

Focused task-based peer interaction in the young EFL classroom: enhancing explicit knowledge and production of present continuous

Elisabet Pladevall-Ballester, Universitat Autònoma de Barcelona

Eloi Puig-Mayenco, King's College London

Abstract

This study explores whether explicit knowledge and production of present continuous are equally affected by a task-based peer-interaction intervention with and without pre-task grammar and interactional instruction. Four groups of 4th-grade Catalan-Spanish bilingual EFL learners, aged 9-10, participated in an 8-week pedagogical intervention (50 minutes per week), while a fifth control group only participated in the testing sessions. Groups were divided according to the type of pre-task instruction received on a weekly basis: Grammar + Interaction Strategy (n=22), Grammar (n=23), Interaction Strategy (n=22), and Task-Only (n=23). The four groups also participated in task-based peer interaction every week (both collaborative and information-gap tasks). All groups took part in a bi-modal, untimed Grammaticality Judgement Task (GJT) and carried out a collaborative picture narrative task before and after the intervention. In the GJT, the analysis revealed that only those groups who received either mixed grammar-interaction or grammar-only instruction showed significant improvement in accuracy, modulated by condition. In the production task, the analysis showed that the four treatment groups exhibited an increase in both production attempts and accuracy and differed from the control group after the intervention, though only the groups that received pre-task grammar instruction improved significantly. Results will be discussed in relation to the pedagogical intervention and the development of explicit knowledge and production of present continuous.

Keywords: task-based peer interaction, young learners, EFL, pre-task instruction

1. Introduction

One of the central questions in Task-based Language Teaching (TBLT) research is how tasks may be more conducive to language development and whether this development can be enhanced by manipulating certain task conditions at different moments of the task cycle (Ellis et al., 2020). One of these moments, the pre-task stage, motivates and prepares learners for task performance and has been particularly studied in relation to task performance itself, as well as motivation and learning improvement (Ellis, 2009; Li et al., 2016; Ortega, 1999; Willis & Willis, 2007, among others). Such preparation might come in the form of pre-task form-focused instruction in an explicit or implicit manner, and its use is based on the claim that the strongest version of TBLT without a focus on form (FonF) does not promote the use and/or learning of specific language structures (Ellis, 2016; Sato, 2010; Willis, 1996). Focused tasks might also prompt the use of a given target structure, allowing researchers to explore whether and to what extent the learning and use of such features occur (Doughty & Varela, 1998; Ellis et al., 2006; Lyster, 2004, among others). However, focused tasks might not make it essential for learners to produce a certain grammatical structure (Loschky & Bley-Vroman, 1993), and learners might not be developmentally ready to learn it through task performance (Pienemann, 2005). Pre-task form-focused instruction might then enhance the learning potential of focused tasks.

From a pedagogical perspective, within Long's (2015) TBLT model, pre-task instruction should be rejected since FonF might occur reactively in the form of corrective feedback while learners perform the task. FonF at the pre-task stage might lead learners to carry out a task as if it

were a language practice exercise rather than a communicative, meaning-focused activity (Ellis, 2003; Willis & Willis, 2007). Other researchers consider explicit teaching of target structures as a required step for learners to use them in task performance and/or for their subsequent development. This perspective is theoretically supported by DeKeyser's (2007; 2015; 2017) Skill Acquisition Theory, which posits that explicit knowledge of a given L2 feature precedes practice activities where such knowledge can be proceduralized and eventually automatized through repeated practice in communicative contexts. Pre-task instruction would then promote learners' declarative knowledge, which would be put into practice during subsequent task performance.

The effects of task-based peer interaction have been widely researched among adolescent and adult learners in instructed settings or children in immersion contexts. Yet child foreign language learners still represent an under-researched population (Collins & Muñoz, 2016; García Mayo, 2021) who can and should inform primary classroom language pedagogy. Young EFL learners have been shown to be able to participate in task-based peer interaction, negotiate for meaning and form in the L2, and produce instances of conversational adjustments, repetitions, and collaboration and feedback strategies (Azkarai & Imaz Agirre, 2016; Azkarai & Kopinska, 2020; Calzada & García Mayo, 2020; 2021; García Mayo & Imaz Agirre, 2017; García Mayo & Lázaro Ibarrola, 2015; Lázaro Ibarrola & Azpilicueta Martínez, 2015; Pinter, 2007; Pladevall-Ballester & Vraciu, 2020; Pladevall-Ballester, 2021, among others). Primary school learners in FL contexts need rich and varied input as well as opportunities to use language in a meaningful way in task performance. Yet, they are also able to activate their explicit learning mechanisms and gradually become more metalinguistically aware learners (Holmes & Myles, 2019). In such contexts, pedagogical interventions in the form of explicit instruction followed by communicative practice might provide the ideal scenario to enhance learning of specific grammatical features, as well as foster communication (Benati & Schwieter, 2019; Nassaji, 2017).

Although research on child EFL peer interaction is increasing, the majority of studies are not classroom-based and do not include any type of classroom intervention or measures of L2 knowledge and use. Knowing to what extent focused task-based peer interaction and explicit instruction followed by task performance result in more frequent use and effective acquisition of the target feature would greatly contribute to our understanding of how to maximize the use of tasks in the primary classroom. This study aims to analyze whether including focused task-based peer interaction in weekly EFL primary school sessions has any effects on the learners' explicit knowledge and use of the present continuous tense and whether adding pre-task explicit instruction of the target structure (Grammar group), pre-task explicit instruction of interactional and corrective feedback strategies (Interaction strategy group), or both (Grammar + Interaction strategy group) during the weekly teaching intervention has any further effects compared to only participating in task-based peer interaction (Task-only group) or not participating in it at all (Control group).

2. Pre-task instruction and its effects on explicit L2 knowledge and production

A few studies have explored whether pre-task instruction impacts learners' oral production of an FL target structure, general task performance, or explicit and implicit L2 knowledge. Although previous research is not abundant, similar studies based on in-class interventions with young EFL learners are, to the best of our knowledge, inexistent.

Regarding the impact of pre-task instruction and oral tasks on L2 knowledge, Li et al. (2016) explored the development of explicit and implicit knowledge of the past passive by Chinese middle school EFL learners. The four experimental groups participated in a two-hour treatment where they performed two dictogloss tasks in groups and an oral reporting session. One group only performed the oral tasks, a second group received explicit instruction and carried out the tasks, a third group received within-task feedback, and a fourth group received both within-task feedback and pre-task instruction. A control group only took part in the pre-, post-, and delayed

post-tests, namely a Grammaticality Judgement Task (GJT) and an Elicited Imitation Test (EIT). Relevant to our study are the results of the GJT and the comparison between the first and second groups, which showed that the group that received explicit instruction prior to the performance of the tasks obtained greater gains than the group who just performed the tasks. In a very similar study, Li et al. (2018) compared three groups of Chinese EFL learners who were split into control, task-only, and explicit instruction+task conditions on their explicit and implicit knowledge of the past passive. Using the same tests, the researchers found that the condition that obtained higher results was the one with explicit instruction on the measure of explicit knowledge. More recently, Khezrlou (2021) investigated the immediate and delayed effects of task repetition (picture narrative tasks) with and without explicit instruction on the explicit and implicit knowledge of past tense among young adults at a language institute in Iran by means of an untimed GJT and an EIT. Two groups participated in the study, the first one being exposed to task repetition only and the second one receiving task repetition and explicit instruction between the first and the second task repetition on consecutive days. Both groups were pre-tested, post-tested after each task repetition, and post-tested two weeks after the last task repetition. GJT findings indicate that the group that received explicit instruction together with task repetition improved significantly at the immediate and delayed explicit post-tests, whereas the group that was only exposed to task repetition improved only after the first immediate test but obtained lower results afterwards.

