Comparative effectiveness of teaching formulaic sequences and academic vocabulary on complexity, accuracy and fluency of EAP speakers

Ayşegül Nergis
Istanbul University, Turkey
https://orcid.org/0000-0002-0498-2705
aysegulnergis@gmail.com

Abstract

The present study aimed to compare effectiveness of two pedagogical interventions for enhancing aspects of L2 oral performance in an EAP program. For five weeks, two groups of EAP students received focused vocabulary instruction as part of their syllabus. One group explicitly learned and practiced formulaic sequences (FS) used in spoken academic context and the other, academic vocabulary (AcaVoc). Participants' oral performance was tested on three occasions and examined in terms of three aspects of speech: accuracy, complexity and fluency. Results showed that the FS group significantly outperformed the AcaVoc group in terms of L2 oral fluency, while the AcaVoc group outperformed the FS group in terms of L2 oral complexity. On the other hand, there was no significant differences between the two groups in terms of L2 oral accuracy. The participants' use of the target items in posttest was also found to be associated with their L2 oral performance. The findings of this study can contribute to our understanding of how explicit learning of academic vocabulary and formulaic sequences shape the quality of L2 learners' oral performance in academic speaking tasks.

Keywords: EAP; academic vocabulary instruction; L2 speaking; formulaic language; oral fluency; complexity

1. Introduction

In recent years, there has been a growing interest in the role of vocabulary in EAP (Coxhead, Dang & Mukai, 2017; Gardner & Davies, 2014; Malmström et al., 2018). Recent research has examined how acquisition of vocabulary contributes to different modalities of L2 production such as writing (e.g., Cai, 2016), L2 oral fluency (e.g., McGuire & Larson-Hall, 2017; Nergis, 2021) and formulating academic word lists for EAP/ESP (e.g., Gardner & Davies, 2014). However, there is need for more research to further our understanding of how learning of vocabulary affects aspects of speech in L2. The current study aims to address this gap by way of comparing gains in L2 oral fluency, accuracy and complexity of two groups of EAP learners after receiving a 5-week vocabulary intervention of FS and AcaVoc.

2. Literature review

2.1. Vocabulary instruction in EAP settings

Recent research has examined the role of vocabulary knowledge in EAP from a myriad of perspectives. For example, Peters and Pauwels (2015) compared the effectiveness of recognition and output focused activities for enhancing the acquisition of academic formulaic sequences. They found that for EFL business students, cued output activities led to better learning outcomes. Cai (2016) instructed MA students in a Chinese university on genre knowledge and genre-specific lexical phrases. During the intervention, participants extended their genre knowledge and also examined FS in corpus concordance reading tasks and produced them in rewriting tasks. The results showed that the instruction led to observable gains in receptive knowledge of lexical phrases and spontaneous use of lexical phrases in rewriting tasks.

   In their 2015 study, Al Hassan and Wood found that instruction of lexical phrases improves writing performance of L2 learners. Bardovi-Harlig and Vellenga (2012) also reported positive effects of FS instruction on pragmatic knowledge and use of L2. Boers et al. (2006) conducted a study with 32 college students. They found that instructional focus on FS helped L2 learners come across as proficient speakers of L2. Wood (2009) conducted an experiment with one Japanese learner of English. The participant completed a fluency workshop with a focus on recognition of FS. Significant gains in fluency of the participant in a posttest were observed. In an earlier study, Taguchi (2007) investigated effects of chunks learning with a group of ($N = 22$) Japanese L2 learners in an American university. Participants practiced target chunks in communicative tasks. Their use of the target items in spontaneous speech tasks were counted twice

in a semester. It was found that on the second data collection session, participants used a larger amount and a more variety of chunks and seemed to possess an enhanced awareness of discourse.

