# The Effect of Foregrounding and Backgrounding on Grammatical Complexity and Lexical Diversity of Oral Production of EFL Learners in a Foreign Language Learning Context

**Seyyed Mohammad Alavi\***
*Associate Professor, University of Tehran*
**Davood Borzabadi Farahani**
*Assistant Professor, University of Tehran*
**Javad Muhammadi Savadroodbari**
*MA, University of Tehran*

## Abstract

This paper examined an influence of two different narrative tasks on the grammatical complexity and lexical diversity of an oral performance of second-language (L2) learners. The participants (N=90) were presented with two tasks with six pictures in each. One of the tasks contained both foregrounding and backgrounding events on it, but the second one had only the backgrounding event. Backgrounding events language refer to the main theme of the story, and foregrounding events refer to the supportive materials that elaborate on the side events taking place in the task. The complexity and diversity of the language in the oral production of the two comparable groups who were presented to foregrounding and backgrounding events were examined in a between subject design. The results showed that L2 performance was affected by the type of the task. Syntactic complexity and lexical diversity were increased by the participants who were given the task with both of the events on them. We reason that the task with backgrounding event on it leads participants into attempting more to talk about the events happening in the background, hence producing more complex and at the same time more diverse

**Keywords:** task-based language teaching, narrative task, grammatical complexity, lexical diversity, foregrounding event, backgrounding event

## 1. Introduction

Grammatical complexity plays an important role in learners' narrative productions. As there are two kinds of vocabularies stored in the learners' mind, passive and active (Melka, 1997), the same claim can stand true for grammar. Some language learners can perform really well on paper tests but are unable to use their grammatical knowledge in their productions. This story is even worse for lexicon. As Melka (1997) shrewdly stated, language learners suffer from their inability to use the large reservoir of vocabulary items stored in their mind.   An attempt has been made to use such tasks that provoke learners to use their stored knowledge and make the best use of them in their narrative production, i.e., transforming their passive grammatical knowledge and lexical items into active grammatical knowledge and lexical items.

## 2. Literature Review

Over the past thirty years, the word 'task' was associated with various definitions from various researchers who have been investigating various aspects of task on various participants in various settings from various perspectives. Furthermore, other similar terms such as "activity", "exercise", and "procedure" pose another difficulty because these terms have often been loosely defined, and some researchers have confused tasks with these terms; as a result of which they have been used interchangeably. To muddy the issue further, some of these definitions have been either too broad to provide any insights into the process of language teaching or too narrow to capture all the applications for which it might be wanted (Ellis, 2003).

Tasks have been making consistent inroads into language teaching, and they constitute a sine qua non of language teaching now. Tasks are mainly associated with communicatively oriented language classes, and language teachers cannot dispense with them. They are so important because they create favorable conditions in language classrooms. Tasks have been supported by theories from mainstream education, second language acquisition, second language research and English for specific purposes. As Ellis (2003a) stated, "proposals for task-based syllabuses then arose out of recognition that it was not possible to specify what a learner would learn in linguistic terms" (p. 208).

From language teaching perspective, tasks are theoretically claimed to be compatible with learners' in-built syllabus and task-based syllabuses conform to acquisitional processes in which certain

grammatical structures should be learned in certain predictable orders at certain stages. Such a claim has long been empirically demonstrated in first language acquisition (Brown & Bellgui, 1964), and task-based instruction is mainly premised on such a view (Long & Crookes, 1992).

From a second language acquisition and research viewpoint, tasks engage learners deeply in the process of language learning either implicitly or explicitly (Long, 1985; Prabhu, 1987), and task-based instruction matches with the cognitive processes involved in second language acquisition and promotes acquisition. Viewed this way, tasks provide a reasonable challenge and will be cognitively involving and motivating (Ellis, 2003a). Task performance also brings about interlanguage change prompting learners to notice and retain information about the target language they use (Swain, 1995). "From this angle, exploring SLA through task performance is interesting for its own sake and is not necessarily directed toward pedagogic applications" (Tavakoli & Foster, 2011, p. 38).