As for the effects of pre-task instruction on the use and production of target forms, De la Fuente (2006) compared immediate and delayed retrieval of L2 Spanish vocabulary in three groups, one which had had a traditional PPP lesson, another one following a task-based lesson with incidental FonF and a third one also following a task-based lesson but incorporating an explicit FonF component at the post-task stage. No immediate effect of the type of lesson was observed but at the delayed post-test, the two task-based groups outperformed the traditional one and the group which had received explicit instruction at the post-task stage obtained higher scores than the one without explicit instruction. The author argues that although the traditional PPP group had received explicit instruction, it did not offer opportunities for output production and explicit FonF was more effective after task performance. Other studies, such as Mochizuki and Ortega (2008) did find pre-task explicit instruction to foster the use of the target structure. Japanese high-school EFL learners carried out an oral story-retelling task under three conditions: no planning, unguided planning and guided planning focusing on relative clauses, the target form. Results showed that the group with pre-task linguistic guidance produced more and more accurate relative clauses than the other two groups, while their global complexity and fluency were similar. Another study which showed positive results of pre-task activities on task performance and the use of the target structure is Van de Guchte et al. (2017). In this study, ninth-grade L2 learners of German were assigned to two groups, both of which watched two videos of two girls describing the school cafeteria before performing the task (where they would have to describe the school cafeteria as well), but one group watched the videos with a focus on language (i.e. through a worksheet which included questions on the target feature, namely locative prepositions) and the other one with a focus on content (i.e. through questions on the girls' persuasiveness and the structure of the presentation). Findings showed that the group that focused on language produced the target structure more often and more accurately, while students focusing on content at the pre-task outperformed the language-focused group in terms of syntactic complexity. The authors concluded that depending on the lesson objectives, pre-task planning should be directed to language or content or both to maximise language development in a more balanced way.

In an attempt to contribute additional evidence to the effects of pre-task instruction on subsequent focused task performance and drawing on the same target structure and tasks as in their previous studies, Ellis et al. (2019) assigned their learners to an explicit instruction+task group, who was given a grammar lesson on the passive voice, practice activities and then carried two dictogloss tasks, and a task-only group, who only performed the tasks. The oral narratives in

the dictogloss tasks showed that the group that received explicit instruction produced the target structure more often but not more accurately, and the task only group outperformed the explicit instruction group on overall task performance including global complexity, accuracy and fluency and along the lines of Van de Guchte et al. (2017). The authors concluded that pre-task instruction might interfere with task performance and might be best carried out at the post-task stage but they also acknowledged that their learners were not familiar with the structure and that they only performed the tasks once. More intensive practice to proceduralize the structure might not affect language use in task performance to the same extent.

Shintani (2019) extended Ellis et al.'s (2019) study by including a condition receiving only metalinguistic instruction (ME) (as well as a task only condition – Task Only, and a metalinguistic instruction+task condition – ME+Task) and multiple treatment sessions where the groups received different treatment. A group of Japanese university EFL learners were tested on an error correction task measuring their explicit knowledge of the target structure (i.e. past conditional) and a focused story reconstruction task at post and delayed post treatment stages. As for explicit knowledge, the ME+Task condition obtained the highest results, followed by the ME condition and the Task Only condition. The ME+Task condition also obtained the highest gains in production and accuracy, but it was followed by the Task Only condition in this case, the ME condition being the least effective one. The author suggests that pre-task instruction helps the learners produce the target structure during communicative practice thus contributing to knowledge proceduralization and lending support to DeKeyser's (2015) Skill-Acquisition Theory.

Research on the effects of pre-task interaction strategy instruction on grammatical development and language use are even scarcer. Sato and Loewen (2018) analysed the impact of metacognitive instruction on clarifications requests and recasts together with activities including the two types of implicit corrective feedback on university EFL learners' use of third person singular *-s* and possessive determiners *his/her* in picture description tasks. Metacognitive instruction with corrective feedback activities was found to lead to higher use of the target structures than either type of corrective feedback activities alone. Research on interactional and metacognitive instruction has mainly explored its effects on the actual use of corrective feedback (Fujii, Ziegler & Mackey, 2016; Sato & Lyster, 2012) and interactional strategies (Sato, 2020) in subsequent tasks, but also on learners' accuracy and fluency and their beliefs about corrective feedback (Sato & Lyster, 2012, a.o.). With young learners, Sato and Dussuel Lam (2021) found that interactional instruction led to a higher use of the L2 and a greater metacognitive knowledge of how oral communication works, although the children's willingness to communicate was not seen to improve.

With the aim of contributing new data on the effectiveness of pre-task grammar and interactional instruction, the present study seeks to explore whether explicit knowledge and production of present continuous in young EFL learners are affected after a task-based peer interaction intervention with and without grammar and interactional instruction. Present continuous was chosen as it is one of the key grammar components of the 4th grade EFL curriculum and it is formally taught in this grade for the first time. The study had a pre/post test design with an 8-week teaching intervention including a control group and four experimental groups (with four types of intervention), and two main research questions were posed:

- (1) Does focused task-based peer interaction lead to increased explicit knowledge and production of present continuous among young EFL learners?
- (2) To what extent does the use of explicit pre-task instruction affect young EFL learners' knowledge and production of present continuous?

Following previous similar research conducted with adults and adolescents, we expect explicit knowledge gains and more frequent (although not necessarily more accurate) use of the target

structure in grammar instructed groups (Ellis et al., 2019; Li et al., 2016; Li et al., 2018; Khezrlou, 2021; Shintani, 2019) and therefore we hypothesise that explicit grammar pre-task instruction will have an effect on the young learners' knowledge and use of the present continuous tense.

3. Methods

3.1 Participants

The participants were 129 young (N=129) EFL 4th grade learners of English (65 females and 64 males), aged 9-10, with an A1 level of English proficiency ($M=36.84$; $SD=14.39$) as measured in the Oxford Placement Test for Young Learners¹, studying in three Catalan state schools sharing sociodemographic characteristics in the Metropolitan area of Barcelona, Spain. The learners from the different schools were comparable in terms of number of hours of previous and current in-class exposure to English and in terms of level of English.² The learners were Catalan/Spanish bilinguals, and the accumulated in-class exposure to English prior to the study was approximately 385 hours over the span of six years (since the beginning of pre-school). As for the role of English in their daily lives, 56.21% of the learners had taken extracurricular English lessons prior to the study, versus 43.79% who had never taken them (with fairly similar proportions across groups). Along similar lines, 47.23% were occasionally involved in leisure activities in English (e.g., videogaming, watching TV), whereas 52.76% were not. One of the three schools acted as a Control group (C) (N=39), which only took the pre- and post-tests. The children in this group were not taught the present continuous and did not participate in any kind of peer-interaction tasks during the study. The children of the other two schools were divided into four intervention groups, which were named according to the different types of intervention they received: the Grammar group (Gr) (N=23), the Interaction strategy group (I) (N=22), the Grammar+Interaction strategy group (Gr+I) (N=22), and the Task-only group (T) (N=23) (see section 3.2). The teachers created the groups and ensured that they were balanced in terms of cognitive and linguistic abilities. All groups were pre- and post-tested with an untimed Grammaticality Judgement Task (GJT) on the children's explicit knowledge of the present continuous tense and carried out a present continuous-focused collaborative picture narrative task to explore their use of the target structure. The four experimental groups participated in the teaching intervention explained below.

