EXPLORING THE USE OF PRE-TASK PLANNING STRATEGIES IN EFL COMMUNICATIVE TASKS IN HIGHER EDUCATION: THE EFFECT ON CAF MEASURES

AITOR GARCÉS-MANZANERA
Universidad de Murcia

1. INTRODUCTION

Planning has been explored from several perspectives, i.e. from oral production to written production. The study of how planning might affect L2 learners' oral performance during task has been of interest in recent decades (for an extensive review, see R. Ellis, 2009). The relevance of planning is highly related to learner-internal variables given the importance of language aptitude, working memory. Additionally, learner-external variables such as the type of task and task complexity have also proved to be pivotal in determining how L2 learners handle a task. Of special relevance for the present study is the influence of strategic pre-task planning on the measures of complexity, accuracy, and fluency (henceforth, CAF measures). Most of the research on pre-task planning has concentrated on observing what effect planning procedures had on these measures. Whilst most empirical studies have delved into the different effects between no planning and pre-task planning, a few others (e.g. Skehan & Foster, 2005; Mochizuki & Ortega, 2008) have manipulated strategic pre-task planning as guided (structured) and unguided (free) planning.

Regarding the type of task, recent empirical endeavors and theorizings (see Garcés-Manzanera, 2021a & 2021b) have manifested the importance of a specific type of task. In these studies, the role of

communicative tasks within the context of Education in Higher Education is underscored.

Hence, the present study intends to contribute to the substantial body of work in pre-task planning in L2 oral production and TBLT. Our study is sustained by three different axes: (i) the manipulation of strategic planning as guided or unguided planning in an attempt to observe (ii) the effect of these pre-task planning procedures on the preparation and task outcomes phases of an (iii) education-oriented communicative tasks which involve the use of vocabulary related to the context of participants.

2. PLANNING IN L2 ORAL PRODUCTION

Crookes (1989) stated that making plans was a day-to-day activity, which is part of our life as a mechanism to anticipate future activities, connected with consciousness. The role of planning is pivotal in the process of language production, either oral or written (Faerch & Kasper, 1983; Crookes, 1989). The importance of the planning process emerged out of its presence in the components within the models of speech production in L1, but it was not until the 1980s when researchers started to explore how these models applied and differed in the L2. Sangarun (2001) points to three different lenses through which planning has been looked into in L2 research. The first has focused on the conceptualization of planning as an indicator of phenomena while using the L2. In this case, L2 speakers use planning to confront issues at speech production and to enhance their L2. Regarding the second strand, as Sangarun (2001) himself stresses out, L2 researchers were primarily concerned with the conceptualization of planning as functional given that L2 learners make conscious use of this process in an attempt to attain the goals propounded in L2 tasks (see O'Malley *et al.*, 1985). Finally, the third line of inquiry has been focused on pre-task planning (this concept will be detailed in ensuing sections) as a *condition* of enhancing L2 production but also advancing L2 development. This development is operationalized through the measures of complexity, accuracy, and fluency. Before moving on to justifying why pre-task

planning is highly regarded in L2 research, an overview of Levelt's (1989, 1999) model of L1 speech production will be provided. A relationship with L2 speech production will be established.

LEVELT'S MODEL OF L1 SPEECH PRODUCTION

The modular model of L1 Speech Production by Levelt (1989, 1999) has been widely used as a basis for L1 and L2 oral production research. Throughout the decades, it has been adapted to the specificities of the L2 research dynamics (e.g. Kormos, 2006).

Levelt's model propounds that there are three main stages in language production. As illustrated in figure 1 on the next page, there are several components in the model. The first step is the *conceptualizer*, where the speaker generates the message and organizes the information obtained from his knowledge or external sources into a given structure. This stage entails the process of planning. The second step is the *formulator*, where the information structured and organized in the conceptualizer is further encoded in the L2. The next step is the *articulator*, in which the information encoded linguistically is further buffered phonologically. Finally, the presence of *monitoring* may be present at all the stages to ensure not only accuracy, but also rightful comprehension.

FIGURE 1. Levelt's (1989) model of speech production.
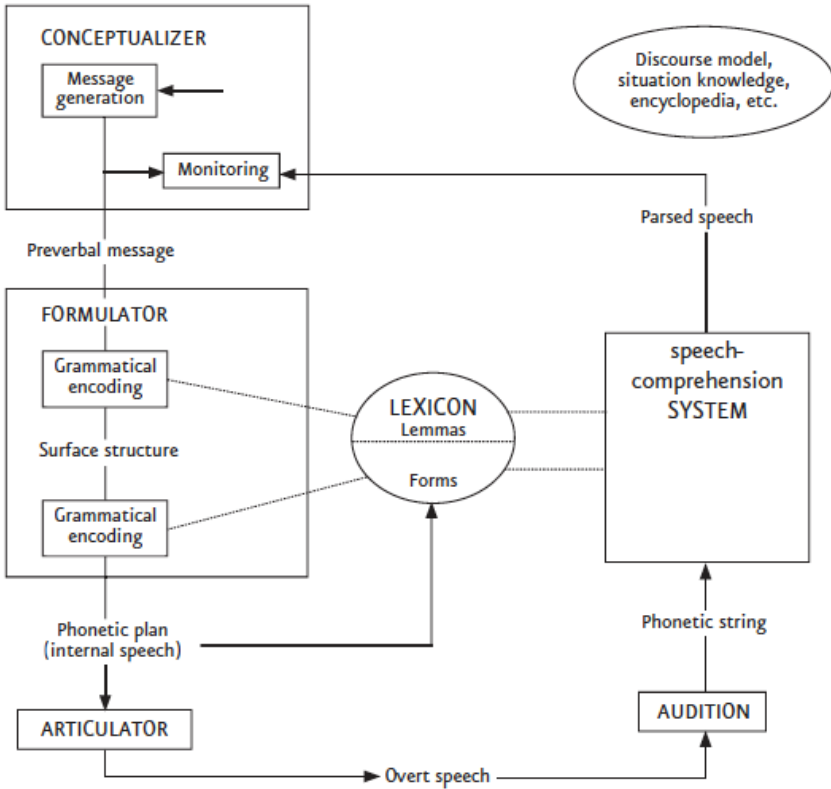
In the context of L2 oral production, the concept of *parallel processing* emerged as the ability to perform two steps of the model of speech production simultaneously. Nonetheless, this dual processing may be only possible in the L1, and in the case of L2 proficient speakers (Kormos, 2006; Skehan, 2014). This implies that L2 speakers are able to conceptualize at the same time they are producing a meaningful utterance. Low proficient L2 speakers are less likely to engage in this dual parallel process as a result of the absence of automaticity in the activation and development of those processes. As Lambert *et al.* (2020) point out, low proficient L2 speakers may resort to serial processing, in which each of the steps has to occur successively. This is the reason why low-proficient learners tend to be more disfluent and hesitant.