From an ESP (English for specific purposes) angle, tasks are ideal units for specifying the content of specific courses (Ellis, 2003a) because specific groups of learners have different needs, and tasks serve learners' specific needs and closely mirror what they need to do with the language so as to learn. This is usually done through a needs-analysis in which learners' needs are first identified and appropriate tasks are accordingly designed (Long & Crookes, 1992).

From a pedagogical standpoint, specifically related to language teaching and the purposes of the present study, tasks provide "empirically sound principles for classroom materials design" (Tavakoli & Foster, 2011, p. 38), which implies that course syllabus designers and language teachers do not have to rely on their intuition to specify materials and activities for their courses and students. By choosing appropriate tasks, teachers can rightly guide their students to focus their attention on the particularly problematic areas of language.

Since the present study aims to examine oral production tasks in a classroom context, the definitions of 'tasks' suggested by Tavakoli and Foster (2011), seem to be appropriate to follow. Tavakoli and Foster define a task as "anything that classroom language learners do when focusing their attention primarily on what they want to say to others or what others are trying to say to them" (p. 39). This definition serves our purpose in the present study. On the one hand, the primary focus is on meaning when the students are involved in either meaning-focused or form-focused activities; the students are supposed to watch some pictures

and recount stories about those pictures. On the other hand, forms are also given priority when the need arises.

*2.1. The Triad of Complexity, Accuracy, and Fluency*

The notions of complexity, accuracy, and fluency (CAF) have been used in a number of studies on the acquisition and use of a second language, although they do not constitute a theory or a research program in themselves. According to Skehan (1998), these three aspects of performance have to be distinguished due to their vulnerability to the types of processing learners adopt. They are mostly employed as the dependent variables to assess variation with respect to independent variable(s). These notions have been defined differently in each study.

It is not, however, our intention in this article to explore this triad. We are only interested in exploring the effects of foregrounding and backgrounding events on grammatical complexity and lexical diversity. Complexity is certainly the most problematic construct of the CAF triad because of its complex nature. In one sense, the same term, complexity, is used to refer to the properties of both tasks and language. In another sense, it "concerns the elaboration or ambition of the language that is produced" (Skehan, 1996, p. 22). In yet a different sense, it has also been defined as "the extent to which the language produced in performing a task is elaborate and varied" (Ellis, 2003, p.  340).

L2 complexity can be interpreted in two different ways: (a) *cognitive complexity* and (b) *linguistic complexity* (Skehan, 1996). They both refer to properties of language or its (sub)systems (Housen & Kuiken, 2009). Cognitive complexity is interpreted from the perspective of language user, whereas linguistic complexity is interpreted from the perspective of the L2 system or its feature. According to the same researchers, cognitive complexity is what we call as difficulty, and refers to the relative difficulty with which language features are processed in L2 performance. It is determined both by subjective and objective factors. The latter is a good example of linguistic difficulty. Skehan (1996) argued that a task is *objectively difficult* when it is hard in nature, and it is different from being *subjectively difficult,* which may be hard for one but not for another. Even restricting the use of complexity to performance description does not suffice. Still complexity can be applied to different aspects of language and communication. Lexical, interactional, propositional, and various kinds of grammatical complexity are some of the aspects mentioned in Ellis (2003).

Is it appropriate to consider all these aspects as dimensions of the same construct? Yet, for each of the constructs mentioned, the word *complex* may have a different meaning. One is merely structural,

"composed of two or more parts" (Merriam-Webster, 2005). Another structural definition of complexity refers to variety, to the existence of multiple alternatives. This is the most common sense of definition in lexical complexity. Some authors like Ellis (2003), identify *complex* with 'acquired late'. The second meaning of complexity is temporal. Ellis (2003), defined it as "the capacity to use more advanced language" (p. 3), although his definition did not reach unanimous agreement. In line with Ellis, Skehan (1996) defined complexity or 'range', to use his term, as:

> the capacity to use more advanced language, with the possibility that such language may not be controlled so effectively. This may also involve a greater willingness to take risks, and use fewer controlled language subsystems. This area is also taken to correlate with a greater likelihood of restructuring, that is, change the development in the interlanguage system. (p. 45)

Acquiring something late does not necessarily mean that it is complex. It is acquired late, maybe because it is infrequent. Therefore, time-related notions should be separated from other senses of complexity, because they are not included in dictionary definitions of it. Foster and Tavakoli (2009) have the same opinion about the definition of complexity. They also believe that complexity is the most problematic construct of all because of its polysemous nature.