3.2 Intervention

The teaching intervention lasted 8 weeks with a 50-minute session per intervention group each week. These weekly sessions were divided into two parts of 25 minutes each and were carried out by the same researchers throughout the intervention. The first 25 minutes were different in each group while the remaining 25 minutes were exactly the same for all four experimental groups.

As for the first 25 minutes, the Task-only group (i.e. T) did not receive any type of pre-task instruction and did in-class EFL games or vocabulary activities which were unrelated to the target

¹ In addition to the Oxford Placement Test for Young Learners, all participants took the ELIAS Grammar Test II (Kersten, et al. 2012), which is a proficiency test designed for young learners that has not been standardized to date. We ran correlational analysis and our results show that the proficiency score in the Oxford Placement Test for Young Learners was strongly correlated with the proficiency score in the ELIAS Grammar Test II, $r(1243) = .62$, $p < .001$, 95% CI [.61; .71].

² A two-way ANOVA was conducted to explore the effects of SCHOOL and INTERVENTION on proficiency scores during the pre-test. There were no significant main effects of SCHOOL [$F(2, 113) = 1.209$, $p = .274$] or INTERVENTION [$F(4, 113) = .739$, $p = .566$], nor was there a significant interaction between the two [$F(3, 113) = 0.596$, $p = .619$].

structure but ensured the children received the same amount of exposure to English. The other three experimental groups (Gr, Gr+I and I) did receive pre-task instruction during this part of their weekly intervention sessions. Such pre-task instruction was structured around the Present and Practice stages of the Present, Practice, Produce (PPP) approach to instruction (see Nassaji & Fotos, 2011; Shintani, 2013, among others) in which children were presented with the focused structures in an explicit way and then practiced them using a variety of activities depending on the group. The Grammar group (i.e. Gr) was provided with metalinguistic explanations and rules for present continuous verb forms in affirmative and negative syntactic contexts and L1/L2 contrasts every week (i.e. Present). This was followed by the controlled production stage (i.e. Practice), where they carried out fill-in-the gap, error correction, translation, sentence completion and jumbled sentences exercises as well as dictation and running-dictation or competition games. The Interaction strategy group (i.e. I) was given explicit instruction of interactional strategies (i.e. conversational adjustments, lexical requests and corrective feedback strategies) through the use of glossaries which included examples and L1 translations of clarification requests, comprehension checks, confirmation checks, repetitions, recasts and lexical requests. We also illustrated such strategies with video modelling with videos we had created specifically for the sessions where children could watch other children interacting and using such strategies and expressions (i.e. Present). Children were asked to identify those expressions in the videos by raising flashcards depicting the expressions and also to repeat the expressions. They also participated in fill-in-the gap exercises, running dictation or sentence completion exercises so as to practice the expressions in a controlled way (i.e. Practice). The Grammar+interaction strategy group (i.e. Gr+I) was provided with a reduced version of the same activities carried out in the previous two groups, since the amount of time was the same. The corrective feedback strategies taught and practiced in this group focused on the target structure of the study (i.e. present continuous) whereas the other interactional strategies did not.

For all the intervention groups, the remaining 25 minutes in every session consisted in task-based peer interaction in the form of procedural repetition of both a collaborative picture narrative and an information-gap focused task carried out in pairs. This was the Produce stage of the PPP approach to instruction used in every session for the pre-task instruction groups (i.e. Gr, I, Gr+I), where children in these groups could actually use the structures (i.e. present continuous and/or interactional and corrective feedback strategies depending on the group) they had been previously taught. The Task-only group also participated in such tasks (without having received any kind of pre-task instruction). The tasks were the same in all experimental groups and all of them were focused on the present continuous tense, the target grammar structure of the study. The picture narrative tasks were collaborative story telling tasks where children were given six to eight vignettes which they had to describe as they told the story together. The prompts were all in the present continuous and vocabulary help was given if needed. The information-gap tasks consisted of a set of questions the students had to answer in relation to different timetables or different sets of pictures each member of the pair had in order to answer a final question with the information learners had been exchanging. Again, all the prompts were in the present continuous and vocabulary help was given if required.

3.3 Tasks

3.3.1 Grammaticality Judgement Task

An untimed Grammaticality Judgement Task (GJT) was designed to tap into the explicit knowledge of present continuous formation in English.³ The task included a total of 36 items

³ The materials for the GJT will be uploaded to the IRIS database (www.irisdatabase.org.)

divided across six conditions (6 items each target condition and 6 items for each of the control conditions) where we manipulated type (control vs. target), grammaticality (grammatical vs ungrammatical) and within the ungrammatical sentences we manipulated the place where the ungrammaticality was introduced (wrong auxiliary [no-yes], wrong inflection in the main verb [yes-no] or wrong auxiliary + wrong inflection [no-no]). For the three ungrammatical conditions, we decided to manipulate where the ungrammaticality was introduced, in case there was an effect related to the type of ungrammaticality. This included whether the sentence lacked the present participle, whether the auxiliary agreement was incorrect, and whether the ungrammaticality arose from both having the wrong auxiliary and lacking the present participle. Table 1 below contains an outline of the how such manipulation led to six different conditions with some example items for each condition.

Table 1

Experimental conditions and example items in the GJT

Type	Grammaticality	Condition	Example items	N=
target	grammatical	Yes-Yes	She is eating pizza at the moment.	6
	ungrammatical	Yes-No	I am read a book right now.	6
		No-Yes	We isn't listening to music now.	6
		No-No	Ann are have a shower right now.	6
control	grammatical	na	My mom doesn't play football.	6
	ungrammatical	na	Susan to the park go can't.	6

As shown in Table 1 above, in addition to the 24 target items, the GJT contained two unrelated conditions with 12 items (ratio of 1-to-2) that served as both distractors and controls. Within these two conditions, one contained 6 grammatical sentences and the other one contained 6 ungrammatical sentences. These items were used to ensure participants could do the task appropriately and that they were able to detect ungrammaticality. We carefully designed the ungrammatical sentences so that they contained clear morphosyntactic violations that the participants would be sensitive to. This was particularly important as the proportion of accuracy in the control conditions was taken as an inclusion-criteria. All items in the sentences were level-appropriate and we carefully controlled for the verbs used throughout the experiment to make sure participants knew their meaning.

Participants first completed four practice items before the main task began. The 36 experimental items were then presented randomized for each participant, and in a written and recorded oral form. Participants were asked to judge the grammaticality using a binary response mode. Alongside the item, they saw three different emojis signalling different responses: happy = grammatical, sad = ungrammatical and confused = I am not sure. Once a participant selected one of the emojis, they were then asked to correct the sentence if they thought the sentence was ungrammatical.⁴ Participants had been given exact instructions as to what the emojis meant and did four practice items alongside the researcher in the room so that they could ask any pertinent questions. Once they felt comfortable with the task, they proceeded at their own pace monitored by the researcher.