All this considered, the planning process has been considered from diverse perspectives depending on how it is developed. In fact, one way to overcome and enhance L2 oral production, and in turn, knowledge is by providing learners with effective preparation. Thus, this preparation, which might be understood as planning or preconceptualization, will allow less proficient learners to produce more accurate, but also more fluent content. Likewise, this might help reinforce the gradual progression from serial processing toward parallel processing (see Lambert & Robinson, 2014; Lambert *et al.*, 2020). A preconceptualization of the content to be orally produced is also thought to promote automatic processing of both content and language, leading to more fluency (see Wendel, 1997; as cited in Sangarun, 2001).

In what follows, the different dimensions of planning will be explored, with a special focus on pre-task preparation procedures.

3. TYPES OF PLANNING

The categorization of planning for L2 oral production has been proposed in different ways until the more synthetic categories proposed by Ellis (2005). A series of seminal works (see MacWhinney & Osser, 1977; Butterworth, 1980; Levelt, 1989; Wendel, 1997) identified several categories of planning: macro and micro planning; pre-planning and co-planning, and finally, online planning and offline planning. Despite the varied terminology used by the authors, all of them point to very similar operationalizations. The nearest definitions to the further L2 oral production planning patterns are pre-planning and offline planning, and co-planning and online planning. In the case of the former, planning occurs at a time where communication is not taking place, i.e. pre-task preparation. The latter refers to planning occurring at the time of the communication, i.e. when the speaker is producing utterances, parallel processing is activated. Butterworth's (1980) definitions regarding macro and microplanning mainly refer to the type of element being planned, that is, whether the speaker is planning content or linguistic structures.

Ellis's (2005) works provided the basis for what would be the main categories used in L2 oral production research, and more specifically, the planning process. The first basic distinction is established between *pre-task planning* and *within-task planning*. The first type of planning involves that learners purposefully prepare the task before performing it. The second type entails that learners are planning while performing the task. Within-task planning can be pressured, that is, with a time constraint to perform the task, or unpressured, in which case no time constraints are set. A further subdivision is proposed in the case of pre-task planning: rehearsal and strategic planning. In what follows, a detailed description of both types will be presented.

3.1. REHEARSAL

As was noted by Ellis (2009), "rehearsal can be seen as a special type of pre-task planning" (p. 476). When learners are rehearsing a task that has already been completed several times, they are in fact preparing their performance for the task. Rehearsal has also been considered as *task repetition* since, as pointed out by Lambert *et al.* (2020), it is the opportunity to practice the performance of a task more than one time. Rehearsal has been extensively examined in L2 research for several reasons: (1) it has been reported to ease online processing demands at the time of task performance given that (2) they are able to produce iterations, lexis, and utterances which may be repeated afterward, with the (3) benefit of reusing the same content or honing it. Lambert *et al.* (2020) distinguish various lines of inquiry on rehearsal: immediate task repetition (Lambert *et al.*, 2017), repetition after a short-term period (Gass *et al.*, 1999), and repetition after a long-term period (Bygate, 1996).

Research on the last two lines of inquiry has pointed to two different terms into which rehearsal may be divided: same task repetition and parallel task repetition. The first refers to the repetition of the same task under the same conditions as on the first occasion whilst the second involves repeating the same task type but containing different content. Although initial studies (e.g. Gass *et al.*, 1999) pointed to no apparent benefits from the use of rehearsal when learners performed the same

task type with different content, more recent research has proved the opposite (Kim & Tracy-Ventura, 2013). In fact, parallel task repetition may reinforce the learners' ability to monitor their attention during task performance. Nonetheless, research has concluded quite successfully that rehearsal favors subsequent task performance of the *same* task. The measures where task repetition has been more prominent have been fluency and complexity. For instance, Bygate's (1996) laboratory study was centered on parallel task repetition, with a 10-week gap to repeat the task. The issue of transferability, as mentioned previously, is of the utmost importance since should gains in task performance not be transferred to similar tasks, learning may not have occurred (Lambert *et al.*, 2020). As argued by Ellis (2009), this issue may raise serious concerns since the ultimate aim would be for learning to occur. Thus, if learners do not perform appropriately in new or similar tasks, then the role of rehearsal may not be regarded as useful in advancing acquisition. To overcome this issue, several voices (e.g. Sheppard, 2006) have posited that task repetition would be useful only when some type of feedback was provided. This goes in accordance with Ellis's (2009) claim that the benefits of rehearsal alone may not go beyond a slight effect on performance. Conversely, the use of feedback techniques may assist learning toward acquisition by reinforcing the three linguistic measures (complexity, accuracy, and fluency). Very interestingly, Lambert *et al.* (2020) noted that parallel task repetition may not have an immediate effect on task performance, but it may result in the proceduralization of fluency. This might be the case, but still, the presence of other variables until the next performance (i.e. learning more vocabulary, lexis or other practice) could be reasonably associated with this increase in fluency or other measures.