Subordination is the most common way of measuring complexity (e.g. number of clauses per T-unit or C-unit). In some studies (see Malvern & Richards, 2009), lexical complexity has been assessed by means of type-token ratio. In Yuan and Ellis (2003), complexity was measured at three levels: (1) syntactic complexity, (2) syntactic variety, and (3) Mean Segmental Type-Token Ratio (MSTTR). As there were not so many elided utterances in the task performance, T-units rather than C-units were used to measure syntactic complexity. Syntactic complexity has been defined as "the ratio of clauses to T-units in the participants' production" (Yuan & Ellis, 2003, p. 13). The total number of different grammatical verb forms was used as the basis for measuring syntactic variety and for the MSTTR, the narratives taken from the participants were divided into segments of 40 words. The total number of different words in each segment was divided by all the words in the segment to calculate MSTTR. According to Malvern and Richards (2009), this way of measuring complexity has some drawbacks which will be elaborated on hereafter.

Foster, Tonkyn, and Wigglesworth (2000) proposed the Analysis of Speech Unit (AS-unit) to measure syntactic complexity in spoken

language. The unit that they used was fundamentally a syntactic unit. The two main reasons that they used this unit as a valid one are:

> First, the studies of pausing in native-speaker speech suggest that syntactic units are genuine units of planning, since many pauses occur at syntactic unit boundaries, and specially clause boundaries. Secondly, our definition allows analysis of speech units which are greater than a single clause since there is evidence from intonation and pause features that speakers may plan multi-clause units. (p. 365)

Foster et al. (2000, p. 365) defined AS-unit as: "… a single speaker's utterance consisting of *an independent clause, or sub-clausal unit,* together with any *subordinate clause(s)* associated with either" . According to this definition, "an independent clause will be minimally a clause including a finite verb" (Foster et al., 2000, p.365). Unlike T-unit and better than C-unit, AS-unit allows us to include independent sub-clausal units, which are an indispensable part of everyday speech. According to Foster et al., an independent sub-clausal unit consists of either one or more phrases that can be expanded to a full clause by providing the omitted parts. Minor utterances which are defined as a type of "irregular sentences" or "Non-sentences", are among independent sub-clausal too, as Quirk, Greenbaum, Leech and Svartvik (1985, p. 835-53) put it.

A 'subordinate clause' on the other hand is defined as a finite or non-finite verb element plus at least one other clause element (Foster, Tonkyn, & Wigglesworth, 2000). False starts, repetitions and self-corrections may cause some problems in relation to the unit. Since they frequently occur in oral data and specifically in second language data, a clear procedure is needed to handle them. False starts are utterances which are begun and left altogether or reformulated in some way (Foster et al., 2000). According to Foster et al., in cases where an AS-unit is produced before the message is abandoned, that part of the utterance which meets the AS-unit criteria will be counted as an AS-unit, with the rest being considered as a false start. Repetitions happen in oral speech more than false starts or self-corrections. The important point is that repetitions which are used for rhetorical effect should be distinguished from those that indicate dysfluency.

When the speaker identifies his error, he will correct himself immediately after the production or stops and reformulates the speech. When this happens, the final version is counted and the previous version is excluded. "Topicalized noun phrases belong to the unit of which they are the topic" (Foster, et al., 2000, p.369). The noun phrases that are

separated from the AS-unit by falling intonation and a pause which is equal or more than 0.5 second will be considered as AS-units. If a coherent analysis is going to be done, any kind of exclusion must be done according to some specified principles. Foster et al. provided criteria for the principled exclusion of data, when it is necessary for the coherent analysis. They proposed three different types of spoken language data and called them levels 1, 2, and 3. In level 1, which is used for a full analysis of data, everything except untranscribable data should be included. In level 2, which is used for highly interactional data, one-word minor utterances are excluded. Level 3 is for special cases where analysis of non-fragmentary AS-unit is needed. In this level, the items that were removed in level 2 are also omitted as well as v-less elliptical As-unit involving ellipses of elements of the interlocutor's speech. In level 3, one or two word greeting and closure, and As-units involving substitution of clause, predicate, or predication level units of interlocutor's speech are excluded, too. In the present study, level 1 was adopted.