3.3.2 Production task

⁴ The analysis of the corrections will be reported in a further study.

A collaborative oral picture narrative task was created to test the frequency and accuracy of production of present continuous at pre and post intervention stages. The task was focused and prompted the use of present continuous. It was called *Sarah's surprise* and consisted of six picture vignettes where a girl named Sarah leaves home leaving her dog alone. A series of animals visit the dog, carry out different activities and they all end up having a party at Sarah's place (see Appendix 1). The children were asked to have a look at the pictures and tell the story in pairs with the following prompt: *This is Sarah. What's happening in the pictures? What's Sarah doing? What are the animals doing in each picture?* Learners were grouped into dyads (the same for pre and post test) by the teachers, who ensured the members of each dyad had different linguistic and cognitive skills so that the pairs were of mixed proficiency type. Children were prompted to use English throughout the task and were reminded they could resort to one another to appeal for help.

3.4 Procedures

The tasks took place in a classroom setting before and after the intervention. The GJT was completed by each participant using their personal computer. We created and administered the GJT using Gorilla Experiment Builder (www.gorilla.sc) (Anwyl-Irvine et al. 2020) as it has been shown to be highly reliable to conduct experiments where reaction times are recorded (Bridges et al. 2020). The task involved the use of the same set of experimental items in both the pre-test and post-test. However, the items were randomized for each participant and during each testing session. The decision to use identical items was made to control for any effects arising from a lack of lexical knowledge. Given that the pre-test and post-test were conducted 10 weeks apart, effects resulting from participants being re-exposed to the same items are unlikely to be observed. The learners also took part in the same collaborative picture narrative task explained in section 3.3.2 at pre and post intervention and they were videorecorded. Learners' oral production in the picture narrative task was transcribed according to the CHAT conventions from the Child Language Data Exchange System (CHILDES) database (MacWhinney, 2000). Additionally, the children completed the Oxford Placement Test for Young Learners, the ELIAS grammar test II (Kersten, et al. 2012) and a questionnaire tapping into their perceptions and motivation towards peer interaction that was designed as part of the project but is not reported in the study. Additionally, caregivers also answered a background questionnaire on the linguistic background of the participants. All participants as well as their caregivers provided informed consent to take part in the study and all procedures regarding this study, as well as the main project that this study belongs to, were given ethical approval by the pertinent research ethics committee.

4. Results

4.1. GJT results

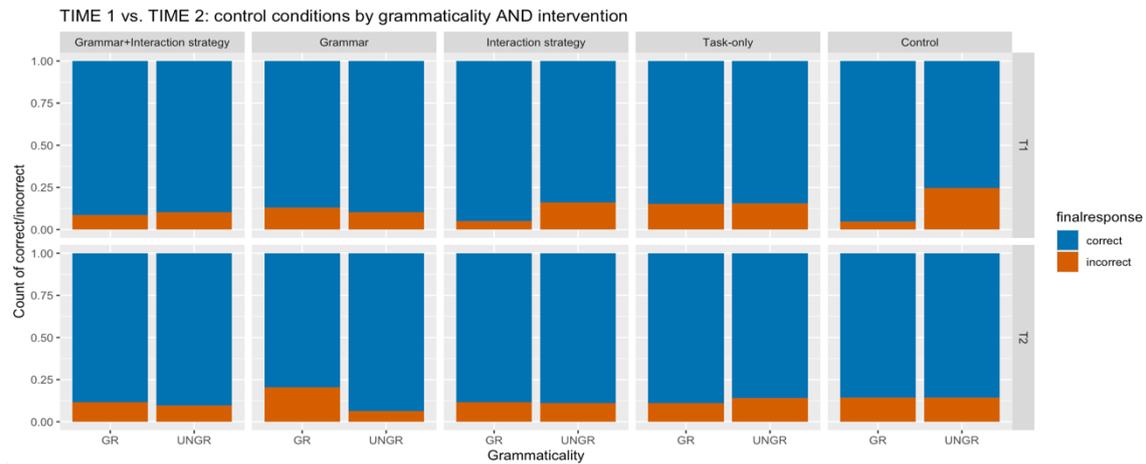
To explore the results in the GJT, we coded the responses for accuracy. Correct responses indicated acceptance of a grammatical sentence or rejection of an ungrammatical one, while incorrect responses indicated rejection of a grammatical one or acceptance of an ungrammatical one. In the coding process, correct responses were assigned a score of 1, while incorrect responses received a score of 0. Responses indicating 'I am not sure' were excluded from the analysis because, although technically incorrect, they did not offer insights into learners' intuitions about sentence grammaticality. Additionally, although the GJT collected metalinguistic corrections from participants, these corrections were not incorporated into the scoring procedure and have not been analyzed in the current chapter. Before conducting any statistical analysis, we applied an exclusion criterion, excluding participants with less than 66% accuracy across the two control conditions. This step was essential to ensure the efficacy of the task with this participant profile. By including only those participants who achieved 66% accuracy or higher in the control

conditions, we ensured their capability in detecting both grammaticality and ungrammaticality. Given that these sentences were crafted to clearly and directly assess grammaticality, we decided to set an accuracy threshold of 66%. This criterion indicates that evaluators detected grammatical and ungrammatical sentences correctly in at least two-thirds of the control items. This exclusion resulted in the removal of nine participants, comprising 6.97% of the total sample size.

We first explored the results of the two control conditions descriptively, and the results showed that participants had high levels of accuracy in accepting grammatical sentences (Time 1: grammatical = 90.81 [$SD=28.89$]; Time 2: grammatical = 86.31 [$SD=34.41$]) and high levels of accuracy in rejecting ungrammatical sentences (Time 1: ungrammatical = 83.83 [$SD=36.74$]; Time 2: ungrammatical = 88.46 [$SD=31.96$]). These overall patterns held across intervention groups, as shown in Figure 1 below. It is important to note that the high percentages of rejection in the ungrammatical condition indicate that participants could detect ungrammaticality in the task, and there was no bias towards positive responses in the data.

Figure 1

Proportion (0 to 1) of accuracy in the control conditions presented by grammaticality, intervention and time.



Turning to the four target conditions, we first present the results descriptively as seen in Table 2 and Figure 2 below. Participants, overall, had high proportions of accuracy at both Time 1 and Time 2 for the grammatical (yes-yes) condition. Across the three ungrammatical conditions, participants had fairly low proportions of accuracy at Time 1 across intervention groups with some moderate improvement modulated by intervention and condition, which we now turn to explore statistically.