3.2. STRATEGIC PLANNING

Strategic planning entails, in Ellis's (2009) words, "planning what content to express and what language to use but without the opportunity to rehearse the complete task" (p. 474). This type of planning has received scholarly attention, and findings have pointed to benefits on fluency measures (e.g. Skehan & Foster, 1997; Ellis, 2005; Wigglesworth

& Elder, 2010). In his comprehensive review, Ellis (2009) traverses through the most important variables that studies on strategic planning have dealt with. Firstly, the participants, i.e. learners, are central to any L2 research study. In the case of strategic planning, studies have relied on L2 and foreign language learners. Secondly, the role of proficiency as a variable is central in strategic planning given the importance of other variables such as negotiation of meaning or peer interaction. In this case, Ellis (2009) indicates that studies on strategic planning have not made it explicit what level of proficiency participants had. Most of the studies on pre-task planning have relied upon three traditional settings: classroom-based research, laboratory research, and testing-based research. Other variables examined in pre-task planning research have to do with the type of task (whether these are interactive or monologic, simple or complex) and the planning variables (i.e. the length of planning, guided or unguided planning, or form-focused or meaning-focused planning).

As strategic planning is the central piece of this empirical study, I will mainly concentrate on the prolific amount of research that has appeared in the last decades regarding this type of planning. More specifically, we will focus on the effect of planning on the three aspects of language, that is, fluency, complexity and accuracy (in other words, CAF).

Fluency: Oral production is mainly a communicative operation where the economy of language applies to speech speed. Hence, fluency plays an important role in how pre-task planning may condition its effect. The operationalization of fluency has been manifold and has ranged from holistic rating measures (e.g. Elder & Iwashita, 2005; Tavakoli & Skehan, 2005) to temporal aspects or other phenomena (such as reformulations), as precisely reviewed by Ellis (2009). To start with, studies conducted in testing-based environments pointed to mixed findings in terms of fluency, some reporting no effects and others negative ones (see Wigglesworth, 2001; Elder & Iwashita, 2005). However, most of the studies, and classroom-based research particularly, have reported positive effects of strategic planning on fluency measures.

Another variable affecting the measure of fluency was proficiency level, where studies on strategic planning have predominantly reported varied findings. In general terms, it has been suggested that low proficient learners are less likely to benefit from planning procedures (e.g. Tavakoli & Skehan, 2005; Mochizuki & Ortega, 2008), contrary to the tendency in higher-proficient learners. In this case, more proficient learners have been shown to take more advantage of strategic planning (Wigglesworth, 1997; Mochizuki & Ortega, 2008). However, as Ellis (2009) noted, "no conclusion regarding the value of planning for learners with different levels of proficiency is possible yet" (p. 493).

Another variable worth considering is task complexity as well as the type of task. Among the many types, e.g. personal information tasks, narrative tasks, and decision-making tasks, the latter is the toughest one given the degree of familiarity with the content to be communicated, and the structure. Several studies have accounted for the apparent advantages of strategic planning (and planning in general) for simpler tasks (e.g. Foster & Skehan, 1996), even if tasks are decision-making ones. The structure of tasks has also been found to be largely influential with a clear effect on fluency measures (Tavakoli & Skehan, 2005). Finally, the length of planning is regarded as a powerful variable on fluency, and a commonly acknowledged finding has been that, when the planning time is longer, fluency increases (see Mehnert, 1998). In the same vein, the type of planning has been reported to influence the degree of fluency. Whilst studies have not been conclusive as to the role of the type of planning, with very diverse findings, some of them have pointed to guided planning as more beneficial for the enhancement of fluency (e.g. Kawauchi, 2005; Sangarun, 2005). However, other studies (e.g. Yuan & Ellis, 2003; Gilabert, 2007) reported fluency gains when unguided planning was used as strategic planning.

Complexity: The operationalization of complexity has been varied and has ranged from measuring the number of T-units, the number of verb forms or word types, or the use of holistic ratings.

The effect of strategic planning on complexity has not been conclusive (Ellis, 2009). Nonetheless, studies have proved that there is an

enhancement in complexity when planning is involved. More specifically, grammatical gains have been reported along with lexical complexity (Gilabert, 2007).

As regards proficiency, findings have pointed to no effect of planning on this variable, with the sole exception of Kawauchi (2005), whose results indicated that advanced learners did not enhance complexity with strategic planning. Similarly, the setting (i.e. classroom-based research or testing conditions) does not seem to affect language complexity. Task factors such as how the task outcome is presented (e.g. monologic or interactive) are not conclusive either, but studies such as Yuan and Ellis (2003) revealed that planning led to more complexity. Other studies, such as Gilabert (2007), did not report any influence of task complexity on lexical complexity.

In the case of types of planning, previous research has indicated that a certain degree of guidance (i.e. guided planning) may lead to greater grammatical complexity (e.g. Foster & Skehan, 1996), even if this only occurred in the short term (Skehan & Foster, 2005).

Accuracy: Previous research has not yielded clear results as regards accuracy, with studies showing mixed perspectives. Nonetheless, most of the research on accuracy with strategic planning has reported gains in this aspect of language. Unlike in the case of complexity, the level of proficiency seems to hold a pivotal role in making planning an effective tool for language enhancement. Results reported in previous research have shown that there are accuracy gains for low-proficient learners (Kawauchi, 2005). Likewise, task type seems to influence accuracy when planning. For instance, decision-making tasks are thought to lead to more accuracy (Foster & Skehan, 1996), although the same authors reported that results were not uniform and therefore inconclusive in a later study (Skehan & Foster, 2005). Along this line, the fact that a task is structured has an equal bearing on the accuracy gains, as Mehnert's (1998) study reported, where an effect was revealed.