## 2.2. *Lexical Diversity*

Diversity in general and lexical diversity in particular have been used in many educational and, even linguistic studies. The variety of active vocabulary used by speakers or writers together with lexical density, accuracy of expression and lack of errors of lexical choice can be a rough definition for lexical diversity (Read, 2000). Measures of vocabulary diversity have been used in a wide range of linguistic and educational studies (Malvern & Richards, 2002; Vermeer, 2000; to name a few). As mentioned in many studies, lexical diversity is difficult to quantify reliably. The best known measurement is Type-Token Ratio (TTR), which is inherently problematic. It takes into account the comparison between the number of different words (types) to the total number of words (tokens). As the tokens in a sample increase, the TTRs decrease (Chotlos, 1994; Richards, 1987). It is, therefore, invalid to compare the TTRs derived from samples with different sizes. In their study of oral interview, Malvern and Richards (2002) solved the problem of calculating lexical diversity by using Mean Segmental Type-Token Ration (MSTTR). Malvern and Richards (2002) describe MSTTR as "the average TTR for successive segments of text containing a standard number of words tokens" (p. 35).

Richards and Malvern (2002) developed a new measure of lexical diversity that overcomes the disadvantages in measuring lexical diversity. They claimed that:

> The solution is based on the observation that the way in which TTR falls with sample size is systematic and that this means that

the probability of new vocabulary being introduced into longer and longer samples of speech or writing can be mathematically modeled. (p. 89)

Their model is a mathematical equation that relates TTR to token size (N) in terms of a third parameter referred to as 'D':

$$TTR = \frac{D}{N}\left[\left(1 + 2\,\frac{N}{D}\right) - 1\right]$$

### 2.3. The Research Study

The prime objective of the present study was to investigate the role of foreground and background on non-native language learners' production, especially in terms of grammatical complexity and lexical diversity. As Hamp-Lyons mentioned, tasks in general and narrative tasks in particular are inherently problematic, and it is the researcher's responsibility to unpack the complexities of how tasks should be addressed so as to help L2 learners benefit from them. One feature of narrative tasks which makes them demanding for L2 learners and affects both production and performance is the amount of information present in the task. The degree of foregrounding and backgrounding in narrative tasks seems to be an under-researched issue. Specifically, how foregrounding and backgrounding events affect grammatical complexity and lexical diversity is an issue which is yet to be empirically investigated. An effort is being made in this study to fill this gap in the literature.

## 3. Method

### 3.1. Participants

These were 90 learners of English as a foreign language in a private institute. They were both females and males equal in number, ranging between 17 to 27 years old. They were given the placement test, and all had been scored as 'intermediate' language learners. The participants were randomly placed into two groups. Each group received a different task, adopted from Tavakoli and Foster (2011), one with foregrounding events, and the other with both foregrounding and backgrounding events.

### 3.2. Data Collection Procedure

A total number of 150 participants took part in the placement test. From among these participants 105 received a score between 480 and 530. Some of them were excluded from the study to have 90 participants. An attempt was made to have equal number of participants in terms of their gender. They had 125 minutes to answer all the questions in the placement test. The TOEFL pbt practice test was chosen as the placement

test as it was easy to place the participants into different levels based on the score they received. No more time was allocated to those who needed more. After choosing 90 participants who were considered to be intermediate language learners, they were divided into two groups of 45. Each group received a task and had 5 minutes to plan what they were going to say. One of the tasks had both backgrounding and foregrounding events on it, but the other one had just the foregrounding events.

The tasks contained 6 pictures which told one story together. As both of the tasks were loose in nature, it was not difficult to unravel the theme of the story. Some of the participants used all their 5 minutes to plan their narration, and an overwhelming majority of them, especially those without backgrounding events on their tasks, used less than the permitted time.

