

Questioning the Effectiveness of Procedural Repetition: The Case of Spanish EFL Primary School Learners L2 Behavior in College Language Learners

AMPARO LÁZARO-ÍBARROLA

IZASKUN VILLARREAL

Public University of Navarre

Received: 21 February 2018 / Accepted: 24 October 2018

ISSN: 1697-7467

ABSTRACT: It has been claimed that the repetition of the same task type with different content, known as procedural repetition (PR), seems to offer opportunities to focus on form and develop syntactic complexity. At the same time, PR is common and deemed appropriate for the language classroom. This quasi-experimental study explores the effects of PR on the overall production of 5 pairs of young Spanish EFL (age 11) learners in a school context with a beginner level of English. The pairs co-constructed and narrated three stories following the same procedure at three testing times over a three-week period. The analysis of the complexity, accuracy, and fluency of their narratives shows that, unlike previous studies, PR does not have any significant impact on the students' performance, except for lexical density, which increases in the third attempt. In light of these results, the validity of PR for child populations is discussed.

Keywords: Task-Based Language Teaching (TBLT), task repetition, English as a Foreign Language (EFL) children, storytelling tasks, Complexity, Accuracy and Fluency (CAF)

La efectividad de la repetición: El caso de alumnos españoles de inglés como lengua extranjera en Educación Primaria

RESUMEN: Varios estudios han mostrado que repetir la misma tarea con diferente contenido, lo que se denomina repetición procedimental (RP), ayuda a prestar atención a la forma y al desarrollo sintáctico. Asimismo, la RP es común y apropiada para el aula. Este experimento investiga los efectos de la RP en la producción de 5 parejas de niños españoles (11 años) aprendiendo inglés como lengua extranjera en un colegio y con un nivel de principiante. Las parejas construyeron y narraron tres historias, una cada semana, siguiendo un procedimiento similar. El análisis de la complejidad, precisión y fluidez de sus narraciones muestra que la RP no tiene un impacto significativo en las narraciones de los estudiantes, excepto en el caso de la diversidad léxica, que mejora en la tercera narración. A la luz de los resultados, se discuten las posibilidades de la RP con niños.

Palabras clave: Aprendizaje de lenguas basado en tareas (TBLT), repetición de tareas, niños, inglés como lengua extranjera (EFL), narración de historias, Complejidad, Corrección y Fluidez (CAF)

1. INTRODUCTION

Since the emergence of Task-Based Language Teaching (TBLT), numerous studies have demonstrated, with different outcomes and in different contexts, that repeating a given task, either identically (same task repetition, henceforth STR) or by modifying its content (procedural repetition, henceforth PR), boosts the language learning process by allowing learners to focus on the language in greater depth (Bygate, 2001; Kim & Tracy-Ventura, 2013; Mackey, Kanganas & Oliver, 2007; Shintani, 2012, 2014).

Among the two basic types of repetition: STR and PR, the latter type is the most frequently employed in real language lessons (Lázaro-Ibarrola & Hidalgo Gordo, 2017a). Suffice it to examine textbooks or examination papers where similar task types appear on every lesson with their corresponding contents. Therefore, it is pedagogically pertinent to further explore the possibilities of this type of repetition.

While PR in general has been claimed to offer benefits that can be observed in the main dimensions conforming overall performance (complexity, accuracy, and fluency (henceforth CAF) (Housen, Kuiken & Vedder, 2012)), the studies to date do not obtain similar results in terms of which aspects improve more; if there is a correlation and/or trade off effect among these three elements; or if the improvements show up and/or decline on the second or third attempt on the task. The reason for these inconclusive results partly rests in the fact that studies have used different contexts, participants, tasks and task procedures. Also, in some recent studies tasks requiring short sentences and short turns to be completed seem to have been an obstacle for complexity growth and even for fluency (Lázaro-Ibarrola & Hidalgo Gordo, 2017a; Sample & Michel, 2014). In any case, all previous research has something in common: the improvement of, at least, one of the components of CAF. This promising finding together with the fact that procedural repetition is a frequent practice in language classrooms (AUTHOR1 & COAUTHOR, 2017b) makes the subject worth of further exploration.