4. COMMUNICATIVE TASKS IN HIGHER EDUCATION

In a university setting, tasks have to be not only cognitively demanding but foster in equal terms several competencies. In this case, the relevance of decision-making tasks shall be conveniently understood. Similarly, these tasks must be participatory in terms of peer interaction, that is, conducted under a group format. Framed within the Task-Based Language Teaching (TBLT) approach, tasks are characterized by their imitating nature with an aim to simulating real-life situations. Tasks are reported to be essentially communicative notwithstanding the varied definitions in the scholarly literature. Several authors (e.g. Long, 1985; Richards *et al.*, 1985) defined tasks as everyday life actions where processing or understanding language took place in a given situation. Another definition by Breen (1987) focused on pragmatics given the nature of communicative tasks as group problem-solving. As reviewed in Garcés-Manzanera (2021b), the *communicative* was added to 'task' after the consideration that they led to a communicative outcome (Willis, 1996; Skehan, 1998). This communication and the goal involved are determined by the criteria that define a communicative task: goal, input, setting, activities, roles, and feedback (see Skehan, 1998). Considering the setting, that is, Higher Education, communicative tasks have to be specially designed to comply with the field of study. In this regard, our study explores the use of communicative tasks in Primary Education undergraduates. Such a specific group of learners requires that tasks are adequately tailored to the context where they are going to develop their professional activity. Recent theoretical and empirical attempts (see Garcés-Manzanera, 2021) have proposed what has been called *education-oriented communicative tasks*, where the tasks do not only hinge on the education context but also provide learners with the necessary subject matter knowledge with specific vocabulary. This was defined as "English for Teaching Purposes", whose focus has been placed on training future Early Years and Primary Education teachers to be able to control the register and have a good command of the education-related aspects in a school context. The main characteristics of communicative tasks for the education context are, according to Long's (1989) distinctions, two-way and open tasks. This means that learners

have all the instructions to the task before commencing it, and are therefore allowed to provide a creative solution.

Hence, communicative tasks such as the ones used in the present empirical study are centered on furthering future teachers' professional skills and their communicative abilities by simulating situations they will very likely encounter in school settings.

5. AIM AND RESEARCH QUESTIONS

The main aim of the present study is to observe the extent of the difference in holistic CAF measures in pre-task planning in both guided and unguided planning conditions and to observe the relationship between the rating score and the speaking time between both conditions.

The research questions guiding this study are:

1. To what extent are the rating scores for CAF measures different when learners follow a guided and unguided planning procedure?
2. Does a relationship exist between the speaking time and the impressionistic CAF rating scores as a function of the type of planning?

6. METHODOLOGY

CONTEXT & PARTICIPANTS

Our pool of participants consisted of 25 undergraduate students, with an average age of 18-19. Their average proficiency level was B1 (intermediate) according to the CEFR.

The context of this study is a Spanish university, with first-year Primary Education undergraduates. All were pursuing an EFL compulsory module, receiving two hours a week. One of these hours was fully devoted to theoretical aspects of language, and the second hour was centered on developing speaking skills. Our study was conducted in this second hour.

Our research design consisted of two phases involving: pre-task and task outcome with language- and content-focused feedback. In this sense, students were distributed into two types of planning: guided planning and unguided planning. Beforehand, they had to form three to four groups of students to perform the task.

Table 1. *Procedure for the development of the task.*

Phase	Guided planning	Unguided planning
1) Pre-task planning	15 minutes preparation Use of pre-task planning note-taking sheets. Participatory and organization.	
2) Task outcome	Monologic (7 min. maximum per group).	

As seen in table 1, in the first phase, students had to prepare their task outcome in 15-minute time. Organized into groups, all participants had to debate and reach a consensus about the task outcome in a participatory manner. In both planning conditions, each group was provided with a pre-task planning note-taking sheet. Only one of the members of the group was allowed to write in this note-taking sheet, in an attempt to foster cooperation and peer interaction. After the 15-minute window of time to prepare the task, each group had to present their task outcomes. The procedure involved that each member of the group had to take part orally in the presentation. Given the time limitation per group (7 minutes), every contribution had to be purposefully organized so as to comply with this requirement.

INSTRUMENTS

Several instruments were used as part of this study:

a. *The communicative task*. The task used in this study is based on the ones presented in previous research (Garcés-Manzanera, 2021b). One of the notable aspects of this communicative task, as announced before, is that it is intimately related to the education context. The task consisted of two internal preparation

phases: firstly, students were encouraged to become interested in the topic of the task with a warm-up (see figure 2 below), and secondly, the task instructions were presented (see figure 3 on the next page). As can be seen, participants, as future teachers, have to engage in the preparation of a task that is highly linked to their field of study and work. Afterward, participants followed the procedure indicated in table 1.

FIGURE 2. *Warm-up of the education-oriented communicative task.*

WARM-UP

We've found that first- and second-grade children, when presented with a variety of drawn cartoon character faces they haven't seen before, have no problem sorting them into "good" and "bad" characters.

In fact, many children have clearly developed ideas and are able to tell us lengthy stories about why they think a particular character might be a hero or villain with minimal information. Sometimes this seems to be based on their belief that a character looks like another media character they've seen. They'll then make the assumption that a face they're shown looks like "a princess" or "someone who goes to jail." With the lack of diversity in the world of children's television, it's not surprising that kids would make associations with so little information. But it's also a bit alarming – given what we know about the prevalence of stereotyping – that children seem so quick to make attributions of who's good and who's evil.