The participants were informed about the purpose and importance of the study and that their performance would not affect their score. Instructions for completing the tasks were given in oral form in Farsi. The time for the completion of the task was different for each participant, but on average it nearly took 10 minutes to complete one task. A digital recorder was used to record the data and was played back as many times as needed to transcribe the data.

## 4. Results and Discussion
### 4.1. Results
For the sake of clarity, the results are examined for hypotheses 1 and then 2. Regarding the first hypothesis, the presence of backgrounding events has significant effect on the grammatical complexity of the productions. It is assumed that foregrounding effects are present in both tasks. As foregrounding is concerned with the main pivotal theme of tasks, its presence is indispensable in narrative tasks. It actually deals with the effects of both foregrounding and backgrounding together on the production of the cartoon prompts. To analyze the first hypothesis (to what extent does the presence of foregrounding on narrative tasks affect grammatical complexity in an EFL setting?), a non-parametric Mann Whitney U-test was used to examine differences of grammatical complexity in the two groups. The results indicated that there was a statistically significant difference between the two groups. U was calculated to be 9, and given the significance level of 0.001, which is smaller than 0.05, it can be concluded that with 95% of confidence level, H1 is supported. In other words, based on Table 2, the differences seen in mean scores are not by chance, and the presence of backgrounding effects was influential in improving grammatical complexity of students'

storytelling. This improvement level was higher in the group which received pictures with backgrounding events (Mean = 67.8).

The results of the descriptive statistics of the grammatical complexity of the two groups are shown in Table 1. Examining the mean scores of the two groups, it was found out that the participants with backgrounding events used more complex structures in comparison with their counterparts. The standard deviation of the grammatical complexity of the group with backgrounding is higher than the group without background events. This shows that, although the participants with background events performed better, their scores in grammatical complexity are more heterogeneous in comparison with the other group.

Table 1.
*Descriptive Statistics of Grammatical Complexity Scores in the two Groups*

| | Group Statistics | | | |
| | Group | N | Mean | Standard Deviation |
|---|---|---|---|---|
| *Grammatical complexity* | -background | 45 | 42.86 | 5.02 |
| | +background | 45 | 70.20 | 8.97 |

To put it simply, pictorial details or backgrounding events are influential in improving participants' grammatical complexity. Paying attention to details in the task lets them decipher the story better, hence using more independent clauses, sub-clausal units as well as more sub-ordinate clauses which led them to use sentences that, according to the design of the study, are considered to be more complex compared with those without these clauses.

Table 2.
*Mann Whitney U-Test to Examine the Difference in the Grammatical Complexity Mean Scores in the Two Groups*

| | Group | N | Mean Rank | Sum of Ranks | Mann-Whitney U | Z | Sig |
|---|---|---|---|---|---|---|---|
| *Grammatical complexity* | Male | 45 | 42.26 | 1986.00 | 858 | 1986 | .218 |
| | Female | 45 | 49.05 | 2109.00 | | | |
| | Total | 90 | | | | | |

To test the second hypothesis (to what extent does the presence of foregrounding and backgrounding on narrative tasks affect lexical diversity in an EFL setting?), an independent samples t-test was run to

compare the lexical diversity mean scores of the two groups. The results of descriptive statistics for lexical diversity are shown in Table 3.

Table 3.
*Descriptive Results of Lexical Diversity Scores in the Two Groups*

|  | Group | N | Mean | Standard Deviation | Standard Error Mean |
|---|---|---|---|---|---|
| *Lexical diversity* | -backgrounding | 45 | .48 | .10 | .01 |
|  | +backgrounding | 45 | .56 | .09 | .01 |

The results of the t-test analysis showed that there was a statistically significant difference (t=4.09, df = 88) in the two groups. Given the significance level of 0.001 which is smaller than 0.05, it can be concluded that with 95% confidence level, $H_1$ is supported. In other words, the differences observed in the mean scores are not by chance, and the presence of pictorial details has been influential in improving students' storytelling and lexical diversity. The group with backgrounding events on their task showed more diverse vocabulary than the second group, i.e., without backgrounding events on their task (Mean=0.56).