The present study, following a quasi-experimental design, aims to gain a deeper understanding of the effects of PR by analysing the production of 5 pairs of young learners (age 11) while telling three stories collaboratively following a similar procedure in each attempt. Stories are linguistically demanding and thus could constitute a valid genre to foster greater fluency and complexity upon PR. Also, including collaborative tasks is of great interest in the field as these tasks have long and repeatedly proved to be beneficial for the classroom and, in turn, their presence in language lessons and teaching materials has increased (García Mayo, 2007; Kim & Tracy-Ventura, 2013; Mackey, Kanganas & Oliver, 2007; Shintani, 2012, 2014). Finally, this study adds to the Second Language Acquisition (SLA) literature by exploring PR among young children, a population which remains underexplored in spite of constituting one of the largest (and fast growing) groups of English learners in the world (Lázaro-Ibarrola & Hidalgo Gordo, 2017b; Collins & Muñoz, 2016).

2. LITERATURE REVIEW

2.1. Task Repetition

The claim that repeating a given task has positive effects on the linguistic production of the learners can be tracked back to the cradle of Western civilization, as reflected in the Latin expression *repetitio est mater studiorum* (repetition is the mother of learning), which probably was and still is one of the main principles of learning in general. More recently, and within the ebullient field of SLA, Bygate (1996, 2001) confirms that repeating, in this case, repeating tasks, seems to bring about important and assorted benefits for L2 learners.

Although task repetition could be operationalized into several different types, two main types have been extensively dealt with in the literature (Kim & Tracy-Ventura, 2013). On the one hand, STR, which consists in the mere repetition of the exact same task and, on the other hand, PR, which implies the repetition of the same task type with different content. PR, the type we explore in this paper, seems to be a frequent pedagogical practice in language lessons (Lázaro-Ibarrola & Hidalgo Gordo, 2017a) where practitioners often repeat the same type of activity with different content as they progress through the course. This is also reflected in the layout of textbooks or even in the layout of examination papers where similar activities are presented while the content is modified to suit the learning or assessment needs. In sum, the bottom line in teachers' minds is that becoming familiar with the procedure is positive for successful task performance.

Empirically, the effects of PR have been analysed by several authors in different contexts and with different populations and tasks (see for instance the task used in Lynch and Maclean (2000) or the time span between tasks in Bret Blasco (2014); García Mayo & Imaz-Agirre (2016); and García Mayo, Imaz Agirre & Azkarai (2018)). Therefore, extrapolating findings from such differing studies is hardly possible. In what follows we will provide a critical review of the main studies to date which show some commonalities with our study and will, therefore, allow us to better understand the effects of PR. These commonalities will relate to the number of repetitions and the time span between these repetitions. Thus, following Bygate and Samuda (2005), we will consider studies that operationalize task repetition as involving task repetition (in our case with different content and same procedure) at intervals of around one or two weeks.

As happens with most of the topics in SLA, most of the research on task repetition has been carried out with adult populations. In a large study comprising 103 undergraduate students of Spanish at a university in the US, Gass, Mackey, Álvarez-Torres and Fernández-García (1999) analyzed the effects of same content, different content repetition and no repetition (control group). The task consisted in the repetition of a story after the participants had watched it. The same content group watched the same video three times while the different content group watched a different one each time (procedural repetition). In addition to this, a week after the experiment had finished, all the participants watched a new video and had to retell it. This final task was used to see if the gains obtained through the repetitions were applicable to a new story. Although timidly, the findings of this study hint at positive effects of repeating in learners' overall production, lexical sophistication and morphosyntax in all groups but, mainly, in the same content group. However, all the gains had vanished in the final video telling (which would be their fourth attempt on the task). These authors

suggest, in line with other studies (Plough & Gass, 1993; García-Mayo & Lázaro-Ibarrola, 2015), that repeating the task so many times can result in students becoming disinterested.

In a more recent study carried out in a different context but with a similar task, Patanasorn (2010) obtained different findings. This author compared PR, STR and content repetition in the production of 92 Thai English as a Foreign Language (EFL) learners (non-English majors) on film-retell tasks. She examined global accuracy, past simple accuracy and fluency and found important differences between the three groups. In the PR group accuracy in the use of the simple past improved; in the content repetition group fluency improved and accuracy seemed to get worse; finally, no changes were observed in the same-task repetition group. As for complexity, it was not analyzed.