FIGURE 3. *Instructions of the education-oriented communicative task.*

TASK

- Watching films and reading books are two activities closely related to enhancing children personal development.
- In groups of 3-4, prepare **two activities** (one for films and the other for books) in order to spark Primary 4 students' interest in both books and films.
- Remember to mention:
    - Justification.
    - Activity (explanation).
    - **How** the activity promotes books or films.
    - **Advantages and disadvantages**.

---

b. *Unguided planning note-taking sheet*. As seen in table 1, partic-
   ipants in the unguided pre-task planning condition were pro-
   vided with this note-taking sheet. Figure 4 displays what
   learners had to write and what other elements were present in
   the planning procedure. All the members of the group had to
   write their names in addition to indicating who was in charge
   of writing the ideas generated out of the discussion. Given the
   fact that it was *free* planning, no indications were provided as
   to how they had to organize it. Nonetheless, a helping chart
   was present at the bottom of the page where useful lexis, gram-
   matical structures, or linkers may be written.

FIGURE 4. *Unguided planning note-taking sheet.*

c. *Guided planning note-taking sheet*. Figure 5 displays that participants in this planning condition had more *guidance* to organize their tasks. The header contains the same information as in the previous one, and the bottom helping table as well. Looking more closely at the note-taking sheet (based on Kayfetz & Stice, 1987), participants may write or draft their position as regards the organization of their presentation. After stating their position, they may organize the ideas with the schematic division proposed. Likewise, they are expected to draw some conclusions for their presentation. In essence, participants are explicitly encouraged to follow an organizational pattern: introduction/aim, ideas/arguments, and conclusion.

FIGURE 5. *Guided planning note-taking sheet.*

d. *Rubric for holistic CAF measures.* To assess the performance of participants after the planning conditions, we opted for a rubric (see figure 6) used in previous research (Iwashita *et al.*, 2001), which entailed the impressionistic rating of each CAF measure (see Wigglesworth, 2001; Tajima, 2003). Each measure was rated from 0 to 8. In addition to this, the researcher also calculated the amount of time that each speaker used in their part of the presentation.

FIGURE 6. *Rubric of CAF measures.*

Fluency

8	Speaks fairly fluently with only occasional hesitation, false starts and modification of attempted utterance. Speech is only slightly slower than that of a native speaker.
7	
6	Speaks more slowly than a native speaker due to hesitations and word-finding delays.
5	
4	A marked degree of hesitation due to word-finding delays or inability to phrase utterances easily.
3	
2	Speech is quite disfluent due to frequent and lengthy hesitations or false starts.
1	
0	Speech is so halting and fragmentary that conversation is impossible.

Accuracy

8	Errors are not unusual, but rarely major.
7	
6	Manages most common forms, with occasional errors, major errors present.
5	
4	Limited linguistic control: major errors frequent.
3	
2	Clear lack of linguistic control even of basic forms.
1	
0	No linguistic control even of the most basic forms.

Complexity

8	Attempts a variety of verb forms (e.g. passives, modals, tense and aspect), even if the use is not always correct. Takes risks grammatically in the service of expressing complex meaning. Regularly attempts the use of coordination and subordination to convey ideas that cannot be expressed in a single clause, even if the result is awkward or incorrect.
7	
6	Mostly relies on simple verb forms, with some attempt to use a greater variety of forms (e.g. passives, modals, more varied tense and aspect). Some attempt to use coordination and subordination to convey ideas that cannot be expressed in a single clause.
5	
4	Produces numerous sentence fragments in a predictable set of simple clause structures. If coordination and/or subordination are attempted to express more complex clause relations, this is hesitant and done with difficulty.
3	
2	Produces mostly sentence fragments and simple phrases. Little attempt to use any grammatical means to connect ideas across clauses.
1	
0	No awareness of basic grammatical means.

MEASURES

The measures used in the present empirical study are the following ones: *speaking time* (in seconds), and the CAF measures ratings, that is, *fluency*, *accuracy*, *complexity*, and the *total score* (rating score 0-8).

STATISTICAL ANALYSES

A series of statistical analyses were performed on the previously mentioned measures. For RQ1, the mean as a measure of central tendency and standard deviation as a measure of dispersion was used to display the data. To observe to what extent both conditions are different, the effect size (Hedges' G) was computed. The magnitude of the effect was interpreted by following Plonsky and Oswald's (2014) guidelines for SLA research, that is, 0.40 (small), 0.70 (medium), and over 1.00 (large).

For RQ2, however, a linear regression was calculated to observe whether there is an association between the different measures and the speaking time depending on the planning condition that learners followed. To perform this linear regression analysis, 'speaking time' was taken as an outcome variable (dependent) while CAF measures were considered as predictor variables. The association was observed by performing a linear regression analysis on the data belonging to each planning procedure separately. As a supplement to the regression analysis, a series of Spearman Rho correlations were also performed between the previously mentioned measures.

7. RESULTS

RQ1) To what extent are the rating scores for CAF measures different when learners follow a guided and unguided planning procedure?

As can be seen in table 2, there are not many practically significant differences across the measures except for the measure of complexity. In this regard, the effect size is small and indicates that learners in the unguided planning procedure achieved a higher score.

Table 2. *Descriptive statistics for the score of CAF measures and speaking time.*

Measure	Type of planning		Effect size
	Guided planning	Unguided planning	Hedges G
Speaking_Time	46.06 (26.79)	51.40 (32.42)	**-0.17**
Score_Total	5.64 (1.07)	5.93 (0.58)	**-0.32**
Fluency_Total	5.60 (1.29)	5.80 (0.63)	**-0.18**
Accuracy_Total	5.53 (0.83)	5.70 (0.82)	**-0.19**
Complexity_Total	5.80 (1.20)	6.30 (1.05)	**-0.42**

RQ2) Does a relationship exist between the speaking time and the impressionistic CAF rating scores as a function of the type of planning?