Table 2

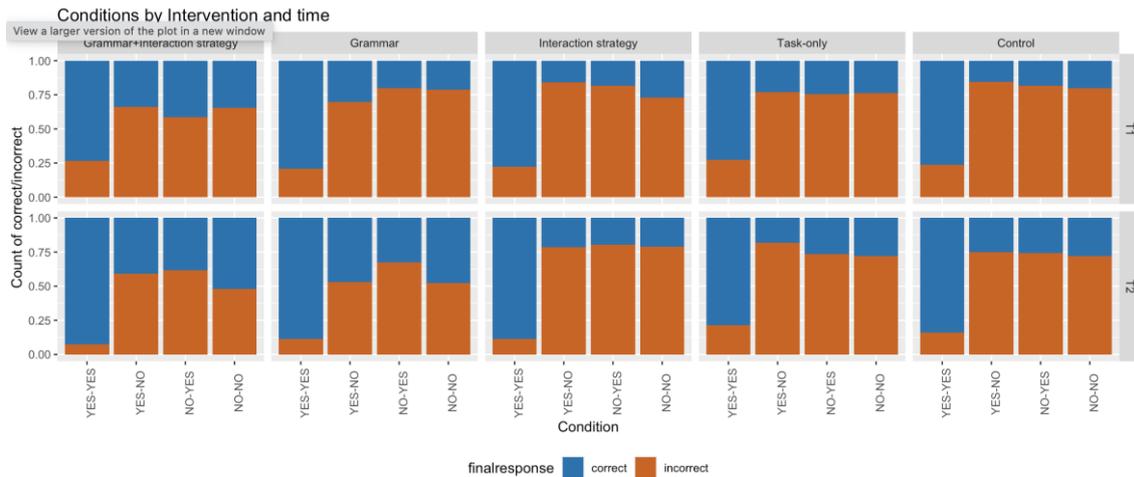
Descriptive statistics [Proportion 0-to-1, Mean(SD)] for the across conditions presented by intervention groups, conditions and times.

	Gr+I		Gr		I	
	T1	T2	T1	T2	T1	T2
YES-YES	.74 (.44)	.93 (.26)	.79 (.41)	.89 (.32)	.78 (.42)	.89 (.32)
YES-NO	.34 (.48)	.41 (.49)	.30 (.46)	.47 (.50)	.16 (.37)	.21 (.41)
NO-YES	.42 (.50)	.38 (.49)	.20 (.40)	.33 (.47)	.18 (.39)	.20 (.40)
NO-NO	.35 (.48)	.52 (.50)	.21 (.41)	.48 (.50)	.27 (.45)	.21 (.41)

	T		Control	
	T1	T2	T1	T2
YES-YES	.73 (.45)	.79 (.41)	.76 (.43)	.84 (.37)
YES-NO	.23 (.42)	.18 (.39)	.16 (.36)	.25 (.43)
NO-YES	.25 (.43)	.26 (.44)	.18 (.39)	.26 (.44)
NO-NO	.24 (.43)	.28 (.45)	.20 (.40)	.28 (.45)

Figure 2

Proportion (0 to 1) of accuracy in the control conditions presented by grammaticality, intervention and time.



To determine which intervention contributed to the greatest increase in accuracy, we employed a generalized linear mixed-effects model with a binomial family using the lme4 (Bates et al., 2015) and multcomp (Hothorn et al., 2008) packages for post-hoc comparisons within the R environment (R Core Team, 2023). Initially, we fitted the fixed effects of condition (Yes-Yes, Yes-No, No-Yes, No-No), intervention group (Gr+I, Gr, I, T, or C), time (Time 1 vs Time 2), proficiency (a continuous variable), and all possible interactions. Random intercepts for participants and items were also included. Following model selection, the model with the best fit included only condition, intervention group, and time as fixed effects along with their interaction terms.

The main outcomes of the model revealed a significant main effect of condition ($\chi^2 = 5.21, p < .001$), time ($\chi^2 = 5.89, p = .015$), and two significant two-way interactions: condition*time ($\chi^2 = 9.46, p = .026$) and intervention*time ($\chi^2 = 14.21, p = .006$). Further exploration through relevant post-hoc comparisons (reported in Table 3 below) indicated significant differences within the condition*time interaction, particularly in the YES-YES and NO-NO conditions. Regarding the intervention*time interaction, significant differences were observed in the Gr+I group and the Gr group, suggesting that these two groups exhibited significant improvement between Time 1 and Time 2.

Table 3

Post-hoc comparisons from the main statistical model for the two significant interactions.

Condition*TIME	estimate	p value
YES-YES: T1 vs YES-YES: T2	= .83	< .001
YES-NO: T1 vs YES-NO: T2	= .32	= .381

NO-YES T1 vs. NO-YES T2	= .29	= .526
NO-NO: T1 vs. NO-NO: T2	= .49	= .016
Intervention*time		
Gr+Interaction strategy: T1 vs T2	= .672	= .011
Grammar T1 vs. T2	= .890	<.001
Interaction strategy T1 vs. T2	= .232	= .978
Task-only T1 vs. T2	= .095	= .999

4.2. Production results

The learners' oral production in the picture narrative task was analysed in terms of number of potential contexts for present continuous, number and percentage of the learners' attempts at producing present continuous and number and percentage of accurate instances of present continuous.⁵

It was particularly important to distinguish between attempts at production and accurate instances because focusing solely on the latter would imply prioritizing the end product over the process. By doing so, we would miss the opportunity to evaluate whether participants were progressing towards developing their use of the present continuous. Consequently, it would be unfair to only analyze accurate instances of the present continuous. This distinction is crucial because research has shown that the procedural stage of language production is essential for the eventual consolidation and automatization of knowledge (DeKeyser, 2007, 2015, 2017). Additionally, Macdonald (2013) demonstrated how language production, whether accurate or inaccurate, shapes language comprehension in general. Furthermore, in an artificial language learning paradigm, Hoppman and Macdonald (2018) showed that learners with extremely limited input benefit more from production-based practice. Their findings indicate that production-based training leads to quicker acquisition of an artificial language than comprehension-based training. This highlights the importance of capturing the extent to which learners in our study produce target forms, even if these forms are inaccurate. Incorrect instances of the present continuous included errors in the auxiliary verb, in the *-ing* morpheme, or both, and mixed L1-L2 instances of the present continuous.

- (1) Dog are playing.
- (2) Is open the door Sarah
- (3) Monkey està (is) jump the lamp.

The percentage of attempts was calculated based on learners trying to produce a present continuous structure in potential contexts, irrespective of whether the structure was target-like or not. As shown in Table 4 and Figure 3 below, all intervention groups attempted to produce present continuous forms around 50% of the potential contexts at Time 1. With the exception of the control group, all groups demonstrated some moderate descriptive improvement from Time 1 to Time 2 of testing.

Table 4

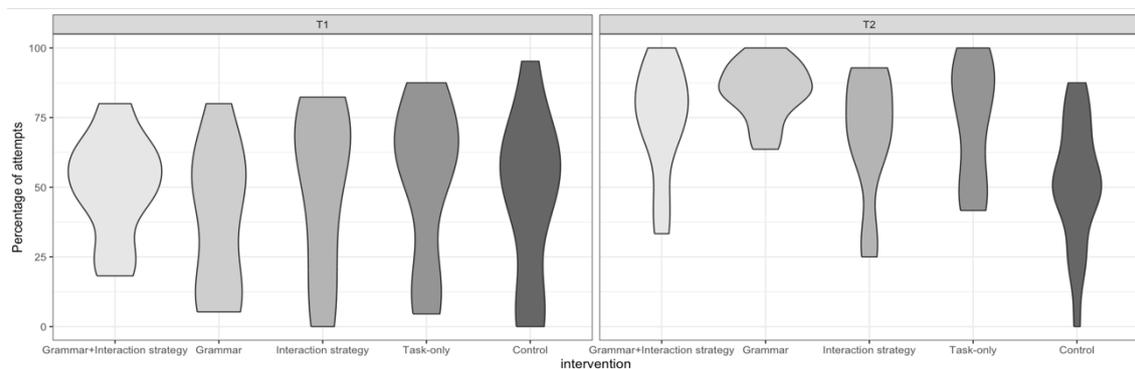
Descriptive statistics means (sd) of production attempts of present continuous presented by intervention group and time.