Table 4.
*Independent Samples t-test Results for Comparing Lexical Diversity Mean*

|  |  | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
|  |  | F | Sig. | T | Df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
|  |  |  |  |  |  |  |  |  | Lower | Upper |
| *Lexical Diversity* | Equal variances assumed | .836 | .363 | -4.09 | 88 | .001 | -.08 | .02 | -.12 | -.04 |
|  | Equal variances not assumed |  |  | -4.09 | 87.2 | .001 | -.08 | .02 | -.12 | -.04 |

*4.2. Discussion*

In the present study, we set out to examine the complex nature of foregrounding and backgrounding events. We also attempted to find out whether language learners used all their knowledge in narrating a task.

Additionally, we wanted to know what effects backgrounding and foregrounding events could have on lexical diversity and grammatical complexity.

The main findings can be summarized as follows: it was found that backgrounding events stimulated language learners to use more complex structures. We also found that when backgrounding events were present, language learners used more diverse lexical items. In what follows, we strive to explain how tenable the findings are, comparing and contrasting them in light of previous studies.

A number of studies have focused on lexical diversity and grammatical complexity being produced by language learners, only few of which have investigated the effects of foregrounding and backgrounding on the above variables (Ahmadian, 2009; Foster & Tavakoli, 2009). The hypotheses were concerned with the effects of foregrounding and backgrounding on (1) grammatical complexity and (2) lexical diversity.

It was hypothesized that tasks with foregrounding and backgrounding events would have positive effect on the participants' grammatical complexity. Narrating a task with both foregrounding and backgrounding events needed more complex structures in comparison with narrating a task with just foregrounding event. Narrating the main theme of the story (foregrounding) as well as the side story that happens alongside the main theme (backgrounding) needed more independent clauses, sub-clausal units, and more subordinate clauses which all together helped to have more complex structures. Therefore, it could be concluded that foregrounding and backgrounding events positively influence the grammatical complexity of the language produced by the participants.

In Skehan and Foster's (1997) study, participants narrated the story, using less complex language when planning time was a factor. The results of the present study do not confirm Skehan and Foster's finding. In the present study, the group with the backgrounding events on their tasks needed more time before starting to narrate, i.e., they used their entire planning time. Therefore, in the present study, those with planning time used more complex language than those without planning time. But is this increase in grammatical complexity because of the nature of the tasks or planning time? Juxtaposing the findings of the two studies, one can conclude that it may have been planning which caused more complex language. In planning their time, participants could think about both the theme of the story and the language they were going to use. Preparing what to say may have equipped them with some formulaic or

prefabricated chunks which helped them using more complex language to narrate the task.

Our finding that planning time resulted in more complex language, however, supports  Skehan and Foster's (1999) study, which investigated the effects of task structure and processing load on narrative retellings. The purpose of the article was to pinpoint how inherent structure and processing conditions could affect performance. The inherent structure of the tasks was concerned with sequencing of the events taking place, and processing conditions with using planning time. Skehan and Foster (1999) found that the inherent structure of the tasks and time pressure had nothing to do with the complexity of the samples being produced, as it remained the same for all tasks. In the present study, those who needed more time to plan their narration had the more complex task, which in the end led to more complex language.

Lexical diversity was the second dependent variable which was examined. The participants with just the foregrounding events on their tasks were required to narrate the story. As the backgrounding events were not present in their task, they had to narrate the main theme of the story. Data gathered showed that the participants with both of the events on their tasks talked more and used more lexical items to narrate the story. The presence of backgrounding events as well as foregrounding events positively influenced the participants' lexical diversity. It was hypothesized that narrating the details of the story needed more lexical items for language learners to elaborate on a specific part of a story. Using their planning time to prepare what they were going to say, using new lexical items to narrate the story, and not making the narration too long were among the many factors that led to more diverse vocabulary for the participants.