Kim (2013a, b) elicited data from 48 Korean EFL learners during collaborative tasks divided into a STR group and a PR group. She found that learners' interest and focus on form (measured by their use of language related episodes) decreased when repeating the same task in comparison to learners who repeated different versions of the same task. Following the same line of research, this same author and Tracy-Ventura (2013) also compared STR and PR with 36 Korean EFL learners carrying out collaborative tasks. These consisted of an information-exchange task and the authors used a pre-test post-test design (before and after the three repetitions). In this case, they found that PR promoted syntactic complexity and that both groups improved in terms of task relevant linguistic features; however, they found no effects on fluency. They explain that, in the completion of collaborative tasks, learners produce many interactive features, such as negotiation strategies which, in turn, may be interpreted as lesser fluency although not necessarily as a sign of a worse performance (Kim & Tracy-Ventura, 2013).

2.2. Task Repetition in Young Learners' Collaborative Tasks

Studies focusing on the effects of task repetition in the case of child learners resolving tasks are scarce. Some authors have analysed the effects of repetition on young learners' use of interactional feedback (Lázaro-Ibarrola & Hidalgo Gordo, 2017a; García Mayo & Imaz Aguirre, 2016; Mackey et al., 2007) or on the functions and quantity of L1 use (Azkarai & García Mayo, 2017; Pinter, 2007b; Shintani, 2012), among other aspects. However, few studies have concentrated on the influence of TR on CAF and their results are far from conclusive. It is noteworthy to say that all these studies have used different types of spot-the-difference tasks which, as mentioned above, generate a type of language where fluency is hard to measure.

The main author looking at TR among children is Pinter (2005, 2007a, 2007b). In her studies she examined the effects of TR (operationalized as repetition of the same task with slight modifications of content) using a spot-the difference task and a direction-giving map task with 10-year-olds. Although she did not measure CAF as in most of the literature (e.g. Ellis, 2008), she reports that her learners became more accurate and fluent (she did not measure complexity).

In another study with children, Sample and Michel (2014) examine not only the three components of CAF separately but also look into a possible relationship between these three dimensions. These authors analyze the productions of six 9-year-old EFL learners performing a spot-the difference task and report that their learners increased their fluency significantly through task repetition. Also, the raw numbers of complexity suggest that students used fewer words but more complicated sentences over the course of the repetitions, although the trend

was not significant. Interestingly, they found the largest and most significant improvements in the second repetition. They discuss their results by making reference to the possibility that, in the third repetition, where gains diminished in reference to the second one, students had learned to complete the task using shorter and more direct utterances, without lengthy descriptions, and also by repeating some of what their partner had said, all of which did not encourage greater complexity.

Quite different were the results obtained Lázaro-Ibarrola and Hidalgo Gordo (2017a). They analyse the interactions of ten pairs of English learners (age 11) who had to repeat a spot-the-difference task (in the form of picture placement) three times. Their analysis of learners' general performance (complexity, accuracy and fluency) indicates that accuracy timidly improved while complexity and fluency remained unaffected. Contrary to Sample and Michel (2014), this happened in the third, and not in the second, repetition. These authors suggest that three repetitions might be necessary for effects to show up and also interpret their results for accuracy as encouraging. In a similar vein with Kim and Tracy-Ventura (2013), they also consider the possibility that complexity and fluency have not improved by means of familiarity with the procedure due to the characteristics of their task. In their task, short sentences and continuous turn taking were the natural way to resolve the activity and, therefore, greater complexity or greater fluency were not necessary to complete the task.

In spite of their different results, all the studies including PR with children resolving spot-the-difference tasks identify some positive effects and suggest that the type of task might be the reason to explain the lack of greater effects in complexity and fluency. Taking the baton from these researchers, the present study focuses on the PR of a task with greater linguistic demands than picture-placement tasks: story-telling. That is, a task more similar to the type of task used in adult studies.

3. METHOD

3.1. Research Questions

The present study sets off to explore the effects of the PR of a story-telling task (three times over a three week period) on the overall performance (CAF) of five pairs of 11-year-old Spanish learners of English in school. Thus, our study includes the following research questions:

- Does the procedural repetition of a story-telling task (three times) have an impact on the linguistic production of our learners?
- If any, what are the affected aspects (complexity, accuracy and/or fluency)?
- Are the effects more relevant on the second or third repetition?

