The Effect of Task Repetition on Oral Task Performance in a Japanese High School EFL Class
日本の高等学校の英語の授業で行われる口頭でのタスクパフォーマンスにおけるタスクの繰り返しが与える効果

Edward McShane
マクシェイン・エドワード

***Abstract*:** The purpose of this paper is to measure the effect of task repetition as part of a task-based language teaching sequence in an EFL classroom. Research has shown that task repetition can lead to improvements in the complexity, accuracy, and fluency (CAF) of spoken task performance; however, most studies lack ecological validity. High school students were recorded performing two similar monologic spoken tasks. Significant gains were made in syntactic complexity and fluency, with no loss of accuracy on the second performance. The results support research that suggests task-based language teaching sequences involving task repetition can lead to inter-language development.

***Keywords*:** task-based, language teaching, repetition, complexity, accuracy, fluency

***要旨*:** 本稿は英語の授業で行われている、口頭でのタスクパフォーマンスにおけるタスクの繰り返しが与える効果測定が目的である。繰り返しタスクを行うことによって、スピーキング能力の CAF（複雑さ・正確さ・流暢さ）が向上するとの結果が得られている。しかし、ほとんどの研究は生態学的妥当性を欠いている。高校生が似た長文のスピーキングタスクを 2 回行ったのを録音したところ、2 回目目には文章上の正確さを失うことなく、複雑な統語を流暢に使う能力が向上した。この結果から、タスクを繰り返を含みながら生徒に継続してタスク型指導を行うことが言語間の発達に繋がるという説を立証する。

***キーワード*:** タスク型、言語教育、繰り返し、CAF（複雑、正確、および流暢）

1. Introduction

Task Based Language Teaching (TBLT), which involves meaning focused classroom activities based on real world activities (Long, 2014), ‘has become a dominant paradigm in the teacher education literature' (Lynch & MacLean, 2000:223). The approach is criticised for prioritising meaning at the expense of form, and for being unsuitable for English as a Foreign Language (EFL) contexts (Swan, 2005; Sheen, 1994); however, research has shown that form and meaning can improve together using TBLT methods in EFL classrooms (Skehan, 2009; Ellis, 2009b). Nevertheless, because of learners' limited working memory capacity, there is a trade-off between two aspects of form:

accuracy and meaning (Skehan, ibid). Task repetition is seen as a way of combining pre-task and during-task planning to ease the load on working memory, allowing learners to improve all aspects of a performance (Bygate & Samuda, 2005); however, most studies involving task repetition have been conducted under controlled conditions (Ellis, ibid), lacking ecological validity. Swan (ibid) also criticises TBLT for abandoning traditional methods without warrant. While some varieties of TBLT do dismiss traditional methods, others integrate TBLT methods with those of traditional approaches (Ellis, 2009a).

The purpose of this study is to investigate how task performance changes with repetition of a task that has been carried out as part of a TBLT sequence in an intact classroom, providing ecological validity.

2. Literature Review

2.1 Task-based language teaching

Advocates of TBLT criticise traditional methods, such as Presentation-Practice-Production (PPP), for being synthetic by teaching linguistic forms out of context, and ignoring that aspects of language are interrelated and cannot be learned in isolation (Foster, 1999; Nunan, 2004; Ellis, 2009a). In traditional approaches, syllabuses and lessons are organised around linguistic features, such as structures, functions, or lexis, whereas TBLT uses the task as the unit of organisation (Long, 2014).

Definitions of *task,* in TBLT vary considerably. For instance, Long's definition (1985) refers to real world activities without any mention of linguistic or pedagogic aspects, whereas Nunan's (2005:4) involves the use of 'target language' in the 'classroom' to 'express meaning'. Ellis (2009a) distinguishes between focused and unfocused tasks, with the former targeting linguistic forms, and the latter selected for general communication. Advocates of TBLT agree that tasks should have a clearly defined outcome and be meaning focused (Van den Branden, 2016), which Ellis (ibid) explains constitutes both pragmatic and semantic meaning. Willis and Willis (2007:ch 1.5) draw on Skehan's (1998) definition to provide a useful list of questions, not to determine whether an activity is or is not a task, but to determine how task-like it is:

a) Does the activity engage learners' interest?
b) Is there a primary focus on meaning?
c) Is there an outcome?

d) Is success judged in terms of outcome? Is completion a priority?
e) Does the activity relate to real world activities?'

Just as there is a lack of consistency in the definition of a task, there is also variation in recommendations for implementation of TBLT (Ellis, 2009a; Skehan, 2003). Nunan's (2004) implementation, where tasks can be selected for the language they are expected to produce and involves pre-task controlled practice of target language, looks very much like a traditional PPP sequence of activities. He argues that it differs from traditional methods by presenting language in context, with a focus on linguistic elements taking place after language has been encountered in a communicative way, helping learners notice the connection between form and meaning. This view that traditional methods present language out of context, separated from meaning may be an unrealistic one (Swan, 2005).

At the other end is Long's (2014) implementation, where all tasks are selected based on a needs analysis, without the influence of target forms. This differs from naturalistic approaches in allowing explicit and implicit attention to form during tasks. This attention to form should always be reactive, which is an unrealistic proposal in some teaching contexts, for example, where the class size is large (Swan, 2005). Ellis (2001) points out that Long himself has violated his own recommendations and used planned focus on form in research, possibly because it is difficult to conduct research on reactive focus on form.

Willis and Willis (2007) recommend an implementation of TBLT that does not eschew traditional practices, but diverges from them much more than Nunan's implementation (2004). Pre-task activities are intended to prime students by introducing the topic, and seeding structures and vocabulary. Learners plan for the main task, which can be focused or unfocused, before carrying it out using whatever linguistic forms they have at their disposal. In a focused task students are unaware of the target forms it is intended to elicit. A post-task phase involves reporting to the class and focus on form activities. Task repetition is another recommended post-task activity. The sequence of classroom activities in this study most closely adhere to Willis and Willis' variety of TBLT.

