrerapausoak emateko asmoarekin.

Edozein argibide edo zalantzarako, idatzi lasai [uloidi002@ehu.ikaslea.es](mailto:uloidi002@ehu.ikaslea.es) [mariapilar.garciamayo@ehu.es](mailto:mariapilar.garciamayo@ehu.es) helbideetara.

Nire seme-alaba \_\_\_\_\_-k ingeleseko ikasgelan egingo duen idatzia aztertzeke baimena ematen dut.

\_\_\_\_\_  
(Sinadura)

\_\_\_\_\_  
(Data)

---

(Guraso edo tutorearen izen-abizenak)

**Formulario de autorización para participar en el proyecto sobre el impacto que las composiciones puedan tener en el aprendizaje del Inglés como lengua extranjera**

Solicitamos su consentimiento para poder estudiar las composiciones que su hijo/hija realizará en la clase de inglés.

El principal objetivo del estudio es determinar en qué aspectos de la lengua centran su atención los alumnos de Bachillerato que aprenden el inglés como lengua extranjera tanto al escribir composiciones como al comparar dichos textos con un modelo. Un segundo objetivo será analizar los efectos que el grado de atención y métodos alternativos de retroalimentación (*feedback*) puedan tener en el desarrollo de la interlengua de estos alumnos. Este análisis nos ayudará a proporcionar guías pedagógicas para el aprendizaje de las lenguas extranjeras en esta etapa educativa.

En todo momento la confidencialidad del alumno está garantizada ya que ni su nombre ni su imagen aparecerán en ningún documento o se publicará en modo alguno. Los datos recogidos se utilizarán solamente para fines de investigación y únicamente con el objetivo de mejorar la práctica educativa.

Si tiene alguna pregunta, escribanos a las siguientes direcciones [uloidi002@ehu.ikaslea.es](mailto:uloidi002@ehu.ikaslea.es) [mariapilar.garciamayo@ehu.es](mailto:mariapilar.garciamayo@ehu.es).

Doy mi consentimiento para el estudio de las composiciones que mi hijo/hija \_\_\_\_\_ realizará en el aula.

\_\_\_\_\_  
(Firma)

\_\_\_\_\_  
(Fecha)

\_\_\_\_\_  
(Nombre del padre, madre o tutor)

## Appendix B: Visual stimuli

Idatz ezazu testu bat beheko irudiak oinarritzat hartuz 1. orrian. Halaber, idatzi ariketan zehar sortzen zaizkizun zalantza eta arazo guztiak 2. orrian.

*Write a text in response to the pictures you see. Use sheet 1 to write the text and sheet 2 to write down all the difficulties you face while doing the exercise.*



**Appendix C: Sheet 1****1. ORRIA / SHEET 1**

Idatz ezazu testu bat emandako irudiak deskribatuz orri honetan.

*Write a text describing the set of pictures given to you in this sheet.*

*Eskerrik asko zure laguntzarengatik  
Thank you for your collaboration*

## Appendix D: Sheet 2

### 2. ORRIA / SHEET 2

**Mesedez, idatzi orri honetan testua idazterakoan izan duzun edozein arazo edo zalantza.** Ongi badatozkizu, ondorengo esaldiak erabili ditzakezu, baina nahiago izanez gero, erabili zure hitzak.

- Ez dakit ..... nola esaten/idazten den ingelesez.
- ..... idatzi dut baina ez nago ziur zuzena ote den.
- Nola esaten da ..... aditza iraganean?
- Ez dakit irudiaren deskribapen zuzena ..... edo ..... den.

*Please write down any difficulty/doubt that you may have had when writing the text.*

*You can use the following expressions if you find them useful, but you can also use your own words:*

- *I do not know how to say/write ..... in English.*
- *I wrote ....., but I am not sure if this is correct.*
- *What is the past tense of .....?*
- *I am not sure whether the picture is describing ..... or .....*

*Eskerrik asko zure laguntzarengatik  
Thank you for your collaboration*



### Appendix E: Likert-scale

Ereduak jasotzeko gogorik ba al duzu? Aukeratu 1etik 5era beheko eskalan eta egin X bat dagokion laukitxoan.

*Choose from a scale of 1 to 5 how eager you are to receive the models.*

1-Bat ere ez <i>Not at all</i>	2- Oso gutxi <i>Slightly</i>	3-Nahikotxo <i>Average</i>	4- Gogotsu <i>Quite eager</i>	5-Oso gogotsu <i>Very much</i>

*Eskerrik asko zure laguntzarengatik  
Thank you for your collaboration*

## Appendix F: The models

Alderatu bi testu eredu hauek zuk idatzitako testuarekin. Zein desberdintasun antzematen dituzu? Erabili **3. orria** desberdintasunak idazteko.