In line with previous studies, we expect to find improvements for accuracy (Kim, 2013a, b; Lázaro-Ibarrola & Hidalgo Gordo, 2017a; Mackey, Kanganas, & Oliver, 2007; Patanasorn, 2010). Also, given that we are using a story-telling task we also contemplate the possibility of a positive impact on complexity and fluency (Gass et al., 1999). As for the number of repetitions, probably the level and task type will allow our learners to show greater improvements in the third repetition (Lázaro-Ibarrola & Hidalgo Gordo, 2017a).

3.2. Participants

The present study analyses the effects of PR on the overall oral performance (complexity, accuracy and fluency) of a group of 10 young EFL learners. The participants were 10-11-year-old learners of EFL in a school context in Spain. As little is known about the effects of task repetition among young populations, targeting young students will allow us to better understand the acquisition processes of this under researched, yet large and growing, population. On the other hand, the 10-11 year-olds were chosen because, according to school teachers, younger students wouldn't be able to carry out the task successfully in English.

The school was a small state school located in a working class area where students had little or no opportunities to reinforce English outside school. One of the school's objectives was to introduce different task types in their regular lessons in order to improve the level of English of the students and, therefore, this school was very suitable to carry out this investigation. The methodology used in the English classroom was mainly communicative and basic interaction in English was common. Classroom tasks including story-telling tasks as well as other tasks that required students' interactions in pairs or small groups were frequent in their lessons. However, they were not familiar with the specific task used in this study where students have to co-construct and narrate a story together.

At the time of data collection all learners had been learning English in the same school having 5 lessons of English per week every year. According to school internal tests, their level could be equated to an A1 level of the Common European Framework of Reference for Languages (Council of Europe, 2001). This level is lower than the level of the students in previous research on PR. Therefore, it will be interesting to see whether similar effects will be found in our participants and this will be relevant to the many groups of young students who still have a low command of the foreign language.

All in all, as explained in the theoretical review, previous studies have demonstrated the validity of procedural repetition with older students and with young students at higher levels of proficiency. Thus, with the present study we would like to test this validity with younger students who have lower levels of proficiency.

3.3. Procedure

The students worked in pairs through the whole process of data collection. The pairs were made with two criteria in mind: similar level of proficiency and fluent relationship with their partner. It was their own teacher together with the researcher who made the pairings.

The 5 pairs had to co-construct and narrate three different stories based on a set of pictures (6 cards, 1 with the title and 5 with pictures) following the exact same procedure. Modelled on previous studies and to minimize the effects of ongoing instruction between repetitions, the time span between narrations was one week. The specific narration of every story went as follows:

- Step 1. The two students in the pair sit on a separate desk with a screen between them so that they can hear one another but cannot see one another. Every student has half of the cards telling the story (3 cards each) on their desk. They are given some time to look at the cards.

- Step 2. The students are given 5 minutes to speak to each other describing what they see in their pictures and trying to work out what is happening in the story.
- Step 3. After the 5 minutes or after having been able to co-build the story, they sit together at one desk, see all the pictures, reconstruct the story and narrate it together.

The oral production throughout the three steps is the object of analysis in the present study. The stories selected belong to different collections of “Oxford Reading Tree” (Oxford University Press) and their titles are presented in Table 1.

Table 1. Selected stories from Oxford Reading Tree.

	Day 1	Day 2	Day 3
Year 6	Spots (Hunt & Brychta, 2003)	By the stream (Hunt & Brychta, 2008)	Super Dad (Hunt & Brychta, 2011)

The researchers made an effort to select stories of similar difficulty both in terms of plot complexity and also of the linguistic elements required to tell them. In order to guarantee this, the school teachers helped to choose the stories and made sure that the type of vocabulary and structures was adequate for their students. Also, the fact that the stories in “Oxford Reading Tree” are classified according to age somehow ensures that the plot complexity was similar and adequate for the age ranges in the present study.