2.2 Measuring task performance

Task performance can be measured in terms of meaning and form, with form being further divided between control and restructuring (Skehan, 1998). Researchers agree that complexity, accuracy and fluency (CAF) are useful measurements of second

language task performance (Palotti, 2009). Skehan (ibid) describes fluency as a measure of meaning reflecting a speaker's ability to produce language in real time; accuracy as a measure of control reflecting consolidation of forms and a speaker's ability to use language in accordance with target-like norms; and complexity as a measure of restructuring, reflecting a speaker's ability to take risks and be ambitious in their language use, using more elaborate structures and lexis to push towards inter-language development.

Levelt's (1989) model of processing is used by a number of researchers to theorise CAF (Skehan, 2009; Bygate, 2001; Ellis, 2009b; Bygate, & Samuda 2005). Summarising their descriptions of Levelt's model, concerning speech production, it involves three overlapping processes: the conceptualiser, the formulator, and the articulator. The conceptualiser accesses knowledge and plans a message, which is passed to the formulator, which accesses lexical and grammatical stores to form a message out of linguistic elements. The formulator passes the message to the articulator, which produces the spoken output. There is also a monitor, which checks the process before and after output.

L1 speakers are able to carry out these processes autonomously in parallel; however, limited working memory capacity results in L2 speakers having difficulty attending to the three processes of speech production in Levelt's model, instead focusing on some areas more than others. Different learners, depending on the circumstances will prioritise some processes over others resulting in a trade-off (Ellis, 2009b; Skehan, 2009), where one aspect of CAF wins out over another.

2.3 The effects of planning and repetition on task performance

Strategic planning takes place before a task and allows learners to plan what they will say and how they will say it, with access to the task materials, differentiating it from other pre-task activities (Ellis, 2005). Planning time allows learners to conceptualise and/or formulate their message before speaking. This can lead to more working memory being available during speaking, allowing the speaker to re-allocate attentional resources and improve some aspect(s) of their performance. Strategic planning tends to improve fluency and complexity, but not accuracy (Bygate & Samuda, 2005; Ellis, 2009b), suggesting strategic planning focuses on conceptualisation.

Online planning is defined by Yuan and Ellis (2003:6) as 'the process by which speakers attend carefully to the formulation stage during speech planning and engage in pre production and post production monitoring of their speech acts'. Online planning

takes place when there is no time pressure on speakers, allowing them to carefully formulate their utterances, focusing on form. Studies examining the effects of online planning show an improvement in accuracy (Ellis, 2009b).

As mentioned, Bygate and Samuda (2005) argue that task repetition can combine the benefits of both strategic and online planning to allow the speaker to improve their ability to conceptualise and formulate, resulting in an improvement in all aspects of performance. They define task repetition as 'repetitions of the same or slightly altered tasks – whether whole tasks, or parts of a task' (ibid: 43). Studies examining the effect of repetition show improvement in fluency and complexity with limited improvement in accuracy, and limited carry-over into other tasks (Ahmadian, 2011); however, Sheppard (2006, cited in Ellis, 2009b) shows that gains can be made in all aspects of performance and that gains can be carried over to other tasks if repetition is supported with attention to form treatment between performances. In a comparison of studies on the effects of planning on CAF, Ellis (ibid) found that none of the three studies focusing on task rehearsal took place in the classroom, and did not include pre and post-task phases.

3. Method

3.1 Participants

The research took place at private senior high school in Japan. The students were senior grade one, aged 15 to 16 years old. Ten intact classes of approximately 40 students each were split equally between two native English teachers (NETs) for Oral Communication classes. The researcher was one of the NETs and all data was collected from his classes. The students were required to pass an English test to enter the school, but above the base level required by the test, the students' proficiencies varied.

3.2 The task sequence

The research took place in the third and final term of the school year, which only included four lessons for one of the days of classes, so the syllabus was designed to be taught over four weeks, with the fourth lesson being used for an end of term test. It was decided that the focus of the syllabus for the term should be on supporting opinions.

The task used in the study, around which lessons were built, was chosen by the researcher and colleagues because it was seen as one which would encourage use of the target language being taught that term while providing an easily gradable task for the end of term test that could be prepared for in the three lessons available. The task was

one in which the students prioritise a number of items for survival in an inhospitable environment, adapted from alternatives to 'the NASA game' in Ur (1981:70). This task satisfies at least four of the five questions posed by Willis and Willis to determine how task-like an activity is: Whether it engaged learners' interest is unknown. There is a primary focus on meaning. There is an outcome upon which success is judged. Although the situation may be unlikely to occur for any students outside the classroom, the activity has interactional authenticity (Ellis, 2009a).

The study took place over three 50 minutes lessons. The first of the ten classes was chosen to pilot the lessons. After completing the lessons for the first time with the pilot class, the researcher was better able to deliver the lessons with less variation to the rest of the classes. Lesson plans and materials can be seen in Appendices A - D.

Lesson 1 is similar to an example task sequence recommended in Willis and Willis (2007). It includes priming, preparation and planning pre-tasks before the target task. The target task for this lesson, which will be referred to as T1, involved the students reporting which items they chose for a given survival situation in order of importance, with reasons. A pair discussion preceding the planning stage is likely to be the target task for other teachers using this task; however, in order to prepare for an assessed performance, the report, which would be considered post-task by some, was the target task. The time allowed for each stage was kept consistent; however time taken to confirm that instructions were understood varied between classes. Students were only allowed to make notes during the planning stage, and were not allowed to write full sentences. A listening task during the reporting stage prevented students from continuing to write ideas for their own report.