*Compare the following models with your own text. What differences do you notice? Use SHEET 3 to write down the differences.*

### A (Native speaker)

At home, a boy is lying on the couch playing video games. He is staring at the screen and focusing on doing the best he can. In the park, two young children are playing on swings. The girl is sitting normally while the boy is swinging face-first. Nearby, a group of kids is rolling a large rubber ball around the playground. Smiling and laughing, they seem to be enjoying themselves as they draw the envy of the rest of the kids. Back inside, a boy is working on a laptop. Listening to headphones, he might be playing a game for fun or working with a study tool. Another boy is sitting on the floor watching television. While enjoying the program, he is also sipping a drink from his cup. On a lawn outside, three girls are playing with hula hoops. They are certainly having a good time as they try to keep the hula hoops up as long as possible. However, it can be a challenging activity as two of the girls are finding it difficult.

### B (Proficient L2 speaker)

In the first picture, a ten-year-old boy is sitting on a sofa playing computer games. He looks bored, maybe due to the fact that he is playing alone indoors, probably at home. In the second one, two kids are playing on swings in a park. There is a big field and a forest in the background of the picture. In the following photograph, a group of children is playing with a big ball in a park. In the background of the image we can see more children and some adults supervising them as well. In the fourth photo, a boy is studying or playing with a laptop. He has got some headphones, thus he may be listening to music too. The kid seems to be in the kitchen because there is a microwave behind him. In the next picture, a little boy is watching TV while having a drink. There are some toys and a handcraft on the floor, so he could be in a playroom. Finally, three girls who may be siblings or friends are playing with the hula hoops. They may be on holidays in a campsite, as there are some tents at the background.

## Appendix G: Sheet 3 – Unguided noticing group (UNG)

### 3. ORRIA/ SHEET 3

**Mesedez, alderatu zure testua emandako bi testu erduekin (A eta B) eta idatz itzazu zure testuaren eta bi erduen artean antzematen dituzun desberdintasunak.**

Ongi badatozkizu, ondorengo esaldiak erabili ditzakezu, baina nahiago izanez gero erabili zure hitzak:

- Ez nekien ..... hitza nola esaten/idazten zen eta A/B testuak honela ..... idazten/esaten du.
- Nik ideia hau honela ..... azaldu dut, baina A/Bk honela ..... azaltzen du.
- Uste nuen ..... aditzaren iragana ..... zela, baina A/Bn honela agertzen da .....

*Please, compare your text with the two models provided (A and B) and note down all the differences that you find.*

*You can use the following expressions if you find them useful, but you can also use your own words:*

- *I could not say/write the word ....., but A/B puts it this way .....*
- *I have expressed the idea ..... this way, but A/B puts it this way .....*
- *I thought the past tense of the verb ..... was ....., but A/B writes it this way .....*

*Eskerrik asko zure laguntzarengatik  
Thank you for your collaboration*

**Appendix H: Sheet 3 – Guided noticing group (GNG)**

**3. ORRIA/ SHEET 3**

**Mesedez, alderatu zure testua emandako bi testu erduekin (A eta B). Idatz itzazu zure testuaren eta bi erduen artean antzematen dituzun desberdintasunak ondorengo koadroan.**

*Please, compare your text with the two models provided (A and B) and note down all the differences that you find following the table below.*

Desberdintasuna / <i>Difference</i>		Mota / <i>Type</i>			Ereduko aukera egokiagoa iruditzen zaizu? Zergatik? / <i>Do you think the model's alternative is more appropriate? Why?</i>
Zure testuan / <i>In your text</i>	Ereduetan / <i>In the model(s)</i>	Hiztegia <i>Lexicon</i>	Gramatika <i>Grammar</i>	Adierazpena <i>Discourse</i>	
Ez nekien ..... hitza nola esaten/idazten zen <i>I could not say/write the word .....</i>	A/B erduak honela ..... idazten/esaten du. <i>A/B puts it this way .....</i>	X			Bai. Ez nuen hitz hori ezagutzen ingelesez. <i>Yes. I did not know this English word.</i>
Nik ideia hau honela ..... azaldu dut <i>I have expressed the idea ... this way ....</i>	A/B erduak honela ..... idazten/esaten du. <i>A/B puts it this way ....</i>			X	Ez. Nire interpretazioak hobeto azaltzen du irudia. <i>No. I think that my interpretation is more in line with what is seen in the picture.</i>
Uste nuen ..... aditzaren iragana ..... zela. <i>I thought the past tense of the verb ... was ....</i>	A/Bn honela agertzen da ..... <i>A/B writes it this way ....</i>		X		Bai. Uste dut erdua zuzen dagoela. <i>Yes. I think that the model is right.</i>


*Eskerrik asko zure laguntzarengatik*  
*Thank you for your collaboration*

**Appendix I: Sheet 4****4. ORRIA / SHEET 4**

Berridatz ezazu testu bat emandako irudiak deskribatuz orri honetan.