We will next report on the multiple linear regression analyses performed on the measures for each planning condition, i.e. unguided planning, and guided planning.

Unguided planning. Tables 3 to 5 display the quantitative information on the multiple linear regression analysis performed on the unguided planning condition.

In this case, a multiple linear regression was calculated to predict speaking time based on the different CAF measures. A significant regression equation was found ($F(3, 13) = 10.511$, $p = .008$), with an R^2 of .840 (see table 3 below).

Table 3. *Regression outcome (1).*

Model	R	R^2	Adjusted R^2	RMSE
H_0	0.000	0.000	0.000	32.428
H_1	0.917	0.840	0.760	15.880

Participants' predicted speaking time is equal to 168.077 – 38.564 (Fluency_Total) + 9.844 (Accuracy_Total) + 8.077 (Complexity_Total). The participants' speaking time decreased 38 seconds for each point in the Fluency measure, as this was a significant predictor of the dependent variable (speaking time).

Table 4. *Regression outcome (2).*

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
H_1	Regression	7951.427	3	2650.476	10.511	**0.008**
	Residual	1512.973	6	252.162		
	Total	9464.400	9			

Note. The intercept model is omitted, as no meaningful information can be shown.

Table 5. *Regression outcome (3).*

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
H_0	(Intercept)	51.400	10.255		5.012	< .001
H_1	(Intercept)	168.077	65.833		2.553	0.043
	Fluency_Total	-38.564	9.029	-0.752	-4.271	**0.005**
	Accuracy_Total	9.844	8.636	0.250	1.140	0.298
	Complexity_Total	8.077	6.875	0.264	1.175	0.285

The previous multiple linear regression was confirmed through a series of Spearman Rho correlations (table 6):

Table 6. *Spearman Rho Correlations between Speaking Time and CAF measures.*

Spearman's Correlations (Free)

	Spearman's rho	p
Speaking_Time - Fluency_Total	**-0.777** **	0.008
Speaking_Time - Accuracy_Total	0.391	0.264
Speaking_Time - Complexity_Total	**0.596**	0.069

* $p < .05$, ** $p < .01$, *** $p < .001$

As can be observed, there is a statistically significant correlation between Speaking Time and the score obtained in Fluency. There is also a moderate correlation between Speaking Time and Complexity, meaning that, the more time, the more complexity. Nonetheless, this correlation was not statistically significant.

Guided planning. Tables 7 to 9 display the quantitative information on the multiple linear regression analysis performed on the guided planning condition.

Multiple linear regression was calculated to predict speaking time based on the different CAF measures. A non-significant regression equation was found ($F(3, 11) = 2.251$, $p = .139$), with an R^2 of .380 (see table 7 below).

Hence, no influence was observed from any of the measures on the dependent variable (speaking time).

Table 7. *Regression outcome (1).*

Model Summary - Speaking_Time				
Model	R	R²	Adjusted R²	RMSE
H_0	0.000	0.000	0.000	26.794
H_1	0.617	0.380	0.211	23.794

Table 8. *Regression outcome (2).*

ANOVA						
Model		Sum of Squares	df	Mean Square	F	p
H_1	Regression	3823.107	3	1274.369	2.251	0.139
	Residual	6227.826	11	566.166		
	Total	10050.933	14			

Note. The intercept model is omitted, as no meaningful information can be shown.

Table 9. *Regression outcome (3).*

Coefficients						
Model		Unstandardized	Standard Error	Standardized	t	p
H_0	(Intercept)	46.067	6.918		6.659	< .001
H_1	(Intercept)	-18.432	47.129		-0.391	0.703
	Fluency_Total	17.568	12.049	0.851	1.458	0.173
	Accuracy_Total	0.798	18.059	0.025	0.044	0.966
	Complexity_Total	-6.603	14.225	-0.297	-0.464	0.652

Additionally, we performed a series of Spearman Rho statistical tests to observe the potential association between Speaking and CAF measures (table 10 below).

Table 10. Spearman Rho correlations between the ratings of CAF measures and Speaking time.

Spearman's Correlations

	Spearman's rho	p
Speaking_Time - Fluency_Total	0.471	0.077
Speaking_Time - Accuracy_Total	**0.525** *	0.044
Speaking_Time - Complexity_Total	0.504	0.055

* $p < .05$, ** $p < .01$, *** $p < .001$

As can be observed, there is a moderate correlation between Speaking Time and Accuracy. This means that the longer the duration of speaking, the more possibilities there are for the learner to obtain a higher score in accuracy.

In the next section, we provide an interpretation of these results.

8. DISCUSSION

The preceding section has provided a descriptive and statistical presentation of the analyses performed. Our first research question asked whether there was any influence on CAF measures as a function of the planning procedure. One of the first findings in this respect was that, overall, both planning conditions had a similar positive effect on the measure of fluency. Also, participants in the unguided planning procedure had a higher score in complexity than those in the guided planning procedure. This increase in complexity is related to other studies, such as Wendel (1997), in which planning led to the same outcome, also in fluency. This is similar to Wigglesworth's (1997) findings, which indicated that planning led to greater fluency.

Another finding pointed out that participants in the guided planning condition used less speaking time in their task outcome as in previous research (Foster & Skehan, 1996). Such a finding is aligned with what

was found in the case of unguided planning, and the potential association between speaking time and fluency score. In this case, we observed that unguided planning favored an improvement of the fluency score. This finding runs counter to Skehan and Foster's (2005) study, which claimed that undetailed planners –in our study, unguided planning condition– led to less filled pauses, and thus less fluency.