Lesson 2 consisted of a sequence of post-task activities, which included a repetition of T1 in pairs, instead of in front of the class, a listening activity using a model answer for T1, a matching activity, and a running dictation. The text and example sentences used in the post-tasks were based on common errors present in T1. A list of common errors was compiled by both native teachers and was based on in-class observation of all the students. It was not based on the recordings that were made by the researcher, as it would be unethical to give greater attention to the needs of the subjects of the research.

The third lesson followed the same plan as lesson 1, using a different location and set of items for the task. The target task from this lesson will be referred to as T2. Students were not allowed access to notes from the previous lessons during this lesson so that conditions would be consistent.

3.3 Analysis

Speakers from seven classes were recorded completing T1 and T2. Two of the nine classes were eliminated because their lesson 2 classes were shortened, meaning the matching task was done for homework. Some recordings were not usable due to speakers being inaudible and interruptions from classmates. In each class, four or five recordings were available for analysis. The recordings were transcribed and analysed using the Computerised Language Analysis (CLAN) program (MacWhinney, 2000). Pruned speech (Ortega, 1999) was transcribed, encoded in a way that would allow the CLAN program to analyse the transcripts to measure performance according to the specifications of this study. Transcripts can be seen in Appendix E.

3.4 Operationalisation of the measures

Although it is agreed that CAF are useful measurements of second language task performance, they can be measured in different ways (Housen & Kuiken, 2009). This section will explain the choices made in operationalising the each of the measures.

3.4.1 Syntactic complexity

The number of clauses per AS-unit was calculated to measure subordination for syntactic complexity. An AS-unit is defined as 'a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clauses associated with either' (Foster et al., 2000:365). The AS-unit was chosen because it is well defined and, unlike other units, does not omit any speech for analysis, such as ellipses. In a comparison by Norris and Ortega (2009) of 16 studies examining the effects of planning on CAF, they found that every study included a measure of complexity that measured subordination, with the AS-unit being favoured in the more recent studies.

3.4.2 Lexical variety

Skehan (2003, 2009) recommends including a measure of lexical performance as an additional measure to those used for CAF. The D measure, which measures lexical variety (Malvern & Richards, 2002), was calculated using the VocD sub program included in the CLAN program. The D measure was chosen over type-token ratio (TTR) because TTR is not reliable for comparing texts of varying lengths (Johansson, 2008; Malvern & Richards, ibid), whereas D is a 'generally acceptable' measure (Skehan, 2009:514), better suited to texts of different lengths (Lu, 2012; Skehan, 2003).

3.4.3 Accuracy

Accuracy was measured as the percentage of error-free clauses (Foster & Skehan 1996,1999; Yuan & Ellis 2003). This was chosen because it was easy to determine whether a clause contained an error, avoiding the unreliability of deciding how many errors occur in a measure such as errors/clause. Incorrect article use errors were ignored because 'useful variance should be maximised' (Foster & Skehan, 1996:304). It was felt that correct article use was outside the students' range.

3.4.4 Fluency

Fluency was measured as the number of words per minute (Nation, 1989) of pruned speech. Although syllables per minute is a commonly used measure of speech rate (Ellis, 2009b), words were chosen over syllables because it could easily be counted using CLAN software, reducing human error.

4. Results

Table 1 shows the descriptive statistics for the four measures for the two tasks. Syntactic complexity, accuracy, and fluency included 30 subjects, and all three showed increases in the means. Lexical variety included only 18 subjects, because some of the task performances did not produce large enough samples for the VocD program to calculate their D measure. Only subjects who provided two samples large enough were included. The means for this measure showed a decrease on the second task.

	Task	n	mean	SD
Syntactic Complexity:	T1	30	1.14	0.14
clauses / AS-unit	T2	30	1.23	0.2
Lexical Variety:	T1	18	34.78	13.4
D measure	T2	18	23.75	6.67
Accuracy:	T1	30	49.12	23.66
% error-free clauses	T2	30	56.1	19.39
Fluency:	T1	30	46.65	14
words / minute	T2	30	54.02	13.82

Table 1 Descriptive Statistics

To test the significance of the results, 2-tailed paired t-tests were conducted. The t-test values were calculated using Microsoft Excel 2011 for Mac and were verified using StatPlus for Excel. Table 2 shows the results of the tests. The gains in complexity and fluency were significant, with fluency gains being highly significant. Lexical variety showed a significant decrease. The t-test showed no significance in the gains in accuracy.

Table 3 shows the differences in scores for each of the measures for the students with the greatest increases and decreases on each measure. For fluency, the student with the second largest increase, S14, is also shown. Features of these students' performances will be discussed to help interpret the quantitative results. Transcripts of these students' performances can be seen in Appendix E.

	Sum of Differences	Sum of Squares of Differences	df	t	p-value 2 tailed
Syntactic Complexity:					
clauses / AS-unit	2.79	1.53	29	2.43	0.021
Lexical Variety:					
D Measure	-198.66	6049.55	17	-3.11	0.006
Accuracy:					
% error-free clauses	209.36	22903.77	29	1.41	0.170
Fluency:					
words / minute	221.25	4494.37	29	4.07	0.000

Table 1 Two-tailed t-test results, with alpha value 0.5

	Difference between tasks			
Student	Clauses / AS-unit	D Measure	% error-free clauses	words / minute
S27	0.58	-20.44	-24	16.07
S09	-0.3	n/a	4	-8.92
S23	0.22	-59.48	32	13.49
S25	0.3	14.18	62	-0.49
S22	0.3	n/a	-37	3.02
S24	0.3	-15.14	12	31.48
S14	-0.22	-3.76	-32	20
S03	0.13	-23.08	21	-13.62

Table 3: Results for students with greatest differences between tasks

5. Discussion

Because of the number of stages involved in the lessons in this study, and the lack of a control group it is difficult to identify which treatments, if any, affected performance. Each measure will be discussed in turn, taking into account features of the language used by the students in Table 3 when it is relevant.