Recent instructional studies have examined the effects of vocabulary learning on EAP speaking. For example, Nergis (2021) found that advanced L2 learners in an EAP setting (n = 37) benefited from formulaic sequences instruction, as compared to explicitly learning single academic vocabulary items, in terms of oral fluency, specifically articulation rate. McGuire and Larson-Hall (2017) also investigated the effects of vocabulary instruction on utterance fluency. In their study, a small group of non-native speakers studying at an American university received FS instruction through noticing activities for five weeks, while another group practiced regular speaking and listening tasks. Both perception-based and objective measures of fluency were used to assess participants' progress from the pretest to posttest. It was found that the FS group significantly outperformed the control group and used a larger amount of FS in their speech at posttest. As can be seen these studies, research on the role of vocabulary in EAP generally acknowledges the idea that L2 speakers in EAP settings need to acquire a large amount of frequently used academic vocabulary as well as context-appropriate formulaic expressions for active use and automatic, efficient speech in communicative tasks; but there is room for more instructional research on aspects of speech.

2.2. Aspects of linguistic performance: Fluency, accuracy, complexity

2.2.1. Defining fluency, accuracy and complexity

Skehan (1996) distinguished between three aspects of speech: fluency, accuracy and complexity. Fluency refers to language learners' capacity to produce speech with few hesitations and pauses. Complexity refers to elaboration of the current language system. It is associated with subordination and lexical richness of speech (Yuan & Ellis, 2003). Accuracy concerns the degree an utterance corresponds to target language norms (Skehan & Foster, 1997).

Fluency, accuracy and complexity, as three fundamental aspects of linguistic performance, are "part of the vocabulary of language teaching" (Fulcher, 2003, p.30). Actually, language teachers have an intuitive understanding of these terms and tend to classify classroom activities as *fluency* or *accuracy* based (Brumfit, 1984). Fluency, as a performance variable, is often associated with proficiency (De Jong et al., 2012) since at the conceptual level, it is associated with effortless processing of linguistic information and reflects underlying cognitive processes (Kormos, 2006; Segalowitz, 2010). Because of its complicated

nature, a number of types and definitions fluency has been formulated for research purposes. For example, Lennon talked about a narrow and broad sense of fluency while defining its relation to performance (narrow) and general proficiency and competence (broad). Segalowitz (2010) proposed that fluency can be explained by way of: (1) examining underlying cognitive processes, (2) mapping measurable variables in utterances and (3) examining impressions and perceptions on the part of listeners. Another framework of analysis was formulated by Skehan (2003) which was taken into consideration in the current study to analyze oral fluency, mainly because it allows a thorough and systematic map that encompasses all aspects of fluency related variables. Skehan's framework examines fluency as a performance variable and contains these three main aspects: (1) *breakdown fluency* refers to pauses in speech, (2) *repair fluency* is concerned with self-corrections and repetitions in speech, and (3) *speed fluency* reflects speed of delivery.

According to Skehan (1996), complexity involves taking risks in language production and restructuring the interlanguage. Researchers associate complexity of speech with existence of subordination in utterances (Clercq & Housen, 2017). Previous studies mostly dealt with the relationship between task planning conditions and complexity of speech. According to Foster and Skehan (1996), pre-task planning results in greater complexity of language production. Yuan and Ellis (2003) argue that accuracy and complexity of speech is enhanced in online planning conditions. In online planning conditions, participants have to allocate their attentional resources into the ongoing task and have to prioritize form over meaning. However, this results in a decrease in fluency. According to Ahmadian and Tavakoli (2011), complexity is also enhanced when learners are given a chance to repeat a task. Theoretical explanation of interaction effects between aspects of speech will be discussed in the following section.

2.2.2. Interaction between aspects of speech

According to Leonard and Shea (2017), three aspects of speech interact during the course of L2 development regardless of the learning context (i.e. natural or instructional setting) and are helpful to explain the processes underlying automaticity of the speaking act; however, they were also observed to develop at different paces during interlanguage development (Housen et al., 2012).

Although the three aspects of speech occasionally confound each other in linguistic performance, they still need to be distinguished from each other for research and testing purposes because they require different type of processing on the learner's part (Skehan, 1996). Due to their limited capacity to process information, human beings are not able to attend to all aspects of a task. This occurs when learners have difficulty in allocating attentional resources and have

to prioritize one aspect over the others (Skehan, 1996). For example, from the perspective of second language teaching, encouraging learners to produce language more accurately allows controlled rather than automatic processes (Ahmadian & Tavakoli, 2011), so for the sake of speaking in an accurate fashion, fluency is forfeited. On the other hand, in some tasks, speakers may draw on formulaic language which results in enhanced fluency; however, they may not be able to refer to the acceptable language norms so efficiently and they may fail to produce accurate and complex speech (Skehan, 1996; Yuan & Ellis, 2003). In fact, when learners are producing more fluent language they are prioritizing meaning over form (Ellis & Barkhuizen, 2005).