3.4. Data codification measures

The oral production of the children was video and audio recorded and was transcribed in CHILDES (McWhinney, 2000). Following previous studies (Ellis, 2005, 2008), among the many possible measures for CAF the following were considered to be the most appropriate for our data.

In order to measure complexity, syntactic and lexical complexity were included (Norris & Ortega, 2009). Syntactic complexity was analyzed in terms of the number of words per AS-unit and of the number of clauses per AS-unit (complexity by subordination). AS-units have been defined as ‘a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause associated with either’ (Foster, Tonkyn & Wigglesworth, 2000: 365). Here, minor utterances (“yes”, “ok”, “no”, etc.) were excluded from our account.

For lexical diversity, D was considered (Malvern & Richards, 2000; Malvern, Richards, Chipere, & Purán, 2004).

The adopted accuracy measures included the proportion of error free clauses as well as the proportion of errors per clause (Yuan & Ellis, 2003). Here all error types were considered: syntactic, morphological and lexical. As for fluency, speech rate was measured in terms of syllables per minute (Kim & Tracy Ventura, 2013).

To ensure coding reliability, 30% of the data were independently coded by the two researchers. The resulting coded transcripts were compared and all the differences in coding were discussed until a total intercoder agreement was reached. Finally, the remaining data were distributed and coded. The two coders were in constant communication and when any of them had any uncertainty about specific codings, those instances were discussed until agreement was reached.

As for the statistical analysis, non-parametric tests were used, given that a Komolgorov-Smirnov test revealed that the data were not normally distributed. Friedman tests were used to examine whether the distribution of the variables varied across tasks. When differences were obtained, a Wilcoxon signed-rank test for related measures was used to see where the distributions varied exactly.

4. RESULTS AND DISCUSSION

4.1. Results

Starting with accuracy, Table 2 presents the results obtained, including the percentages of error free clauses and the percentage related to number of errors per clause. The table features the results for the three tasks.

Table 2. Accuracy: Error free clauses and percentage of errors by task.

	Task 1	Task 2	Task 3
Number of Errors	148	177	192
Clauses	351	398	418
Error free clauses	202	239	259
% Error free clauses	57.55%	60.05%	61.96%
% Errors per clause	0.42%	0.44%	0.46%

When looking at the raw numbers in Table 2, there seems to be a timid but steady trend towards a higher percentage of error free clauses from task 1 to task 2 to task 3; yet, the opposite trend applies when looking at the total number of errors and the number of errors per clause, which slightly increases at each testing time suggesting no improvements in correction. That is, there are more correct clauses but, when looking at clauses that contain errors, the number of errors per clause is higher. This contradictory result seems to indicate that there is no clear effect for accuracy. In fact, the most relevant result is that none of these aspects, error free clauses, total number of errors, or number of errors per clause, presents important variations and therefore what can be said is that accuracy remains rather stable across tasks. This stability is confirmed by the statistical analysis which yielded no significant differences for any of these aspects across task attempts.

Next, Table 3 presents the results for fluency in each repetition.

Table 3. Fluency: Syllables per minute.

	Task 1	Task 2	Task 3
Syllables	832	999	1186
Time in minutes	9.19	8.28	11.81
Syllables per minute	94.90	124.67	101.76

As we can see in Table 3, when examining the number of syllables per minute, there is a remarkable increase from task 1 to task 2 but then, a clear drop from task 2 to task 3. However, and as claimed for accuracy, the most relevant finding is the stability across tasks with no statistical differences in any of the measures or task attempts.

Next, in Table 4, we present the overall results for complexity measures.

Table 4. Complexity: Words, Clauses, AS-units, and D by task

	Task 1	Task 2	Task 3
Words	2348	2723	2797
Clauses	351	398	418
AS-units	115	134	150
Words per AS-unit	20.42	20,32	18,65
Clauses per AS-unit	3.05	2,97	2.79
Lexical diversity D	22.96	21.81	26.93

As Table 4 features, if we look at amount of production, the number of words and clauses sharply increases from task 1 to task 2 and remains stable from task 2 to task 3. As for the complexity entailed in this amount of linguistic items, however, words and clauses per AS-unit remain unaffected across tasks, as confirmed by the lack of statistical significance for any of these aspects. Finally, the lexical diversity measure shows identical results in task 1 and task 2 but an increase in the last task (task 3), which is statistically significant. In fact, this is the only statistically significant result obtained in our analysis (Friedman test: $\chi^2= 8.4$; $df= 2$, $p= 0.015$; Wilcoxon signed-rank test results: Story 1 vs 2: $z=-1.483$; $p= 0.138$; Story 2 vs 3: $z= -2.023$; $p= 0.043$; Story 1 vs 3 $z= -2.023$; $p= 0.043$).