*Rewrite your text describing the set of pictures given to you in this sheet.*

*Eskerrik asko zure laguntzarengatik  
Thank you for your collaboration*

## Appendix J: Exit questionnaire

### AMAIERAKO GALDETEGIA

Izen-abizenak:

*Mesedez, erantzun ondorengo galderak erantzun egokia aukeratuz edo dagokion lekuan zure ideiak ahalik eta zehatzen azalduz.*

*Gogoratu galdetegia guztiz anonimoa dela. Zure identitatea ez da inolaz ere argitaratuko, izen bakoitza kode batekin ordezkaturiko baita.*

#### 1. Eman iritzia egindako ariketa honi buruz: **Testu bat idatzi irudi sorta bat oinarri hartuta.**

a. Eskolan, ingeleseko klaseetan, horrelako ariketak egin ohi al dituzu?

Bai  Batzuetan  Ez

b. Ariketa gustatu al zaizu?

Bat ere ez  Gutxi  Besterik gabe  Nahikotxo  Asko

Zergatik? \_\_\_\_\_

c. Zure ingelesa hobetzeko erabilgarria dela uste al duzu?

Bat ere ez  Ez oso erabilgarria  Besterik gabe  Nahikoa erabilgarria  Oso erabilgarria

Zergatik? \_\_\_\_\_

d. Gustatuko al litzaizuke irakasleak horrelako ariketak erabiltzea aurrerantzean?

Bai  Ez

Zergatik? \_\_\_\_\_

#### 2. Eman iritzia egindako ariketa honi buruz: **Zure testua beste testu eredu batzuekin alderatu.**

a. Eskolan, ingeleseko klaseetan, horrelako ariketak egin ohi al dituzu?

Bai  Batzuetan  Ez

b. Ariketa gustatu al zaizu?

Bat ere ez  Gutxi  Besterik gabe  Nahikotxo  Asko

Zergatik? \_\_\_\_\_

atzean jarraitzen du



c. Zure ingelesa hobetzeko erabilgarria dela uste al duzu?

Bat ere ez	Ez oso erabilgarria	Besterik gabe	Nahikoa erabilgarria	Oso erabilgarria
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Zergatik? \_\_\_\_\_

d. Gustatuko al litzaizuke irakasleak horrelako ariketak erabiltzea aurrerantzean?

Bai

Ez

Zergatik? \_\_\_\_\_

Oharren bat idatzi nahi baduzu:

---

---

---

*Eskerrik asko zure laguntzarengatik*



## EXIT QUESTIONNAIRE

Name: \_\_\_\_\_

*Please answer the questions **by choosing the right answer or by explaining your thoughts as fully as possible**, where necessary. Remember that the questionnaire is **anonymous**. Your name will never be published, as it will be replaced by a code.*

**1. Give your opinion on the following activity: Write a text in response to some pictures.**

a. Do you usually do this type of activities in your English classes, at school?

Yes                                       Sometimes                                       No

b. Did you enjoy the activity?

Not at all                      Not so much                      So-so                      Quite a lot                      Very much

Why? \_\_\_\_\_

c. Do you think it might be useful to improve your English?

Not at all                      Not so much                      So-so                      Quite a lot                      Very much

Why? \_\_\_\_\_

d. Would you like your teacher to continue using this type of activity?

Yes                                       No

Why? \_\_\_\_\_

**2. Give your opinion on the following activity: Comparing your text with models.**

a. Do you usually do this type of activities in your English classes, at school?

Yes                                       Sometimes                                       No

b. Did you enjoy the activity ?

Not at all                      Not so much                      So-so                      Quite a lot                      Very much

Why? \_\_\_\_\_

c. Do you think it might be useful to improve your English?

Not at all                      Not so much                      So-so                      Quite a lot                      Very much

Why? \_\_\_\_\_

It continues behind



d. Would you like the teacher to continue using this type of activity?

Yes

No

Why? \_\_\_\_\_

Please write anything else you would like to add - suggestions, comments etc:

---

---

---

*Thank you for your collaboration*