Most relevant to our study is that of Tavakoli and Foster (2008). These researchers were curious to find out how story line complexity (backgrounding, and foregrounding) and inherent narrative structure (loose or tight) would affect the complexity, fluency, accuracy, and lexical diversity of the participants. As there were two groups of participants, one in Tehran, and the other one in London, the interpretations were somewhat more complex.  The four tasks were football (-backgrounding, tight), journey (-backgrounding, loose), picnic (+backgrounding, tight), and Walkman (+backgrounding, loose). The + backgrounding tasks led to more complex language on the part of the participants in comparison with the –backgrounding tasks. Comparing the tight narrative structure tasks with the loose narrative structure tasks, they found more complex language for tight narrative structure tasks. The

independent variables in the study did not affect lexical diversity in any discernable ways, although the participants in London produced more diverse language than those in Tehran.

In the present study, +backgrounding events helped the participants produce more complex and, at the same time, more diverse vocabulary. In Tavakoli and Foster's study, lexical diversity was compared between the participants who took part in Tehran and London, and the presence of more diverse vocabulary items was because of the exposure of the participants to English-native speakers in London. What these researchers failed to account for was why the same participants in the same setting produced more diverse vocabulary. Looking more closely into their study may reveal more diverse vocabulary in the language produced by the participants with +backgrounding events in Iran, or the language produced by the participants with +backgrounding in London. One of the participants' mother tongue in our study was English. Although he participated in the study as an intermediate language leaner, he produced much more diverse vocabulary compared with others.

Foster and Tavakoli's (2009) findings are closer to those of this study. They just used native speakers as their participants. They concluded that participants with backgrounding and foregrounding events on their tasks, which are simply harder to process and narrate, would produce more complex language. Tavakoli and Foster's (2011) study yielded similar results with non-native speakers. We can argue that living or spending some time in a country in which the target language is used, the presence of backgrounding events, and having time to plan may positively impact lexical diversity, but the presence of these factors does not necessarily lead to more grammatical complexity.

As these findings reveal, grammatical complexity seems to be affected by the structure of the tasks more than lexical diversity does. Other external factors including environment, planning time, and years of residence in a native language country might affect lexical diversity, too.

## 5. Conclusion and Pedagogical Implications

The first conclusion which is warranted in the context of the present study is that grammatical complexity and lexical diversity are malleable to different contextual factors, with either grammatical complexity or lexical diversity affected by some of these factors, but not the others. While, for example, residence in native-speaking country, presence of backgrounding events, and planning time may help language learners use

more diverse vocabulary, grammatical complexity does not necessarily lend itself to such factors and is more likely to be affected by the structure of the task.

The second tentative conclusion concerns the role of setting in determining the amount of lexical diversity produced. The setting per se is inconclusive in pushing language learners to use more diverse vocabulary. The level of language proficiency seems to play a more vital role than exposure to language in a native-speaking country. It does not make a difference whether language learners are exposed more to the target language in English-speaking countries. What matters most is language learners have attained a high level of language proficiency-whether it is in an EFL or ESL setting.

Based on the findings of the present study, it can also be concluded that having pictorial details in the tasks lead the L2 learners into using more complex language. This finding can have both theoretical and practical implications for learners, teachers and syllabus designers. Theoretically, pictures are among the most motivating stimuli for readers of a book to continue reading, and even more motivating for language learners to study the intended language on books. These pictures should be selected and designed in a very careful and thoughtful way. The focus here is on the kind and content of the pictures. What a picture conveys and how it does it is of great importance in helping language learners to make their passive knowledge and especially their passive vocabulary active.

Great emphasis should be put on selecting pictures and sequencing them if they tell one story all together. Those in charge of selecting the pictures have to select them in a way that tells the story clearly and makes the story a visual image in the participants' minds.