4.2. Discussion

This quasi-experimental study aimed at exploring the effects of PR on the overall performance (CAF) of 5 pairs of young EFL learners with a low level of proficiency. Unlike previous studies with child populations, in which spot-the-difference tasks have been typically

included, the task included in our study was a story-telling activity with the aim of generating opportunities for a type of discourse encompassing greater complexity and fluency.

Contrary to our predictions, from all the aspects under scrutiny, only lexical diversity in the third repetition has shown an increase that is supported by statistical significance. This change might be the only effect of PR in our data, however, our results need to be taken with caution for two main reasons. First and foremost because we do not have a control group and, therefore, can only present this growth in lexical density as a probable effect of PR. Second, because variety of vocabulary, more than any of the other aspects analyzed in this study, can depend on the content of the story itself, that is, it might be the case that the different content of the third story required a greater variety of words for its narration.

On the other hand, the trends hinted at by the raw numbers, albeit not statistically significant, should be discussed.

In the case of error free sentences their number grows in the second attempt and is further augmented in the third one, which would go in line with our predictions; however, and contrary to our expectations, this is not accompanied by a lower number of errors, leaving us with a contradictory scenario for accuracy which leads us to highlight the importance of not using a single measure for this component.

In the case of fluency, the number of syllables per minute increases in the second attempt but declines again in the third one, which, again, suggests that this fluctuation might hinge on variables that are beyond our control.

In the case of complexity, learners are able to produce more words in each repetition but, on the contrary, they do not generate more complex sentences. Perhaps this increase in amount of production, that is, the fact that they are able to speak more overall, is a first step into greater complexity.

In spite of these possibilities, the fact that a significant growth has only been located in the lexical component and in the third repetition, together with the small and contradictory signs of change observed in the raw data, lead us to conclude that there is no clear impact from PR and, if any, it would be restricted to the lexis, which, in turn is the aspect that can more strongly depend on a particular story.

This lack of clear effects of PR seems to contradict previous research and thus opens new avenues to reflect on what PR really offers or on how its effects can be measured. Nevertheless, it is important to note that it does not differ much from studies with child populations which have only reported small effects. For instance, in Lázaro-Ibarrola and Hidalgo Gordo (2017a), when analyzing CAF among young EFL learners on a spot-the-difference task, only timid gains in accuracy in the third repetition (significant in the number of errors per clause but not in the number of error-free sentences) were reported. As for the study carried out by Sample and Michel (2014), their six 9 year old EFL learners performing a spot-the-difference task only increased significantly in fluency in the second repetition.

All this leads us to question the effects of PR on CAF among children and calls for more research with this population including larger numbers of students and a wider variety of measures.

In the case of our participants, we can also discuss the possibility that their low proficiency level might have been a key factor. Their level is lower than the level reported in all previous studies and, perhaps, this is why only gains in vocabulary have appeared. In fact, studies have demonstrated how low level students performing communicative tasks

place most of the focus on the lexis (Leeser, 2004). In line with Bygate (2001), we would like to suggest that, with low-proficient students, such as the ones in the present study, mass repetition over a short time period might be necessary to obtain greater effects or, at least, to further explore and understand the possibilities of task repetition.

On the other hand, the fact that error free sentences and total number of errors present contradictory percentages calls into question the validity of using only one measure for a particular component (in this case for accuracy), as this could lead researchers to jump to conclusions on incomplete evidence.