2.2.3. Vocabulary knowledge and aspects of speech

Although research on the effects of vocabulary instruction on L2 speaking is rare, a number of studies have recently investigated the role of vocabulary knowledge in enhancing specific aspects of L2 oral performance. For example, Uchihara and Saito (2019) found that for freshmen university students with varying levels of proficiency, fluency (speech rate) in L2 oral performance was predicted by productive vocabulary knowledge. In their 2012 study, De Jong et al. found that for advanced L2 learners productive vocabulary knowledge was a strong predictor of speaking fluency in Dutch as L2. In a similar vein, Koizumi and In'nami (2013) found that productive vocabulary knowledge, measured via L1-L2 translation, predicted fluency, accuracy and syntactic complexity of low level L2 learners. With a group of advanced L2 speakers in a UK university (*N* = 46), Uchihara and Clenton (2018) did not find a significant relationship between receptive vocabulary knowledge and lexical complexity in L2 speech. According to Uchihara and Clenton, these results might suggest that speakers can achieve communication even with less sophisticated vocabulary; indeed they usually prefer to do this to be more comprehensible to their listeners. In a recent study, Tavakoli and Uchihara (2020) examined the contribution of multiword units to L2 oral fluency and found that fluent performers used multiword units comparably more than less proficient groups of learners. In a recent research, Saito (2020) compared the effects of single and multi-word units on comprehensibility of lexical appropriateness of L2 speech. The results showed that L2 oral comprehensibility and lexical appropriateness of Japanese learners of English was determined by use of multi-word units.

From the works cited above, it is clear that there is a growing interest in vocabulary learning and use in EAP. There is also need for more instructional research to guide efforts to enhance effectiveness of EAP programs in terms of helping L2 learners acquire academic vocabulary to improve the quality of their

oral performance in classrooms; however, it is still unclear how learning of academic vocabulary and FS affects all three aspects of speech. The current study aims to address this gap.

3. The study

EAP practitioners allocate some teaching time to learning and practice of vocabulary. However, there is still much to investigate about the role of learning academic vocabulary for enhancing aspects of L2 speech. Previous research has examined this pedagogical issue in correlational studies (De Jong et al., 2012; Saito, 2020) and by assessing oral performance variables through perceptive measures (Uchihara & Saito, 2019). Instructional studies, on the other hand, are quite rare (McGuire & Larson-Hall, 2017; Taguchi, 2007; Wood, 2009). To address this gap, the present study aims to explore comparative effectiveness of two vocabulary interventions for L2 oral performance. Given the research outlined above, it was hypothesized that the FS group would outperform the academic vocabulary group in terms of L2 fluency. It was also predicted that after receiving vocabulary intervention, the participants' oral performance would be in association with their production of the target items. The following research questions were addressed in this study:

RQ1: Does an academic-vocabulary-focused intervention or formulaic-sequences-focused intervention lead to better oral performance of advanced EAP learners?

RQ2: Are the variances in L2 oral performance related to the participants' use of learned vocabulary items?

3.1. Research design

Table 1 summarizes the research design. The sampling method of this study was convenience sampling; two intact EAP classes were randomly assigned either as the FS group or the AcaVoc group. Pedagogical intervention took 10 hours in each group and was delivered in five consecutive weeks. The posttest data were collected one week after the completion of intervention and a delayed posttest was delivered seven weeks after the posttest.

Table 1 Research design

FS group	Pretest ($N = 22$)	Instruction of formulaic sequences	Posttest ($N = 20$)	Delayed Posttest ($N = 10$)
AcaVoc group	Pretest ($N = 21$)	Academic vocabulary instruction	Posttest ($N = 20$)	Delayed Posttest ($N = 9$)

As can be seen, at the pretest, there were 22 participants in the FS group and 21 in the AcaVoc group. As participation was voluntary, some participants left the experiment after the pretest, leaving 20 participants in the FS group and 20 participants in the AcaVoc group at posttest. Due to smaller sample size on the third test time, the second delayed posttest was not used in final analysis.