5.1 Syntactic complexity

That syntactic complexity improved significantly suggests 'development in the inter-language system' of the students (Ellis, 2009b:475). The planning time provided in lesson 1 would have allowed the speakers to conceptualise and formulate their message for T1. As has been mentioned, learners tend to use strategic planning time to conceptualise, leading to a more syntactically complex product. That this measure improved on T2 suggests repetition can help to extend this effect of strategic planning.

Pressure of a public performance causes speakers to pay attention to accuracy (Skehan, 2009). The pressure to focus on formulation would have taken attentional resources away from conceptualisation. That complexity increased despite this, is possibly due to less formulation work required on the second task, it already having been done, allowing a more complex message to be conceptualised.

Student S27 used the structure "...can be used to..." for three of the five items producing eight clauses over three AS-units. This contributed to an improved syntactic complexity score, but the repeated use of the structure decreased the lexical variety score. As not a single student used this structure on the first task, it would seem that the

student noticed this during the post tasks. That he was able to produce multiple clause AS-units, which require fluency, shows signs of automation of the structure.

5.2 Lexical variety

There seems to be competition between lexical variety and structural elaboration. This is exemplified by students S27, who has already been discussed, and S03, who both used more complex structures during T2. However, they repeated the structures a number of times, reducing lexical variety. Three of S03's four multi-clausal AS-units followed the pattern '...if I ... I can...' which contributed to her reduced D measure. This task is not conducive to lexical variety however, with the product being quite short, structured, and potentially repetitive. This phenomenon may be less likely to occur in a more open task.

S23 increased in all areas except lexical variety, which decreased considerably. His first performance was very broken, omitting many function words. He used a high percentage of content words, which are not repeated as frequently as function words. The second performance was much improved in this respect, resulting in a lower lexical variety score. As an example, in T1 he spoke 'is' twice, and 'the' only once. On T2, he spoke those words six and ten times respectively.

With the largest gain in both lexical variety and accuracy, and also a slight gain in syntactic complexity it is surprising that S25 showed only a slight drop in fluency. On the first performance he lacked the verbs to go with the items, resulting in low accuracy and lexical variety scores. It is possible that for T2 he made better use of the priming stage of the lesson where the teacher seeded some vocabulary.

5.3 Accuracy

The lack of improvement in accuracy in contrast to the gains in syntactic complexity and fluency supports the trade-off found by Skehan and Foster (1997), where complexity and accuracy compete with each other. This is exemplified by S22 who took an exceptionally conservative approach for T1. She failed to give reasons for her choices, producing short simple utterances to indicate her choices. On the second task she was much more ambitious, resulting in a larger percentage of clauses with errors. Despite the reduced accuracy score, she actually produced more error free clauses on T2.

S03 and S23 both improved their article use, showing that this certainly was not beyond the range for all the students. Choosing to ignore incorrect article use errors may have been a misguided choice.

Although the two tasks involved the same interactions, the vocabulary involved was considerably different. Using the same vocabulary provides one of the stronger benefits of repeating a task (Bygate & Samuda, 2005). The students would have been limited in their ability to draw on experience from the first performance, reducing the workload on the conceptualiser less than in studies where the same task was repeated, which would account for a lack of improvement in accuracy.

5.4 Fluency

Significant improvements in fluency suggest students improved their ability to express meaning. Studies have shown that providing pre-task planning time consistently leads to improved fluency (Ellis, 2009b). As with syntactic complexity, it seems that task repetition can add to the benefits provided by strategic planning. S24 and S14 had the most significant fluency gains. They both seemed to have difficulty recalling their choices during T1, with long pauses, and S14 muttering *nandake* (What is it?). The lack of long pauses during T2 suggests they worked harder to memorise their choices, which may have contributed to a decrease in accuracy for S14. When task details have to be kept in working memory, accuracy suffers (Skehan, 1998).

5.5 Validity of the study

While this study may have ecological validity through its use of intact classes, and by following a recognised task-based teaching framework, the same factors reduce the scientific validity of the study. It is very difficult to create a scientific study in a classroom (Brown, 1997). With no control group, one cannot be sure that the performance gains were a result of the treatment.

Efforts were made to ensure that lessons were delivered in the same way each time. By having a pilot class, variation in input language was reduced. Timing of stages was also controlled. Nonetheless, with seven classes completing the task sequence two times there would obviously have been some variation. Recordings were made of the vocabulary elicitation stages at the beginning of lessons involving the target tasks. There was no noticeable difference between the classes for either lesson. However, there were considerable differences between the two lessons for each class. The elicitation stages for T2 included more vocabulary input from the teacher and a slightly greater variety of structures. This is a consequence of using a different scenario and items for the task.

The order in which the students performed the target tasks would have had an impact on their performance. Those who did not speak first were able to plan further while listening to others. They were not able to continue writing ideas thanks to the listening task; however, they were still able to plan mentally and also borrow ideas from other students. Although this violates the independence of observations (Brown, 1992), it is inevitable in an intact class. It can be seen as a pedagogic benefit of doing this kind of activity and is a typical feature of a TBLT framework.

6. Conclusion

This study measured the performance of students completing two similar target tasks. Each target task included pre-tasks that helped prepare the students. Between these two task sequences, students engaged in further rehearsal and conscious raising tasks. It has been argued that this is compatible with TBLT framework recommendations. However, there are many ways to implement TBLT, and it is likely that the same tasks would be carried out differently in other classrooms.