5. CONCLUSIONS

This study has explored the effects of PR on the overall performance (CAF) of 5 pairs of young EFL learners with a low level of proficiency while performing a story-telling task. However, our results seem to indicate no clear effect of PR on the complexity, accuracy and fluency of our learners' production of oral narratives. Therefore, it would be interesting to explore other possible aspects that might have been affected by the repetition. In this sense, it would be interesting to assess the effects of PR not only on CAF but also, from a more holistic perspective, including L1, interaction strategies, turn-taking, pragmatics, task-tactics, capacity to resolve the task, etc. In fact, in our analysis the shortest amount of clauses, the shortest time on task and the lowest number of words is found in task 1. Perhaps the PR is allowing these students to produce more English although they are not able to do so more fluently, correctly or with greater sophistication. Of course, research on the effects of collaborative vs. individuals and comparing STR vs. PR is essential to further substantiate these findings and to identify the most relevant variables. Thus, to end with a positive reflection, this more abundant production might be an advantage in itself, PR might have had a releasing effect on these low proficient students opening up their ability to use English to communicate, in this case, to narrate a story with their partner.

It is, of course, important to acknowledge the limitations of the present study. The number of students is small and there is no comparison group performing a different type of task repetition. Also, the three stories might have presented different levels of difficulty to the students, in particular, regarding the vocabulary that they needed to use. Therefore our results need to be taken with caution and further research should be carried out to confirm or contradict our findings.

Notwithstanding the limitations, we would like to say that our lack of effects calls into question the frequently demonstrated validity of task repetition for CAF with certain populations. This is relevant for both teachers and SLA researchers. First of all, it is relevant for language teachers, who should not expect that the PR of stories, with young and low level students, will necessarily result in greater accuracy, complexity or fluency. However, as suggested above, teachers should also be aware that PR might be useful to help students speak more on each attempt. Finally, our findings are relevant for SLA researchers. On the one hand, the fact that our results do not go in line with previous research evinces, once more, the complexity of SLA as a field of research due to the conjoint effect of the always

numerous variables involved in a given experiment. On the other hand, our results encourage further studies which could explore the benefits of PR beyond CAF, by including a more thorough analysis which considers elements such as quantity of production, use of interaction strategies and negotiations of meaning, use of the L1 and ability to complete the task.

6. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the funding from research grant FFI2012-32212 (Spanish Ministry of Economy and Competitiveness). Special thanks go to the school José María de Huarte in Pamplona for granting permission to conduct the study and to the school teachers for donating their precious class time. Our most heartfelt gratitude to the students for their enthusiasm and help. Last but not least, thanks to Maite and Euri for being there 24/7.

7. REFERENCES

- Azkarai, A. & García Mayo, M. P. (2017). Task repetition effects on L1 use in EFL child task-based interaction. *Language Teaching Research*, 21(4), 480-495.
- Bret Blasco, A. (2014). *L2 English young learners' oral production skills in instructed second language contexts: A longitudinal study of CLIL and EFL settings* (Unpublished doctoral dissertation). Universitat Autònoma de Barcelona, Barcelona, Spain.
- Bygate, M. (1996). Effects of task repetition: appraising the developing language of learners, in J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136-146). Oxford: Macmillan.
- Bygate, M. (2001). Effects of task repetition on the structure and control of language, in M. Bygate, P. Skehan, & M. Swain (Eds.), *Task-based learning: Language teaching, learning, and assessment* (pp. 23-48). London: Longman.
- Bygate, M. & Samuda, V. (2005). Integrative planning through the use of task repetition, in R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 23-48). Amsterdam/Philadelphia: John Benjamins.
- Collins, L. & Muñoz, C. (2016). The Foreign Language Classroom: Current Perspectives and Future Considerations. *The Modern Language Journal*, 100, 133-147.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge, Cambridge University Press, available from: http://www.coe.int/t/dg4/linguistic/source/Framework_EN.pdf, accessed 16 November, 2017.
- Ellis, R. (2005). Planning and task-based performance: Theory and research, in R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3-36). Amsterdam: John Benjamins.
- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, 21 (3), 354-375.
- García Mayo, M. P. (Ed.). (2007). *Investigating tasks in formal second language learning*. Clevedon: Multilingual Matters.
- García Mayo, M. P., & Imaz Agirre, A. (2016). Task repetition and its impact on EFL children's negotiation of meaning strategies and pair dynamics: an exploratory study. *The Language Learning Journal*, 44 (4), 451-466.
- García Mayo, M.P., Imaz Agirre, A., & Azkarai, A. (2018). Task repetition effects on CAF in EFL child task-based oral interaction, in M. Ahmadian & M.P. García Mayo (Eds.), *Recent*