3.2. Participants and the context of the study

Participants all were found to be successful in a standardized proficiency exam (fulfilling the university entrance requirement of TOEFL IBT 80 or IELTS academic module 6.5) the year before data collection and were accepted into their faculties as first-year students. They were from two intact first-year classes in an English-medium university in Istanbul. The participants were enrolled in a speaking course which was a four-credit Academic Skills course with a focus on academic speaking. The EAP course was designed around the content of world problems and aimed to help students equip with skills and knowledge that they can use in their academic studies.

The students enrolled in this course were from a variety of disciplines and the medium of instruction in these departments was English. The age range of the participants was 18-22 (*M* = 19.65). 21 participants were male and 19 were female. L1 of the participants was Turkish. Data were not collected from students whose L2 is different from English (i.e., students learning English as L3) and students who reported to have speaking anxiety or speaking/hearing impairment. Participation was voluntary and all participants signed an informed consent form.

3.3. Intervention

Two groups of EAP students received pedagogic intervention as part of their syllabus. While one group received intervention on FS, the other group received intervention on AcaVoc. Two class hours were allocated to each group for 5 weeks. The regular syllabus was task- and content-based EAP syllabus with a special focus on academic communication skills. Both groups were instructed by the same instructor who was an experienced EAP instructor and a PhD candidate in English Language Teaching. In order to address possible threats to validity, video-recorded pilot studies were examined by a small of group of experienced instructors and researchers one year before the experiment.

During the intervention, both groups completed the same activities in each session. At the beginning of each session, the students received a handout that contained the list of the target structures (6-8 items for each session) and concordances that showed how they are used in real-life academic spoken discourse. Sample handouts can be seen in Appendices D and E. At the end of each

session, students completed a worksheet that tested their learning of the target items. Concordances used in the handouts and end-of-session worksheets were also elicited from MICASE website and adapted for classroom use by way of eliminating repetitions and pauses markers such *erm* and *uhh* so that students would not feel encouraged to use them in their speech. The rationale for using the concordances from spoken academic corpus was to make learners familiar with the context in which target items are used in spoken academic discourse and therefore, make learning of these items more meaningful for them.

Interventions were implemented as follows: Each session started with a 5 minute warm-up session that involved brainstorming with pictures, videos and question-and-answer. Students examined pictures demonstrating global problems and answered questions related to the issues discussed in the weekly assigned reading and/or video(s). All visuals materials were presented on a white screen via a projector. After the warm-up session was finalized, students received the handouts described above. Sample concordances used in the intervention can be seen in Figures 1 and 2 below.

that was quite different, right? it may change.	that's my point	it might change, I say.

Figure 1 Sample concordance for the FS group

it doesn't matter if it's high carbon dioxide, it doesn't matter if it's an	accumulation	of waste products, it just simply doesn't matter.

Figure 2 Sample concordance for the AcaVoc group

In order to present the concordances of the target items, the instructor projected slides on the white screen and read them aloud and instructed the students to repeat them aloud after her. After that, the students examined the target items and concordances and discussed their use in real-life academic contexts in pairs or small groups for about 5 minutes. Following that, they completed communicative EAP tasks such as group discussions, pair and group research and oral summary activities and presentations and actively used the target items in their speech. Task requirements for each task were projected on a slide on the white screen. The students were instructed to use 3 to 5 target items in each task by referring to the worksheets delivered in the beginning of the lesson. This part of the session took around 20 minutes.

In the remaining time, the students were involved in less controlled tasks such as short oral reports or pair and group presentations and were instructed to

use 3 to 5 target items actively in their speech. After practising the target vocabulary in structured tasks before this stage, the participants were more familiar with them to a great extent and were able to use them in their oral production.