⁵ See Appendices 2 and 3 for raw numbers of learners' attempts, accurate attempts and potential contexts.

Production attempts		
	T1	T2
Gr+I	50.71 (19.20)	73.77 (20.60)
Grammar	38.53 (26.51)	84.58 (10.63)
Interaction	49.58 (30.08)	68.80 (22.57)
Task-only	52.94 (28.86)	73.63 (21.73)
Control	46.75 (29.48)	51.68 (22.61)

Figure 3

Violin plot showing the distribution of the responses (percentage of attempts) divided by group and time.



A two-way ANOVA was conducted to examine the effect of intervention and time on the production attempts of the present continuous. The analysis revealed a significant main effect of time ($F(1, 112) = 21.87, p < .001$) and a significant interaction of intervention*time ($F(4, 112) = 2.541, p = .047$). Subsequent Tukey's HSD Test for multiple comparisons indicated that only the Grammar group demonstrated significant improvement from Time 1 to Time 2 ($p < .001, 95\% \text{ CI} = [12.74, 79.36]$); all other comparisons were non-significant. We further explored the comparison between groups within Time 1 and Time 2 and results showed that at Time 1 there were no significant between-group differences and that at Time 2 the Gr+I, Gr and I group had significantly different percentages of attempts from the Control group (all $p < .001$) and that the T group was close to significant when compared to the Control group ($p = .061$). Regarding the comparison between the four intervention groups, no significant differences were found.

Regarding accuracy, the percentage of accurate attempts was calculated at both Time 1 and Time 2 with respect to the number of potential contexts. As shown in Table 5 and Figure 4 below, participants, regardless of intervention group, exhibited fairly low percentages of accuracy (around 10% across groups) at Time 1. At Time 2, all intervention groups displayed some moderate descriptive improvement, while the Control group showed no descriptive improvement.

Table 5

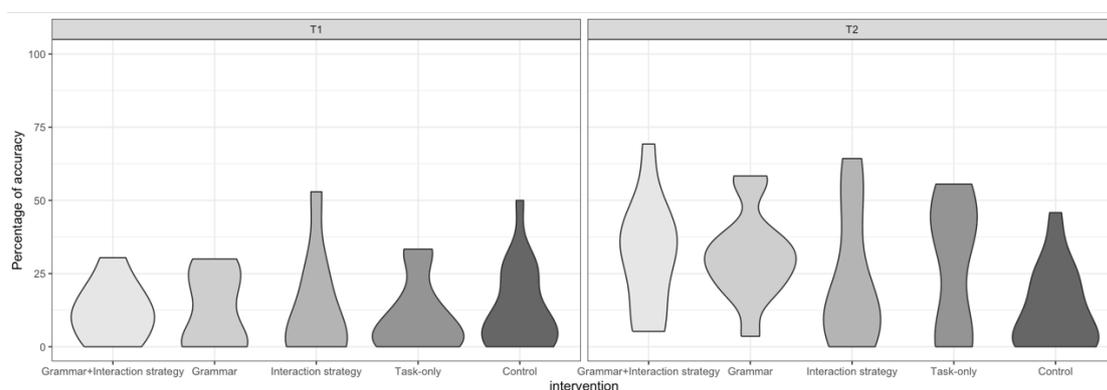
Descriptive statistics means (sd) of production accuracy of present continuous presented by intervention group and time.

Production accuracy		
	T1	T2
Gr+I	12.18 (10.01)	31.31 (19.57)

Grammar	12.40 (12.72)	31.18 (15.27)
Interaction strategy	12.76 (18.05)	22.99 (22.91)
Task-only	10.37 (12.29)	27.50 (22.66)
Control	12.78 (13.88)	12.73 (13.70)

Figure 4

Violin plot showing the distribution of the responses (percentage of attempts) divided by group and time.



We conducted another two-way ANOVA to compare the effects of intervention and time on the production accuracy of the present continuous. The results showed a significant main effect of time ($F(1, 112) = 14.26, p < .001$) and a significant interaction of time*intervention ($F(4, 112) = 4.28, p < .001$). We further explored the interaction with Tukey's HSD Test for multiple comparisons, which showed that only the Grammar ($p < .001, 95\% \text{ CI} = [1.89, 41.45]$) group had some significant improvement from Time 1 to Time 2. The Gr+I group showed a close to significant improvement ($p = .054, 95\% \text{ CI} = [3.92, 43.46]$). We explored further comparisons between intervention groups within Time 1 and Time 2, respectively. Comparisons showed that at Time 1, groups did not differ from one another and at Time 2, all four intervention groups significantly differed from the Control group ($p < .05$ for all comparisons). When the four intervention groups were compared, no significant differences arose.

5. Discussion

The study set out to explore whether an 8-week task-based peer interaction intervention and its different intervention types led to increased explicit knowledge and production of present continuous among 4th grade EFL learners and whether the use of pre-task instruction had an effect on such potential gains. The two research questions in the study will be explored together, first in relation to explicit knowledge and then in relation to production and use of the target feature.

Findings from the GJT indicate a high proportion of accurate judgements in the grammatical condition at both times and by all groups, possibly indicating some pre-study familiarity with the target structure. Although the present continuous is specifically dealt with in the 4th grade, students might have surely come across it during their previous years of EFL instruction. We

should also acknowledge that the structure follows the same pattern as their L1 equivalent, making it more accessible to the learners. This high proportion of accurate judgements cannot be attributed to a response bias towards grammaticality, as shown in our initial assessment of participants' sensitivity to grammaticality in both the control grammatical and ungrammatical conditions. In the case of the ungrammatical conditions, all groups had low proportions of accurate judgements at Time 1 and a different degree of improvement according to condition and type of intervention at Time 2. Accurate judgements of the ungrammatical NO-NO condition were significantly higher at Time 2 generally across groups but only the two groups where explicit pre-task grammar instruction was carried out (i.e. Gr+I and Gr) did show significantly more accurate judgements in all ungrammatical conditions. The Gr group devoted half of their weekly sessions to explicit learning of the target structure previous to focused task practice in the second half of the session. The Gr+I group devoted less time to explicit learning of the target structure but they did receive explicit instruction of the present continuous together with explicit instruction on interactional and corrective feedback strategies. These corrective feedback strategies that the researcher taught in the sessions were also focused on the target structure and therefore reinforced the explicit grammar instruction, which was then put to practice in task performance in the remaining of the session. The children in these two groups were provided with declarative knowledge every week which they could then put into practice through focused task procedural repetition over the 8-week intervention. This enhanced their knowledge of the target structure, which was reflected in the GJT results.