Syntactic complexity and fluency scores increased on the second task, while lexical variety decreased and accuracy showed no significant change. There seems to have been a trade-off between accuracy and complexity, which has been shown in other studies, and also a trade-off between syntactic complexity and lexical variety. Both of these trade-offs may have been fostered by the design of the tasks, with the closed nature and repetitiveness leading to a decrease in lexical variety, and increased difficulty of the second task leading to a lack of improvement in accuracy.

It is difficult to compare to other published studies due to the differences between the tasks being too great, the presence of additional treatment through pre and post-tasks, and the dependent variables for accuracy and fluency differing from those used by others. Despite this lack of generalisability, the results are positive for TBLT in general. The improvement in quality of performance on a different task hints at acquisition taking place, and signs that the conscious raising tasks contributed to lexicalisation of new forms are present. This study indicates that TBLT can lead to inter-language development with no expense to accuracy.

References

Ahmadian, M.J. (2011) The effect of “massed” task repetitions on complexity, accuracy and fluency: does it transfer to a new task? The Language Learning Journal, 39(3), pp.269–280.

Brown, J.D., (1992) Statistics as a Foreign Language—Part 2: More Things to Consider in Reading Statistical Language Studies*. TESOL Quarterly, 26(4), pp.629–664.

Brown, J.D. (1997) Designing a language study. In Griffee, D. T. & Nunan, D. (eds) Classroom teachers and classroom research. Tokyo: The Japan Association for Language Teaching, pp. 55-70.

Bygate, M. (2001) Effects of task repetition on the structure and control of oral language. Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing, pp.21–48.

Bygate, M. & Samuda, V. (2005) Integrative planning through the use of task repetition. In Ellis, R. (ed) Planning and Task Performance in a Second Language, pp. 37-74. Philadelphia: John Benjamins, Publishing Company. Ebrary. Web. 23. Jan. 2014.

Edge, J. & Garton, S. (2009) From Experience to Knowledge in ELT. Oxford: OUP.

Ellis, R. (2001) Introduction: Investigating Form-Focused instruction. Language Learning.

Ellis, R. (2005) Planning and task-based performance: Theory and research. In Ellis, R. (ed) Planning and Task Performance in a Second Language, pp. 3-34. Philadelphia: John Benjamins, Publishing Company. Ebrary. Web. 23. Jan. 2014.

Ellis, R. (2009a) Task-based language teaching: sorting out the misunderstandings. International Journal of Applied Linguistics, 19(3), pp.221–246.

Ellis, R. (2009b) The Differential Effects of Three Types of Task Planning on the Fluency, Complexity, and Accuracy in L2 Oral Production. Applied linguistics, 30(4), pp.474–509.

Foster, P. (1996) 'Doing the task better: How planning time influences students' performance' in J. Willis and D. Willis (eds): Challenge and Change in Language Teaching. Heinemann.

Foster, P. (1999) Task-based learning and pedagogy. ELT Journal, 53(1), pp.69–70.

Foster, P., Tonkyn, A. and Wigglesworth, G. (2000) Measuring spoken language: A unit for all reasons. Applied linguistics, 21(3), pp.354–375.

Foster, P. & Skehan, P. (1996) The influence of planning and task type on second language performance. Studies in Second language acquisition, 18, pp.299–324.

Foster, P. & Skehan, P. (1999) The influence of source of planning and focus of planning on task-based performance. Language Teaching Research, 3(3), pp.215–247.

Housen, A. & Kuiken, F. (2009) Complexity, Accuracy, and Fluency in Second Language Acquisition. Applied linguistics, 30(4), pp.461–473.

Johansson, V. (2008) Lexical diversity and lexical density in speech and writing: A developmental perspective. Lund Working Papers in Linguistics, 53, pp.61–79.

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

Lu, X. (2012) The relationship of lexical richness to the quality of esl learners' oral narratives. The Modern Language Journal, 96(2), pp.190–208.

Lynch, T. & Maclean, J. (2000) Exploring the benefits of task repetition and recycling for classroom language learning. Language Teaching Research, 4(3), pp.221–250.

Long, M. (1985) A role for instruction in second language acquisition: task-based language teaching. In K. Hyltenstam and M. Pienemann (eds.), Modelling and assessing second language acquisition. Clevedon: Multilingual Matters.

Long, M. (2014) Second Language Acquisition and Task-Based Language Teaching (1). Somerset, US: Wiley-Blackwell. Retrieved from http://www.ebrary.com

MacWhinney, B. (2000) The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates

Malvern, D. & Richards, B. (2002) Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. Language Testing, 19(1), pp.85–104.

Nation, P. (1989) Improving speaking fluency. System, 17(3), pp.377–384.

Nunan, D. (2004). Task-Based Language Teaching. Cambridge: Cambridge University Press.

Norris, J.M. & Ortega, L. (2009) Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. Applied linguistics, 30(4), pp.555–578.