Likewise, we also found out that, in the guided planning condition, more speaking time was associated with an increase in accuracy score. This might be related to the orientation of the task since participants were familiar with the type of task they were undertaking. Additionally, the provision of a guided planning note-taking sheet might have favored more concentration on the forms. In the case of the unguided planning condition, as mentioned previously, there was a positive association between speaking time during task and complexity scores. One of the potential explanations behind this finding might be related to the participants' proficiency level, as a potential influential variable. This was particularly evident in Tavakoli and Skehan's (2005) study, where they found that proficiency level was a pivotal variable influencing CAF measures.

As mentioned, the difference between the guided and unguided planning conditions was particularly visible in the association between speaking time during the task and the fluency score. This was also corroborated by a recent study (Mochizuki & Ortega, 2008), where secondary students participated in a one-way communicative task under three different conditions (no planning, unguided planning, and guided planning). Although the variables in this study differ from the ones in our study, certain parallelism may be established. Additionally, Mochizuki and Ortega (2008) found that guided planners produced more accurate language than unguided planners. In our study, we observed that participants in the guided planning condition were more inclined to increase their accuracy scores when speaking time, as an interrelated variable, increased. This was highly confirmed with the highly correlated values obtained between the CAF ratings, which might indicate that learners in the guided planning condition displayed a more balanced performance.

9. CONCLUSIONS

Our study has attempted to provide empirical evidence to the wide array of studies on pre-task planning in communicative tasks. Following previous research (e.g. Skehan & Foster, 2005; Gilabert, 2007; Mochizuki & Ortega, 2008), our study explored the potential effect of guided and unguided planning on impressionistic CAF measures.

One of the main conclusions to be drawn is that not one specific type of planning condition (i.e. guided or unguided) has proved to be better than the other. Despite the visible differences in terms of association of measures, the between-group analyses have shown that no special differentiations were observed in these conditions.

Furthermore, fluency was reported to be highly similar in both guided and unguided planning. However, the guided planning group spoke more fluently, even if the effect on fluency ratings was not as consistent, which occurred in previous research (Guará-Tavares, 2008; Mochizuki & Ortega, 2008).

In terms of complexity, unguided planning increased this measure and was moderately correlated with speaking time. We can conclude that the more time spent on speaking, the more complexity there is.

A question that arises in this context is whether proficiency-related factors might have affected considerably the performance in each of the groups, and also, whether individual differences such as language aptitude or affective variables may have had some influence on how they perceived, but most importantly, handled each type of planning condition.

Our study is not without its limitations: (i) we made use of a small pool of participants which, undoubtedly, might have affected some of the results, and (ii) only one proficiency level was taken into consideration.

Future research avenues should bear in mind these limitations since a larger pool of participants would ensure that our results are extrapolated to a greater extent. Similarly, the relevance of the proficiency-level variable is unarguably pivotal for exploring how pre-task planning may

be different not only as a function of task type but also within-subjects. Another important gap that has to be filled in future research is the exploration of the pre-task planning conditions on CAF measures with other types of tasks in the same classroom-based environment. In our case, the present empirical research was conducted in a tertiary level context. It would be convenient to observe whether there are any considerable differences between education stages. Similarly, other variables such as negotiation of meaning and peer-interaction should be borne in mind for future research insomuch as both could shed light on what occurs during the preparation of the task outcome. Finally, further examination of the qualitative data from the note-taking sheets may bring to light what trends participants follow to organize their tasks.

10. REFERENCES

Breen, M. P. (1987). Contemporary Paradigms in Syllabus Design Part II. *Language Teaching*, *20*(3), 157-174.

Butterworth, B. (1980). Evidence from Pauses in Speech. In B. Butterworth (Ed.), *Language Production: Speech and Talk* (pp. 155-176). Academic Press.

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). Macmillan.

Crookes, G. (1989). Planning and Interlanguage Variation. *Studies in Second Language Acquisition*, 11, 367-383.

Elder, C., Iwashita, N. & McNamara, T. (2002). Estimating the Difficulty of Oral Proficiency Tasks: What Does the Test-Taker Have to Offer? *Language Testing*, *19*(4), 347-368.

Ellis, N. C. (2005). At the Interface: Dynamic Interactions of Explicit and Implicit Language Knowledge. *Studies in Second Language Acquisition*, 27, 305-52.

Ellis, R. (2009). Task-Based Language Teaching: Sorting Out the Misunderstandings. *International Journal of Applied Linguistics*, *19*(3), 221-246.

Faerch, C. & Kasper, C. (1983). Plans and Strategies in Foreign Language Communication. In C. Faerch & C. Kasper (Eds.), *Strategies in Interlanguage Communication* (pp. 20-44). Longman.

Foster, P. & Skehan, P. (1999). The Influence of Source of Planning and Focus of Planning on Task-Based Performance. *Language Teaching Research*, 3, 215-247.

Garcés-Manzanera, A. (2021a). El contexto educativo en las materias de lengua extranjera (inglés) en los grados de Educación Primaria de las universidades españolas. In R. Pérez Cabana & S. A. Flores Borjabad (Coords.), *Nuevos retos y perspectivas de la investigación en Literatura, Lingüística y Traducción* (pp. 1406-1430). Dykinson.

Garcés-Manzanera, A. (2021b). Implementing Communicative Tasks for the Education Context: Primary Education Undergraduates' Perceptions of a Didactic Proposal. In REDINE (Coord.), *Medios digitales y metodologías docentes: mejorar la educación desde un abordaje integral* (pp. *n. d.*). Adaya Press.

Gass, S.M., Mackey, A., Álvarez-Torres, M. J. & Fernández-García, M. (1999). The Effects of Task Repetition on Linguistic Output. *Language Learning*, *49*(4), 549-581.