The other two treatment groups (i.e. I and T) patterned like the Control group, which indicates that participation in repeated focused task performance did not have any effects on the learners' explicit knowledge of the present continuous. Recall that the I group devoted the first half of the session to explicit instruction of interactional strategies (i.e. conversational adjustments) and corrective feedback strategies. In order not to interfere with the Gr+I group, the corrective feedback strategies in the I group included a variety of structures and lexical items and were not focused on the target structure in this case. The Task-only group devoted their first half of the session to vocabulary games which were unrelated to the target structure. The reason why this group was included in the study was to explore whether the mere presence of repeated focused task practice without any type of pre-task instruction led to gains in explicit knowledge and use of present continuous. The fact that these two experimental groups did not yield different results from the Control group, which did not carry out any focused task-based practice over the 8-week period, suggests that procedural repeated focused task practice alone might not lead to increased explicit knowledge of the target feature. The procedural repetition of focused tasks might be enough to trigger the use of the present continuous, as will be seen below, but clearly does not affect the children's explicit knowledge of the structure. This confirms previous research on the effectiveness of pre-task explicit instruction of a given grammatical structure (e.g. Khezrlou, 2021; Li et al., 2016; Li et al., 2018; Shintani, 2019) where groups that receive explicit grammar instruction previous to task performance exhibit the highest improvement. It is crucial to observe in previous research that only receiving grammar instruction has been shown not to lead to the same levels of explicit knowledge as instruction followed by task performance, as is shown in Shintani (2019), thus indicating that having the opportunity to practice and produce the target structure reinforces and consolidates the children's explicit learning, which further supports the importance of communicative practice for the development of declarative and proceduralized knowledge (DeKeyser, 2007; 2015; 2017).

Regarding the children's production of present continuous in the picture narrative task, no differences in the percentage of attempts to produce the target structure emerged between groups at Time 1. Unlike the Control group, all treatment groups increased their attempts at producing instances of present continuous, reaching percentages of around 70% and above at Time 2. However, only the group that received explicit grammar pre-task instruction (i.e. Gr) before procedural task repetition significantly improved their percentage of attempts. At Time 2, the treatment groups that received some kind (i.e. either grammar, interactional strategy or both) of

pre-task instruction (i.e. Gr, Gr+I and I) obtained significantly higher percentages of attempts than the Control group. As for the Task-only group, the percentage of attempts was close to significantly higher than the Control group at Time 2, all of which points not only to a clear effect of pre-task instruction but also to the benefits of focused repeated task practice on the children's attempts to use the target structure during task performance. The analysis of attempts (both accurate and inaccurate) allows us to capture the extent to which children try to produce the target structure and thus explore their learning process through production-based practice (Hoppman and Macdonald, 2018; Macdonald, 2013).

As for production accuracy, results were generally low at Time 1, where no differences between groups were found. All experimental groups showed moderate improvement from Time 1 to Time 2 while the Control group remained stable, and again, the group that received explicit grammar pre-task instruction (i.e. Gr) showed a significant increase in the number and percentage of accurate instances of present continuous. This time, the Gr+I group reached marginal significance in their improvement to Time 2. After the intervention, all treatment groups (i.e. Gr, Gr+I, I, T) produced a significantly higher number and percentage of accurate instances of present continuous than the Control group, thus corroborating the learning benefits of introducing focused task-based peer interaction on the use of the target structure during task performance. Previous research with older populations, such as Ellis et al (2019), found more frequent but not more accurate use of target structures as a result of pre-task instruction, although other similar studies did find higher accuracy in the use of the target structure in the groups receiving pre-task explicit grammar instruction, as in our study (Mochizuki & Ortega, 2008; Shintani, 2019; Van de Guchte et al., 2017).

6. Conclusion

This study suggests that the weekly introduction of focused task-based peer interaction in young learners' EFL sessions seems to enhance the frequency of use and accuracy of the present continuous, which is more noticeable if explicit pre-task grammar instruction is added before task performance. Focused task-based peer interaction alone does not appear sufficient to trigger improvement in explicit knowledge of the present continuous, where explicit pre-task instruction seems necessary for significant development. Our study contributes new data that reinforce the claim that task-based peer interaction can and should be introduced in the young EFL learners' classroom, particularly in contexts with limited exposure or opportunities for interaction in a natural environment. Transforming such task-based peer interaction sessions into task-supported ones by adding explicit pre-task instruction can provide learners with additional tools to foster their learning process. This explicit instruction supports young, low-proficiency learners when engaging in communicative practice and can lead to the consolidation of grammatical structures with ample production opportunities (DeKeyser, 2007; 2015).

A few limitations need to be acknowledged. A longer and more intensive intervention might have provided more robust results and potentially shown a stronger effect of repeated focused task communicative practice. Importantly, our results suggest that incorporating similar interventions targeting different structures regularly could lead to longer-term learning gains and potentially enhance overall task performance (i.e., complexity, accuracy, and fluency), which were not analyzed in this study. Although a delayed post-test was not feasible, it would have allowed us to assess the retention effects of such interventions. Further research is clearly needed to corroborate the benefits of task-based intervention studies in young learner EFL settings.

References

- Anwyl-Irvine, A.L., Massoné J., Flitton, A., Kirkham, N.Z., & Evershed, J.K. (2020). Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods*, 52(1), 388–407.
- Azkarai, A., & Imaz Agirre, A. (2016). Negotiation of meaning strategies in child EFL mainstream and CLIL settings. *TESOL Quarterly*, 50, 844–870.
- Azkarai, A., & Kopinska, M., (2020). Young EFL learners and collaborative writing: A study on patterns of interaction, engagement in LREs, and task motivation. *System*, 94(1).
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme 4. *Journal of Statistical Software*, 67, 1–48.
- Benati, A., & Schwieter, J. W. (2019). Pedagogical Interventions to L2 Grammar Instruction. In J. W. Schwieter & A. Benati (Eds.) *The Cambridge Handbook of Language Learning* (pp. 477-499). Cambridge University Press.
- Bridges, D., Pitiot, A., MacAskill, M., & Peirce, J. (2020). The Timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ*. 10.7717/peerj.9414
- Calzada, A., & García Mayo, M.P. (2020). Child EFL grammar learning through a collaborative task. In W. Suzuki and N. Storch (Eds.), *Language in language learning* (pp. 19–39). John Benjamins.
- Calzada, A., & García Mayo, M.P. (2021). Child learners' reflections about EFL grammar in a collaborative writing task: When form is not at odds with communication. *Language Awareness*, 30(1), 1–16.
- Collins, L., & Muñoz, C. (2016). The foreign language classroom: Current perspectives and future considerations. *The Modern Language Journal*, 100, 133–47.
- DeKeyser, R. M. (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge University Press.
- DeKeyser, R. M. (2015). Skill Acquisition Theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94-112). Routledge.
- DeKeyser, R. M. (2017). Knowledge and Skill in ISLA. In S. Loewen and M. Sato (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition* (pp. 15-32). Routledge.
- De la Fuente, M. (2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research*, 10(3), 263-295.
- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). Cambridge University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30, 474-509.
- Ellis, R. (2016). Focus on form: A critical review. *Language Teaching Research*, 20(3), 405–428.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28(3), 339–368.
- Ellis, R., Li, S., and Zhu, Y. (2019). The effects of pre-task explicit instruction on the performance of a focused task. *System*, 80, 38-47.
- Ellis, R., Skehan, P., Li, S., Shintani, N., and Lambert, C. (2020). *Task-based Language Teaching. Theory and Practice*. Cambridge University Press.
- Fujii, A., Ziegler, N., & Mackey, A. (2016). Peer interaction and metacognitive instruction in the EFL classroom. In M. Sato & S. Ballinger (Eds.), *Peer interaction and second language learning: Pedagogical potential and research agenda* (pp. 63–89). John Benjamins.
- García Mayo, M. P. (2021). *Child interaction in task-supported EFL/CLIL contexts*. In M. Ahmadian and M. H. Long (Eds.), *The Cambridge Handbook of Task-Based Language Teaching* (pp. 397-415). Cambridge University Press.
- García Mayo, M. P., & Imaz Agirre, A. (2017). Child EFL interaction; age, instructional setting and development. In J. Enever and E. Lindgren (Eds.), *Researching the complexity of early language learning in instructed contexts* (pp. 249–68). Multilingual Matters.