Ortega, L. (1999) Planning and focus on form in L2 oral performance. Studies in Second language acquisition, 21(1), pp.109–148.
Pallotti, G., (2009) CAF: Defining, Refining and Differentiating Constructs. Applied linguistics, 30(4), pp.590–601.
Sheen, R. (1994) A Critical Analysis of the Advocacy of the Task-Based Syllabus. TESOL Quarterly, 28(1), pp.127–151.
Skehan, P. (1996) A framework for the implementation of task-based instruction. Applied linguistics, 17(1), pp.38–62.
Skehan, P. (1998) Task-based instruction. Annual Review of Applied Linguistics, 18, pp.268–286.
Skehan, P. (2003) Task-based instruction. Language Teaching, 36(1), pp.1–14.
Skehan, P. (2009) Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. Applied linguistics, 30(4), pp.510–532.
Skehan, P. & Foster, P. (1997) Task type and task processing conditions as influences on foreign language performance. Language Teaching Research, 1(3), pp.185–211.
Swan, M. (2005). Legislation by Hypothesis: The Case of Task-Based Instruction. Applied Linguistics, 26(3), 376–401. http://doi.org/10.1093/applin/ami013
Ur, P. (1981) Discussions that Work: Task-centred Fluency Practice, Cambridge: Cambridge University Press
Van den Branden, K. (2016) Task-based language teaching. The Routledge Handbook of English http://doi.org/10.1017/CBO9780511667282
Yuan, F. & Ellis, R. (2003) The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. Applied linguistics, 24(1), pp.1–27.
Willis, D. & Willis, J. (2007) Doing Task-Based Teaching: A practical guide to task-based teaching for ELT training courses and practising teachers, OUP, Oxford, Kindle Edition

---

Received on December 7, 2016.

Appendix A: Plans for Lessons 1 and 2

Lesson 1 Plan – priorities – survival games - Sahara Desert

Greet class, check attendance **5 min**

Priming: Elicit ideas **8 - 10 min**

- Teacher leads brainstorm of Sahara desert, writing student ideas on board
- Give students item lists
- Listen and repeat item names
- Elicit ideas for using items in the Sahara, seeding vocab and structures if students don't provide them

What can you use [item] for? What can [item] be used for? What can you do with [item]? What is [item] used for? Why do you need [item] in the Sahara?	you can [ ] with [item] it can be used to… if we … we can … if we don't … we will …	Africa, desert, dry, hot, cold at night, no water, dangerous animals drink, climb eat, light a fire protect ourselves from sunlight keep warm, cook,

Preparation: Individuals prioritise **5 min**

- Students individually put items in order of importance

Preparation: Pairs reduce list to five **5 Min**

- Students in pairs decide on 5 items and prioritise

Planning: Individuals prepare speech **5 min**

- Students individually prepare a report on the decision they made and why each decision was made
- make notes that wont be available during presentation
- students should make notes only – not write a speech

Target task: Report decision **10 – 20 min**

- students stand at front of class and give report
- while student is speaking, others listen and complete form

Lesson 2 Sahara Desert Follow-up Plan

Practice

- Students review choices and notes from previous week for a few minutes
- Students find a new partner and tell them about their choices

Listening

- Teacher gives speech, students write items in order in notebooks, make notes on reasons
- Teacher asks students for order and reasons
- Teacher echoes student ideas, correcting their errors and using language from listening text to rephrase broken utterances