3.4. Selection of target items

A total of 82 academic vocabulary items and 80 formulaic sequences were learned and practiced during the intervention. Target items were systematically selected from popular coursebooks (for FS) and classroom-materials (for AcaVoc). The FS elicited from seven coursebooks (with a special focus on speaking) were then checked in two academic spoken corpora: Michigan Corpus of Academic Spoken English (MICASE) and British Academic Spoken English (BASE). A total of 80 FS were elicited. According to Ellis et al. (2008), it is important to triangulate instructor insight and corpus data in order to obtain a pedagogically valid list of vocabulary items to teach. Therefore, the target FS were also triangulated for their teachability. A group of experienced EAP instructors rated the target items' perceived teachability in an EAP context. Maximum reliability score of six raters were Cohen's k= 98 and minimum intercoder reliability was k=0.67. The final list of items were later categorized according to their function and added into the FS intervention syllabus (Appendix A).

As for the items used in AcaVoc instruction, two experienced language instructors scanned reading texts used in the course in order to elicit academic vocabulary that were worth teaching in the current EAP program. The items marked by the instructors were checked by the researcher in the Academic Word List (Coxhead, 2000). 82 academic vocabulary items were found in the Academic Word List and were used in the AcaVoc instruction (Appendix B).

3.5. Data collection

Three different task types were used to collect oral data from the participants on pretest, posttest and delayed posttest: 1) picture description, 2) oral narration and 3) oral argumentation. According to De Jong et al. (2012), all are complex and formal tasks which are appropriate for collecting speech data from advanced level L2 speakers participating in this study. The oral narration task was of persuasive nature, while the first two were descriptive tasks. Different task types were used because different task types have different cognitive demands on the learner (Bialystok, 2001) and that oral performance is sensible to different task types (Segalowitz, 2010). Participants were given 30 seconds to plan their speech and 3 minutes to complete each task. Prompts used in data collection sessions, conditions and timing of tasks were determined by corresponding

to previous studies and piloted before the experiment to ensure feasibility and avoid possible problems. Task prompts can be found in Appendix C.

Participants were also given a familiarity test that measured their gains in knowledge of target items (from pretest to posttest time). The test contained fill-in-the-blanks items adapted from MICASE and finalized by referring to opinions of experienced EAP teachers.

3.6. Data analysis

All speech data were transcribed and pauses and hesitations were marked in order to calculate oral complexity, accuracy and fluency measures. The data were later analyzed on PRAAT software (version 5.3.49) using a script that was created by De Jong and Wempe (2008, 2009) to calculate length of pauses and number of syllables (for oral fluency score). This script was used in previous research with different languages (De Jong et al., 2013). In this study a total of 6 measures were used. Details about the each measure are given below.

3.6.1. Measuring oral fluency

According to Skehan (2003), oral fluency as a performance variable contains three main aspects: *breakdown fluency* (pauses), *repair fluency* (repetitions, self-corrections) and *speed fluency* (speed and rate of speech). Adopting Skehan's framework, in this study, *pruned speech rate* (PSR) was used to measure L2 oral fluency because it takes into account speed factor and all dysfluency variables such as pause types and length, self-repairs and hesitations and was reported to be the most efficient oral fluency measure for research purposes (De Jong, 2016). *Pruned speech rate* was consistently used in previous studies as a global measure for L2 oral fluency (Lennon, 1990; Towell, 2002). It is the total number of syllables minus the number of repairs and repetitions and filled pauses divided by phonation time.

3.6.2. Measuring oral accuracy

SLA researchers have different views about how accuracy can be best measured (Ellis & Barkhuizen, 2005). Yuan and Ellis (2003) operationalized accuracy of speech as the percentage of error-free clauses and percentage of correct verb forms. The same method was used in this study because it was persistently used by previous researchers (Kormos & Dörnyei, 2004; Mackey & Gass, 2005).

The present study used two L2 oral accuracy measures: 1) *percentage of error-free clauses* and 2) *percentage of correct verb forms.* In this study, errors

are defined as syntactic, morphological and collocation errors; correct verbs forms include accurately produced verbs in terms of tense, aspect, modality and subject-verb agreement. However, these principles were applied only to calculate the two accuracy scores, not to categorize the errors in clauses and verb forms.