With regards to practice, warming up has always been a good way to make learners ready for the task. Teachers should be aware of the power warming up has in activating the learners' knowledge and use it whenever needed. Assuming that the pictures are selected wisely, teachers can draw learners' attention to the events happening in the pictures and help them make their competence come to surface. Teachers can use the foregrounding effects present in the picture to make the main theme of the story clear and use the backgrounding events, if there are any, to activate the intended structures and vocabularies. Learners also should be apprised by their teachers of the benefits tasks and especially pictures have in laying the ground for preparing them for the main task. Learning a language out of context would be hard if not impossible to be put into use. Being taught a specific grammar or lexicon by the target language

teachers need to be put into practice and tasks is the best ways to do so. Foregrounding and backgrounding events can act as consciousness raisers for the language learners. They will bring the underlying knowledge to surface. The intended structures and vocabularies can be used authentically by the learners if enough and careful attention is paid to the effects present in the tasks.

## References

Bardovi-Harlig, K. (1992). The telling of a tale: Discourse structure and tense use of learner's narratives. In L. Bouton & Y. Kachru (Eds.), *Pragmatics and language learning* (Monograph No. 3, pp. 144–161).Urbana–Champaign, IL: University of Illinois, Division of English as an International Language.

Bardovi-Harlig, K. (1998). Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition, 20*(4), 471–508.

Brown, R., & Bellugi, U. (1964). Three processes in the child's acquisition of syntax. *Haward Educational Review, 34*(2), 133-151.

Chotlos, J.W. (1944). Studies in language behaviour: IV. A statistical and comparative analysis of individual written samples. *Psychological Monographs, 56*, 75–111.

Crookes, G. (1986). Task classification: A cross-disciplinary review. Technical report No 4. Honolulu: Center of Second Language Classroom Research, Social Science research Institute, University of Hawaii.

Dry, H. (1992). Foregrounding: An assessment. In S. J. J. Hwang & W. R. Merrifield (Eds.), *Language in context: Essays for Robert E. Longacre* (pp. 435–450). Arlington, TX: The Summer Institute of Linguistics and the University of Texas at Arlington.

Ellis, R. (2003a). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2003b). Designing a task-based syllabus. *RELC Journal, 34*(2), 64-81.

Foster, P., Tonkyan, A., & Wigglesworth, J. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(2), 374-375.

Gariele, P. (2009). CAF: Defining, refining, and differentiating Constructs. *Applied Linguistics, 30*(4), 590 – 601.

Hamp-Lyons, L. (2011). Writing assessment: Shifting issues, new tools, enduring questions. *Assessing Writing, 16*(1), 3-5.

Hooper, P.J., & Thompson, S. (1980). Transitivity in grammar and discourse. *Language, 56*, 251-299.

Housen, A., & Kuiken, F. (2009): Complexity, accuracy, and fluency in Second Language Acquisition. *Applied Linguistics, 30*(4), 461-473.

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 acquisition* (pp. 77-99). Clevedon, UK: Multilingual Matters.

Long, M. H.,  & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 24*(3), 27-56.

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

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmit & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 84-102), Cambridge: CUP.

Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks. *English language Teaching Journal, 47*(3), 203-210.

Polanyi-Bowditch, L. (1976). Why the whats are when: Mutually contextualizing realms of narrative. *Berkeley Linguistics Society, 2,* 59-77.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language.* Harlow: Longman.

Read, J. (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.

Reinhart, T. (1984). Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics, 22*, 779-809.

Richards, B. J. (1987). Type/token ratios: what do they really tell us? *Journal of Child Language, 14*, 201–249.

Richards, J. C., Platt, J., & Weber, H. (1992). *A Longman dictionary of applied linguistics*. London: Longman.

Richards, J. C., & Theodore, S. R. (2001). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(2), 38-62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan , P., & Foster , P . (1999). The influence of structure and processing conditions on narrative retellings. *Language Learning, 49*(3), 93 – 120.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 245–256).Oxford: Oxford University Press.

Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 61*(Suppl. 1) 1, 37–72.

Tomlin, R. (1984). The treatment of foreground-background information in the on-line descriptive discourse of second language learner. *Studies in Second Language Acquisition, 6*(3), 115–142.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing,* 17, 65–83.

Von Stutterheim, C. (1991). Narrative and description: Temporal reference in second language acquisition. In C. A. Ferguson & T. Huebner (Eds.), *Crosscurrents in second language acquisition and linguistic theories* (pp. 385-403). Amsterdam: Benjamins.

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), 1-27