- García Mayo, M. P., & Lázaro Ibarrola, A. (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System*, 54, 40–54.
- Holmes, B. and Myles, F. (2019). White Paper: Primary Languages Policy in England – The Way Forward. RiPL: www.ripl.uk/policy/
- Hopman, E., & MacDonald, M. (2018). Production Practice During Language Learning Improves Comprehension. *Psychological Science* 29(6), 961-971.
- Hothorn T., Bretz F., Westfall P., & Heiberger R. M. (2008). Multcomp: simultaneous inference in general parametric models. R package version 1.0-0: R Foundation for Statistical Computing Vienna, Austria.
- Keating, G. D., & Jegerski, J. (2015). Experimental designs in sentence processing research: a methodological review and user's guide. *Studies in Second Language Acquisition*, 37, 1–32.
- Kersten, K., Piske, T., Rohde, A., Steinlen, A. K., Weitz, M., Couve de Murville, S., & Kurth, S. (2012). ELIAS grammar test II. *Studies on Multilingualism in Language Education*, 1.
- Khezrlou, S. (2021). Explicit instruction through task repetition: effects on explicit and implicit knowledge development. *Language Awareness*, 30(1), 62-84.
- Lázaro Ibarrola, A. & Azpilicueta Martínez, R. (2015). Investigating negotiation of meaning in EFL children with very low levels of proficiency. *International Journal of English Studies*, 15, 1–21.
- Li, S., Zhu, Y., & Ellis, R. (2016). The effects of the timing of corrective feedback on the acquisition of a new linguistic structures. *Modern Language Journal* 100(1), 276–295.
- Li, S., Ellis, R. & Zhu, Y. (2016). Task-Based Versus Task-Supported Language Instruction: An Experimental Study. *Annual Review of Applied Linguistics*, 36, 205-229.
- Li, S., Ellis, R., & Kim, J. (2018). The influence of pre-task grammar instruction on L2 learning: an experimental study. *Studies in English Education*, 23(4), 831-857.
- Loewen, S. (2015). *Introduction to instructed second language acquisition*. Routledge.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Wiley.
- Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning* (pp. 123–163). Multilingual Matters.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition* 26(3), 399–432.
- MacDonald, M. (2013) How language production shapes language form and comprehension. *Frontiers in Psychology* 4. DOI=10.3389/fpsyg.2013.00226
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk: Transcription format and programs (3rd ed.)*. Lawrence Erlbaum Associates Publishers.
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classroom: A study of guided planning and relativization. *Language Teaching Research*, 12(1), 11-37.
- Nassaji, H. (2017). Grammar Acquisition. In S. Loewen & M. Sato (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition* (pp. 205-223). Routledge.
- Nassaji, H. & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative contexts*. Routledge.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109-148.
- Pienemann, M. (2005). *Cross-linguistic aspects of processability theory*. John Benjamins.
- Pinter, A. (2007). Some benefits of peer-peer interaction: 10-year-old children practicing with a communication task. *Language Teaching Research*, 11, 189–207.
- Pladevall-Ballester, E. (2021). Pair dynamics and language-related episodes in child EFL task-based peer interaction. *Language Teaching for Young Learners*, 3(2), 189-213.
- Pladevall-Ballester, E. & Vraciu, A. (2020). EFL child peer interaction: Measuring the effect of time, proficiency pairing and language of interaction. *Studies in Second Language Learning and Teaching* (Special Issue), 10(3), 449-472.
- R Core Team. (2023). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. <https://www.R-project.org/>
- Roehr-Brackin, K. (2024). Explicit and implicit knowledge and learning of an additional language: a research agenda. *Language Teaching*, 57, 68-86.

- Sato, R. (2010). Reconsidering the effectiveness and suitability of PPP and TBLT in the Japanese EFL classroom. *JALT Journal*, 32(2), 189–200.
- Sato, M. (2020). Metacognitive instruction for collaborative interaction: The process and product of self-regulated learning in the Chilean EFL context. In C. Lambert & R. Oliver (Eds.) *Using tasks in second language teaching: Practice in diverse contexts* (pp. 215-236). Multilingual Matters.
- Sato, M. & Dussuel Lam, C. (2021). Metacognitive instruction with young learners: A case of willingness to communicate, L2 use, and metacognition of oral communication. *Language Teaching Research*, 25(6).
- Sato, M. & Loewen, S. (2018). Metacognitive instruction enhances the effectiveness of corrective feedback: Variable effects of feedback types and linguistic targets. *Language Learning*, 68(2), 507–545.
- Sato, M., & Lyster, R. (2012). Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization. *Studies in Second Language Acquisition*, 34, 591–626.
- Shintani, N. (2013). The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of L2 vocabulary by young beginning-level learners. *TESOL Quarterly*, 47, 36-62.
- Shintani, N. (2019). The roles of explicit instruction and guided practice in the proceduralization of a complex grammatical structure. In R.M. DeKeyser & G. Prieto Botana (Eds.), *Doing SLA Research with Implications for the Classroom: Reconciling Methodological Demands and Pedagogical Applicability* (pp. 83-106). John Benjamins.
- Van de Guchte, M., Rijlaarsdam, G., Braaksma, M., & Bimmel, R. (2017). Focus on language versus content in the pretask: Effects of guided peer-video model observations on task performance. *Language Teaching Research*, 23(3) 310-329.
- Willis, J. (1996). *A framework for task-based learning*. Longman.
- Willis, D. & Willis, J. (2007). *Doing Task-Based Teaching*. Oxford University Press.

Appendix 1

Picture narrative task: *Sarah's surprise*



Appendix 2. Average number of attempts and accurate attempts by group and time

	Attempts		Accurate attempts	
	T1	T2	T1	T2
Grammar+Interaction	5.44 (3.46)	13.61 (5.66)	1.44 (2.12)	5.41 (2.95)
Grammar	7.10 (5.80)	14.25 (5.54)	2.21 (2.39)	4.91 (2.67)
Interaction	7.22 (4.81)	7.36 (3.74)	2.22 (3.15)	3.54 (2.80)
Games	11.36 (7.61)	11.58 (6.05)	2.27 (2.64)	4.16 (4.11)
Control	9.68 (8.21)	8.78 (5.45)	2.36 (2.54)	2.41 (3.01)

Appendix 3. Average number of potential present continuous contexts by group and time

	T1	T2
Grammar+Interaction	11.44 (5.89)	18.61 (6.39)
Grammar	18.20 (7.14)	16.83 (6.31)
Interaction	14.88 (4.19)	10.45 (2.94)
Games	19.91 (6.07)	15.08 (4.77)
Control	19.00 (10.06)	16.21 (5.64)