Two experienced EAP teachers (the first rater was the researcher, the second rater was an experienced EAP instructor with an M.A. degree in ELT) independently calculated the oral accuracy measures on 25 per cent of data. The external coder was trained by the researcher about how to do the calculations for accuracy and complexity scores. Intercoder reliability coefficient was calculated using Cohen's *kappa* (*k*). It was found that *kappa* value was 0.93 for the first measure and 0.91 for the second measure, indicating high inter-coder agreements for the measures.

3.6.3. Measuring oral complexity

In this study, L2 oral complexity was operationalized in terms of two measurements: (1) *syntactic variety* and (2) *mean segmental type-token ratio.* These measurements were also used by Yuan and Ellis (2003) and found to be efficient by Mackey and Gass (2005, p. 241). *Syntactic variety* is the number of different grammatical verb forms. *Mean Segmental Type-Token Ratio* is the number of different words divided by the total number of words in every 40-word segment. After transcriptions were completed, two experienced EAP teachers (the first rater was the researcher, the second rater was an experienced EAP instructor with an M.A. degree in ELT) independently calculated the complexity scores on 25 percent of data. *Kappa* value was found to be $k = 0.91$ and $k = 0.93$ for the complexity measures, showing high inter-coder agreement.

4. Findings

4.1. Main effects of intervention

This study examined effects of two types of instruction on L2 oral performance. First of all, learning of items was tested via the Familiarity test. The results showed that both groups had significant gains in their knowledge of target items from the pretest to posttest. On average, the FS group knew 25% of the target FS at the pretest; this figure increased to 95% at the posttest. As for the AcaVoc group, on average, they knew 16% of the target items at the pretest, but this figure increased to 93% at the posttest.

Following this, a repeated measures MANOVA test was conducted in order to determine the main effects of group (FS vs. AcaVoc) and time (pretest - posttest)

and also the effect of interaction between group and time. The results of the multivariate statistical test Pillai's Trace were reviewed to determine if there were significant differences between the two groups on a linear combination of the dependent variables. The Pillai's Trace was used because it is considered the most powerful multivariate test to be used with a small sample (Field, 2009). The assumption of equality of covariance matrices was first tested through Box's M Test. The result of this test was significant. To overcome such violations of equality of covariance matrices, Field (2009) suggests using the Pillai's Trace for test power. The assumption was then tested through Levene's test of normality. The results showed normal distribution of scores, eliminating the chance of making a Type I error. Table 2 shows the result of the repeated measures MANOVA.

Table 2 Multivariate effects for the interaction between intervention type and L2 oral performance scores in posttest

Source	Df	SS	F	Partial $\acute{\eta}^2$
Time	1	115687	1.949 ***	.813
Intervention	1	488402	.640	.216
Time X Intervention	1	4192E	4.247 ***	.999

* $p < .05$ ** $p < .01$ *** $p < .001$ ($N = 40$)

The results of repeated measures MANOVA showed that time and intervention interaction [$F (1, 30) = 4.247$, $p < .001$, partial $\acute{\eta}2 = .999$] had a statistically significant effect on the mean scores of the vector representing 2 oral complexity scores, 2 oral accuracy scores and 1 oral fluency score. Table 3 shows the result of Univariate ANOVAs.

Table 3 Significant univariate effects according to intervention type

Dependent variable	df	F	Intervention	Means	95% Confidence Interval	
					Lower Bound	Upper Bound
Syntactic variety	1	1.612 *	AcaVoc	13.33	12.47	14.15
			FS	13.04	12.24	13.96
Mean segmental type-token ratio	1	2.305 **	AcaVoc	.84	.81	.86
			FS	.83	.80	.86
Percentage of error-free clauses	1	.758	AcaVoc	.87	.83	.88
			FS	.86	.83	.88
Percentage of correct verb forms	1	.544	AcaVoc	.94	.91	.95
			FS	.92	.90	.94
Pruned speech rate	1	2.614 **	AcaVoc	2.85	2.62	3.01
			FS	3.37	3.13	3.59

* $p < .05$ ** $p < .01$ *** $p < .001$ ($N = 40$)

Follow-up univariate ANOVA tests showed that the AcaVoc group had significant gains in *syntactic variety* [F(1, 30) = 1.612, $p < .05$, partial $\acute{\eta}2$ = .997, Pillai's trace= .999] and *mean segmental type-token ratio* [F(1, 30) = 2.305, $p < .01$, partial $\acute{\eta}2$ = .997, Pillai's trace= .990], while the FS group significantly outperformed the AcaVoc group in *pruned speech rate* [F(1, 30) = 2.614, $p < .001$, partial $\acute{\eta}2$ = .997, Pillai's trace = .990].