- Perspectives on Task-based Language Learning and Teaching* (pp.9-28). Boston/Berlin: Mouton de Gruyter.
- García Mayo, MP and Lázaro Ibarrola, A (2015). Do children negotiate for meaning in task-based interaction? Evidence from CLIL and EFL settings. *System*, 54, 40-54.
- Gass, S. M., Mackey, A., Álvarez-Torres, M. & Fernández-García, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49, 549-580.
- Housen, A, Kuiken, F. & Vedder, I. (Eds.). (2012). *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA*. Amsterdam/Philadelphia: John Benjamins
- Hunt, R. & Brychta, A. (2003). *Spots*. Oxford University Press.
- Hunt, R. & Brychta, A. (2008). *By the Stream*. Oxford University Press.
- Hunt, R. & Brychta, A. (2011). *Super Dad*. Oxford University Press.
- Kim, Y. (2013a). Promoting attention to form through task repetition in a Korean EFL context, in K. McDonough and A. Mackey. (Eds.), *Second language interaction in diverse educational settings* (pp. 3-24). Amsterdam: John Benjamins.
- Kim, Y. (2013b). Effects of pretask modelling on attention to form and question development. *TESOL Quarterly*, 47, 8-35.
- Kim, Y. & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction?. *System*, 41, 829-840.
- Lázaro-Ibarrola, A. & Hidalgo, M. A. (2017a). Procedural repetition in task-based interaction among young EFL learners: Does it make a difference? *International Journal of Applied Linguistics*, 168(2), 183-202. doi: 10.1075/itl.16024.laz
- Lázaro-Ibarrola, A. & Hidalgo, M. A. (2017b). Benefits and limitations of conversational interactions among young learners of English in a CLIL context. In M. P. García-Mayo (Ed.), *Learning Foreign Languages in Primary Schools: Research Insights* (pp. 86-102). Bristol: Multilingual Matters.
- Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8, 55-81.
- Lynch, T. & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4, 221-250.
- Mackey, A., Kanganas, A.P. & Oliver, R. (2007). Task familiarity and interactional feedback in child ESL classrooms. *TESOL Quarterly*, 41, 285-312.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Malvern, D.D. & Richards, B.J. (2000). Validation of a new measure of lexical diversity, in M. Beers, B., van den Bogaerde, G., Bol, J. de Jong & C. Rooijmans (Eds.), *From Sound to Sentence: Studies on First Language Acquisition* (pp.81-96). Amsterdam: John Benjamin.
- Malvern, D., Richards, B., Chipere, N., & Purán, P. (2004). *Lexical diversity and language development*. New York: Palgrave Macmillan.
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Applied Linguistics*, 30, 555-578.
- Patanasorn, C. (2010). *Effects of procedural content and task repetition on accuracy and fluency in an EFL context* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff, USA.
- Pinter, A. (2005). Task repetition with 10-year-old children, in C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 113-126). London: Palgrave Macmillan.
- Pinter, A. (2007a). Some benefits of peer-peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research*, 11(2), 189-207.

- Pinter, A. (2007b). What children say: Benefits of task repetition, in K. Van den Branden, K. Van Gorp, & M. Verhelst (Eds.), *Tasks in action: Task-based language education from a classroom-based perspective* (pp. 131–158). Newcastle, UK: Cambridge Scholars Publishing.
- Plough, I., & S. Gass. (1993). Interlocutor and task familiarity: effects on interactional structure, in G. Crookes & S. Gass (Eds.), *Tasks and language learning: integrating theory and practice* (pp. 35-56). Philadelphia: Multilingual Matters.
- Sample, E., & Michel, M. (2014). An exploratory study into trade-off effects of complexity, accuracy and fluency in young learners' oral task repetition. *TESL Canada Journal*, 31 (Special Issue 8), 23–46.
- Shintani N. (2012). Repeating input-based tasks with young beginner learners. *RELC Journal*, 43, 39–51.
- Shintani N. (2014). Using tasks with young beginner learners: The role of the teacher. *Innovation in Language Learning and Teaching*, 8, 279–294.
- Yuan, F., & R. Ellis. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 24, 1-27.