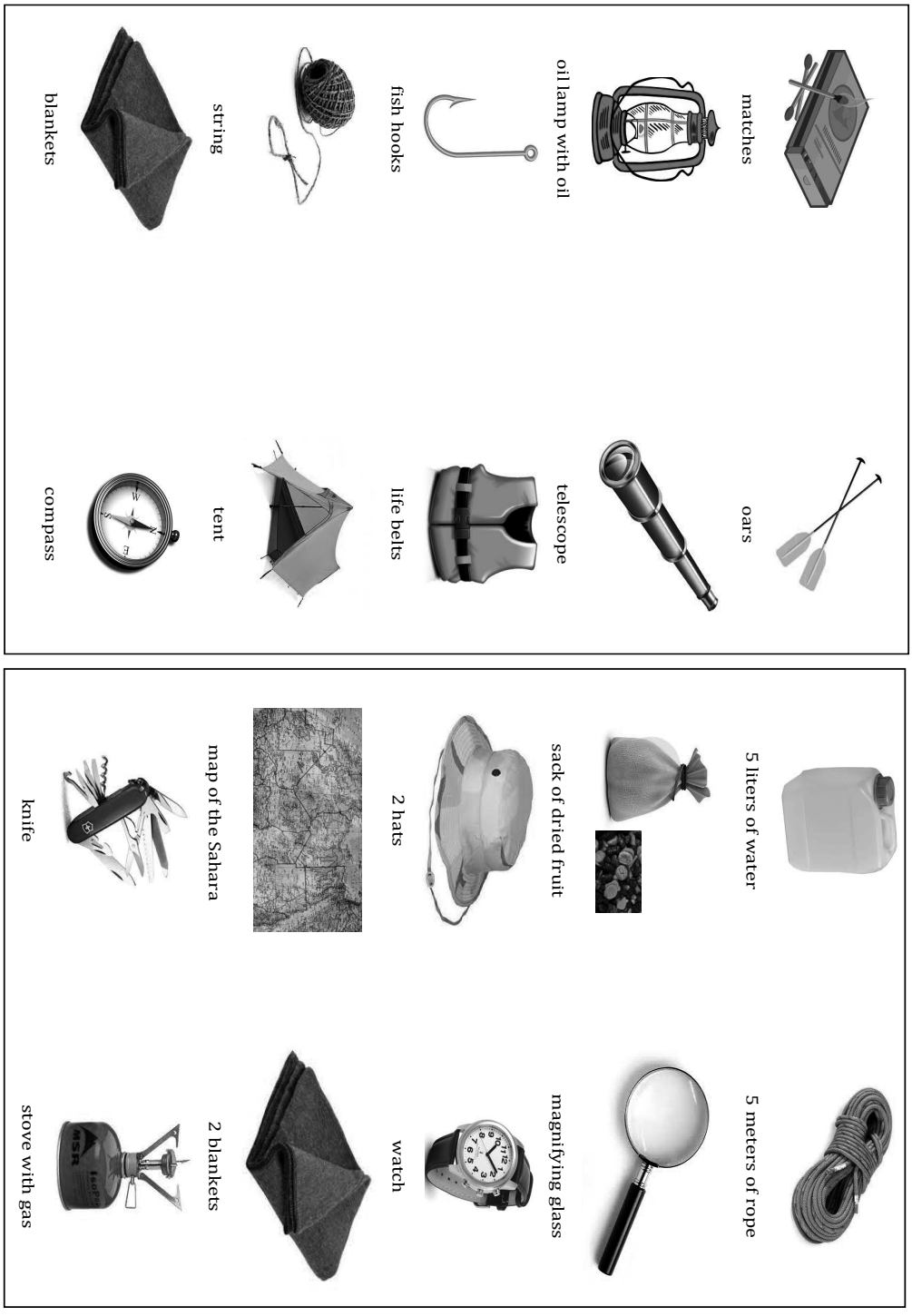
Running Dictation

- Make groups of 4
- 1 student from each group approaches teachers desk to read the speech
- student returns to team, sits down, tells groups what they read
- another student from each group repeats
- continue until groups have accurately completed the task
- first team to finish correctly wins
- give students the print when finished

Match Sentences

- Pairs match items to sentences

Appendix B: Items for T1 and T2



Appendix C: Pre-task Worksheet for T1

Survival: Sahara Desert

You and a friend are in the Sahara desert. You need to travel 100km to the nearest village.

- Choose which items are most important.
- Number them 1-10
- 1 is the most important. 10 is the least important.

map of the Sahara ___
5 meters of rope ___
watch ___
knife ___
2 blankets ___
sack of dried fruit ___
stove with gas ___
magnifying glass ___
2 hats ___
5 liters of water ___

- With a partner, choose 5 items to take. Number 1 is the most important. Number 5 is the least important.

1. _____
2. _____
3. _____
4. _____
5. _____

- Prepare a short speech for the class.
- Tell the class which items are most important **and your reasons**.
- You can make notes, but you can't use them when you speak to the class.

Appendix D: Post-Task Materials

Sahara Desert: Example Sentences

Match the sentences to the items. There may be more than 1 correct answer.

a) map of the Sahara
b) rope
c) watch
d) knife
e) blanket
f) sack of dried fruit
g) stove with gas
h) magnifying glass
i) hat
j) water

1. We will be hungry so we need this for energy. *f*
2. We need it to protect ourselves from dangerous animals. _____
3. We can't live without this. _____
4. It can be used as a compass. _____
5. It is very hot and dry so we need this or we will die. _____
6. It can be used to cut wood or rope. _____
7. It is very cold at night, so we need it to keep warm. _____
8. The Sahara desert is very big so we need it to find our way. _____
9. We should take this in case we need to cook something. _____
10. It can be used to carry things or climb cliffs. _____

Model Answer for Listening and Running Dictation

The most important item is the 5 litres of water. If we don't drink water we will die. There isn't any water in the desert, so we need to carry a lot.

The second most important items are the hats. It's very hot and sunny in the Sahara, so we need hats to protect ourselves from the sun.

The next most important item is the sack of dried fruit. Eating the fruit gives us energy. If we don't have energy we can't walk.

The next most important item is the magnifying glass. It can be used to make fire. Fire can keep us warm at night and keep away dangerous animals.

The knife is the final item. It can be used for many things such as cutting. If we find wood we will need to cut it to make a fire.

Appendix E: Transcripts

Key

Clauses are marked with [^c] at the end and take a new line

Clause errors are marked with [*]

The end of AS-units are marked with a full stop

S27 (pruned and coded file unavailable)

T1	**T2**
I think the most important item is ahh 5 lit- 5 litres of water … because it's too heavy to carry but we can't live without water. Second item is *etto* ah! map of Sahara ah … because ahh [] its its ah its important to know where where we stay or where we go Next Third item is dried fruits because … food is ah so important Next, fourth item is 2 hats because sun- sunlight is [ ] harmful And last five um fifth item is because its usual … it … I can use it … anytime uh anywhere	I think the most important item is oars. It can be used to row a boat or and move ahh somewhere. And the second most important item is life jackets. It needs for emergency situation. And third third impor- most important item is compass. It can be used to find direction and go there. And fourth most important item is tent. Its needs it needs sleep ahh it needs needs to sleep to stay at night Finally the most important item is blankets. It it needs ahh keep It can be used to warm keep warm.

S09

*S09: one is five litre of water [^c] [*] . *S09: I think I thirsty [^c] [*] . *S09: two is two hats [^c] [*] . *S09: because I think [^c] I become hot [^c] [*] . *S09: three is snack of dried fruit [^c] . *S09: I think [^c] I need something to eat [^c] . *S09: three is two blankets [^c] [*] . *S09: I think it is very cold [^c] . *S09: five is watch [^c] [*] . *S09: because I think [^c] I want to read time [^c] [*] .	*S09: first is oars [^c] . *S09: it can row a boat [^c] [*] . *S09: second is compass [^c] . *S09: it can know way [^c] [*] . *S09: third is telescope [^c] . *S09: because it can find island [^c] [*] . *S09: fourth is fish hooks [^c] . *S09: because it can catch fish [^c] [*] . *S09: finally is blankets [^c] [*] . *S09: because it is cold [^c] .

S23

*S23: I think [^c] the most important item is five litres of water [^c] . *S23: because drink it [^c] [*] . *S23: and cool down hot body [^c] [*] . *S23: next dried fruits [^c] [*] . *S23: because eat it [^c] [*] . *S23: so important for human [^c] [*] . *S23: third is map of Sahara [^c] . *S23: no map no life [^c] [*] . *S23: next two blankets [^c] [*] . *S23: so hold warm at night [^c] [*] . *S23: and u_v cut [^c] [*] . *S23: next watch [^c] [*] . *S23: so look at time [^c] [*] .	*S23: the most important item is the life jackets [^c]. *S23: because save the life [^c] [*] and so don't die [^c] [*] . *S23: the second important item is oars [^c] . *S23: because move boats [^c] [*] . *S23: next the most important item is compass [^c] . *S23: so because check the direction [^c] . *S23: next the most important item is blankets [^c] so because hold warm [^c] [*] . *S23: and so sea is very cold [^c] [*] . *S23: fish hooks is the final item [^c] . *S23: so catch the fish [^c] [*] and eat the fish [^c] .