Gilabert, R. (2007). The Simultaneous Manipulation of Task Complexity Along Planning Time and (+/– Here and Now): Effects on Oral Production. In M. Mayo (Ed.), *Investigating Tasks in Formal Language Learning* (pp. 44-68). Multilingual Matters.

Iwashita, N., Elder, C. & McNamara, T. (2001). Can We Predict Task Difficulty in an Oral Proficiency Test? Exploring the Potential of an Information Processing Approach to Task Design. *Language Learning*, 21, 401-436.

Kawauchi, C. (2005). The Effects of Strategic Planning on the Oral Narratives of Learners with Low and High Intermediate L2 Proficiency. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language* (pp. 143-165). John Benjamins.

Kayfetz, J. L. & Stice, R. L. (1987). *Academically Speaking*. Heinle & Heinle Publisher.

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.

Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Elrbaum.

Lambert, C., Aubrey, S. & Leeming, P. (2020). Task Preparation and Second Language Speech Production. *TESOL Quarterly*, *55*(2), 331-365. https://doi.org/10.1002/tesq.598

Lambert, C. & Robinson, P. (2014). Learning to Perform Narrative Tasks: A Semester-Long Classroom Study of L2 Task Sequencing Effects. In P. Robinson, M. Baralt & R. Gilabert (Eds.), *Task Sequencing and Instructed Second Language Learning* (pp. 207-230). Coniuum. https://bit.ly/3A3wDC1

Lambert, C., Kormos, J. & Minn, D. (2017). Task Repetition and Second Language Speech Processing. *Studies in Second Language Acquisition*, *39*(1), 167-196.

Levelt, W. (1989). *Speaking: From Intention to Articulation*. The MIT Press.

Levelt, W. (1999). Producing Spoken Language: A Blueprint of the Speaker. In C. Brown & P. Hagoort (Eds.), *The Neurocognition of Language* (pp. 83-122). Oxford University Press.

Long, M. (1998). Focus on Form in Task-Based Language Teaching. *University of Hawai'i Working Papers in ESL*, 16, 35-49.

Long, M. H. (1985). A Role for Instruction in Second Language Acquisition: Task-Based Language Teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modeling and Assessing Second Language Development* (pp. 77-99). Multilingual Matters.

MacWhinney, B. & Osser, H. (1977). Verbal Planning Functions in Children's Speech. *Child Development*, 48, 978-985.

Mehnert, U. (1998). The Effects of Different Lengths of Time for Planning on Second Language Performance. *Studies in Second Language Acquisition*, 20, 83-108.

Mochizuki, N. & Ortega, L. (2008). Balancing Communication and Grammar in Beginning-Level Foreign Language Classrooms: A Study of Guided Planning and Relativization. *Language Teaching Research*, 12, 11-37.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L. & Russo, R. P. (1985). Learning Strategies Used by Beginning and Intermediate ESL Students. *Language Learning*, *35*(1), 21-46.

Plonsky, L. & Oswald, F. L. (2014). How Big Is "Big"? Interpreting Effect Sizes in L2 Research. *Language Learning*, *64*(4), 878-912.

Richards, J., Platt, J. & Weber, H. (1985). *Longman Dictionary of Applied Linguistics*. Longman.

Sangarun, J. (2001). *The Effects of Pre-Task Planning on Foreign Language Performance*. [PhD dissertation, University of Toronto]. Institutional Repository U of T. https://bit.ly/3B3Ura1

Sangarun, J. (2005). The Effects of Focusing on Meaning and Form in Strategic
    Planning. In R. Ellis (Ed.), *Planning and Task Performance in a Second
    Language* (pp. 111-143). John Benjamins.

Sheppard, C. (2006). *The Effects of Instruction Directed at the Gaps Second Language
    Learners Noticed in Their Oral Production* [Unpublished PhD thesis,
    University of Auckland].

Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford University
    Press.

Skehan, P. (Ed.) (2014). *Processing Perspectives on Task Performance* (Vol. 5). John
    Benjamins.

Skehan, P. & Foster, P. (2005). Strategic and On-Line Planning: The Influence of
    Surprise Information and Task Time on Second Language Performance. In
    R. Ellis (Ed.), *Planning and Task Performance in a Second Languag*e (pp.
    193-218). John Benjamins.

Tajima, M. (2003). *The Effects of Planning on Oral Performance of Japanese as a
    Foreign Language* [PhD dissertation, Purdue University]. Institutional
    Reposiory PU. https://bit.ly/3iorFda

Tavakoli, P. & Skehan, P. (2005). Strategic Planning, Task Structure and
    Performance Testing. In R. Ellis (Ed.), *Planning and Task Performance in a
    Second Language* (pp. 239-277). John Benjamins.

Wendel, J. (1997). *Planning and Second Language Narrative Production.*
    [Unpublished PhD dissertation, Temple University].

Wigglesworth, G. & Elder, C. (2010). An Investigation of the Effectiveness and
    Validity of Planning Time in Speaking Test Tasks. *Language Assessment
    Quarterly*, *7*(1), 1-24.

Wigglesworth, J. (2001). Influences on Performance in Task-Based Oral
    Assessments. In Bygate, M., Skehan, P. & Swain, M. (Eds.), *Researching
    Pedagogic Tasks: Second Language Learning, Teaching, and Testing* (pp. 186-
    209). Longman.

Willis, J. (1996). A Flexible Framework for Task-Based Learning. In D. Willis & J.
    Willis (Eds.) *Challenge and Change in Language Teaching* (pp. 52-62).
    Heinemann.

Yuan, F. & Ellis, R. (2003). The Effects on Pre-Task Planning and On-Line
    Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral
    Production. *Applied Linguistics*, *21*(1), 1-27.