S25

*S25: first I need to five litres of water [^c] [*] .	*S25: the most important item is life jacket [^c] .
*S25: because don't want to die [^c] [*] .	*S25: because I don't want to die [^c] .
*S25: second sack of dried fruit [^c] [*] .	*S25: second item is fish hooks [^c] .
*S25: because don't want to die [^c] [*] .	*S25: second item is oars [^c]
*S25: third I need to map of Sahara [^c] [*] .	because move a boat [^c] [*] .
*S25: because I need to map [^c] [*] .	*S25: and I want to go home [^c] .
*S25: fourth I need to watch [^c] [*] .	*S25: third item is fish hooks because fishing [^c] [*] .
*S25: because to know the direction [^c] [*] .	*S25: and eat fish [^c] .
*S25: fifth I need to magnifying glass [^c] [*] .	*S25: fourth important item is oil lamp with oil [^c] .
*S25: because I need to fire [^c] [*] .	*S25: because guide around me [^c] [*] .
	*S25: five important item is compass [^c] [*] because know a direction [^c] [*] .

S22

*S22: I need five litres of water [^c] .	*S22: the most important item is oars [^c] .
*S22: because water is important for a human [^c] .	*S22: because go far [^c] [*] .
*S22: next I want sack of dried fruit [^c] .	*S22: second important item is life jacket [^c] .
*S22: next map [^c] [*] .	*S22: if boat broken [^c] [*]
*S22: next I want a knife [^c] .	I can float [^c] .
*S22: next I need a stove with gas [^c] .	*S22: next item is compass [^c] .
	*S22: because I can know [^c]
	how to go Japan [^c] [*] .
	*S22: next item is telescope [^c] .
	*S22: because find big boat [^c] [*]
	help me [^c] [*] .
	*S22: next is match [^c] [*] .
	*S22: because make a fire [^c] [*] .

S24

*S24: the most items I need [^c] [*]	*S24: the most important item I think [^c]
is the water [^c] .	is the life jackets [^c] .
*S24: people may die [^c]	*S24: the life jackets is able [^c] [*]
if they don't drink water for long times [^c] [*] .	to float on the sea [^c]
*S24: the second item is the map of Sahara [^c] .	and then we can wait for the help [^c] .
*S24: people may don't know [^c] [*]	*S24: the second is the telescope [^c] .
where we are [^c] .	*S24: we can look far [^c]
*S24: and don't know the direction [^c] [*] .	and find the island and the boats [^c] .
*S24: the third item is the sack of dried fruit [^c] .	*S24: the third item is the oars [^c] .
*S24: the fourth is the watch [^c] .	*S24: we can row the oars [^c] [*]
*S24: we want [^c]	and move the boats [^c] .
to know the time [^c] .	*S24: the fourth is the oil lamp with oil [^c] .
*S24: the last is two blankets [^c] .	*S24: in night ocean is dark [^c]
*S24: we can keep our body warm [^c] [*] .	so the oil lamp with oil is useful [^c] [*] .
	*S24: finally is the compass [^c] [*] .
	*S24: we can know [^c]
	which direction to move the boats [^c] .

S14

*S14: the most of important is five litres of water [^c] [*] .	*S14: the most important item is oars [^c] .
*S14: because it is important to live [^c] .	*S14: it can row the boat [^c] [*] .
*S14: second is map of the Sahara [^c] .	*S14: the next important item is life jacket [^c] .
*S14: third is sack of dried fruits [^c]	*S14: it can float the sea [^c] [*] .
because it is important to live [^c] .	*S14: the third important item is blanket [^c] .
*S14: third is two hats [^c] .	*S14: it can to keep warm [^c] [*] .
*S14: because daytime is very hot [^c] .	*S14: the second important item is telescope [^c] .
*S14: fourth is two blankets [^c]	*S14: it can see far [^c] [*] .
because night is very cold [^c] .	*S14: the last important item is oil lamp with oil [^c] .
*S14: five is stove with gas [^c] [*] .	*S14: it can light [^c] [*] .
*S14: because it is very useful [^c] .	

S03

*S03: I think [^c]
the most important things is five litres of water [^c] [*] .
*S03: because I need to something [^c] [*]
to drink to live [^c] .
*S03: next important things is sack of dried fruit [^c] [*] .
*S03: because if I cant eat [^c]
I will die [^c] .
*S03: next is map of Sahara [^c] .
*S03: because Sahara desert is very huge [^c] .
*S03: and I can know the way to back [^c] [*] .
*S03: next is stove with gas [^c] .
*S03: because fire can use many things [^c] [*] .
*S03: last things is two hats [^c] [*] .
*S03: because I can protect the sunshine [^c] [*] .

*S03: I think [^c]
the most important item is oars [^c] .
*S03: because I can go through my way [^c] [*] .
*S03: the next item is fish hooks [^c] .
*S03: because I can catch the fishes [^c] .
*S03: the next important item is blankets [^c] .
*S03: because if I have blankets [^c]
I can keep the warm [^c] [*] .
*S03: the next item is matches [^c] .
*S03: because if I have matches [^c]
I can make fire [^c] .
*S03: last item is strings [^c] .
*S03: because if I tie fish hooks [^c] [*]
I can catch the fishes [^c] [*] .