3.2. Relationship between use of target items and oral performance scores

Spearman's rank order correlation was run with the count of FS and AcaVoc items used in the intervention and oral performance scores of each group in the posttest to explore the contribution of learning of target vocabulary to variances in oral performance. On average, the FS group used 55 target FS in the posttest (69 % of the target FS) and 45 target FS in the delayed posttest (56 %) while the AcaVoc group used 58 target academic vocabulary in the posttest (71 % of the target academic vocabulary items) and 40 target items in the delayed posttest (49 %). As mentioned earlier, the delayed posttest data was not used in the final analysis due to small sample size. Table 4 and 5 summarize the correlation coefficients between the count of items used and L2 oral performance scores of the participants in the posttest.

Table 4 Spearman correlation coefficients of the count of FS in posttest and dependent variables (N = 20)

Variables	Percentage of error-free clauses	Percentage of correct verb forms	Syntactic variety	Mean Segmental type-token ratio	Pruned speech rate
FS count	-.195	-.055	-.324	-.306	.353
Percentage of error-free clauses		.456	.374	.091	-.112
Percentage of correct verb forms	.456		.353	.035	.035
Syntactic variety	.374	.353		.029	.381
Mean Segmental Type-Token Ratio	.091	.035	.029		.026

Table 5 Spearman correlation coefficients of the count of AcaVoc in posttest and dependent variables (N = 20)

Variables	Percentage of error-free clauses	Percentage of correct verb forms	Syntactic variety	Mean Segmental type-token ratio	Pruned speech rate
AcaVoc count	.192	.173	.325	.375	-.136
Percentage of error-free clauses		.490	.368	-.374	.135
Percentage of correct verb forms	.173		.343	-.150	-.162
Syntactic variety	-.239	.343		-.196	-.113
Mean Segmental Type-Token Ratio	-.418	-.150	-.196		-.222

As summarized in Table 4, a moderate correlation was observed between FS count and *pruned speech rate* (*r* = .353) and a negative correlation between FS count and *syntactic variety* (*r* = -.324). These findings indicate that enhanced, fluent performance of the FS group can be linked to their use of target FS in the posttest. It should be noted that this analysis included only the target items used in the intervention.

As can be seen in Table 5, AcaVoc was found to be in moderate correlation with *syntactic variety* (*r* =. 325) and *mean segmental type-token ratio* (*r* = .375) of the AcaVoc group in the posttest. Similar to the previous analysis, only the target items used in the intervention were included in this analysis. These findings support the effects of the AcaVoc instruction on L2 oral complexity. As hypothesized, increased complexity of speech was linked to participants' (the AcaVoc group) use of target items in the posttest.

5. Discussion

The main research question in this study asked whether an academic-vocabulary-focused intervention or formulaic-sequences-focused intervention would lead to better oral performance of advanced EAP learners. Previous findings had suggested that academic vocabulary instruction contributes to increase in certain aspects of L2 performance. For example, Tavakoli and Hunter (2018) examined previous research and concluded that instruction of multiword units promotes fluency. On the other hand, Gatbonton and Segalowitz (2005) argued that in a communicative classroom, repetitive tasks involving formulaic expressions, should improve both fluency and accuracy. Jiang and Nekrasova (2007) further argued that when learnt properly, FS can enhance fluency as well as accuracy in linguistic tasks. Foster, Tonkyn and Wigglesworth (2000) indicate a direct link between proficiency and use of formulas in speech, asserting that it is only the proficient speakers who can incorporate fixed expressions appropriately into their speech. As for general academic vocabulary knowledge, it is widely discussed that a large vocabulary size leads to higher L2 proficiency (De Jong et al., 2012).

Guided by the theoretical and empirical research summarized above, it was hypothesized that the FS group would outperform academic vocabulary group in terms of L2 fluency. This hypothesis was confirmed. The FS group outperformed academic vocabulary group in *pruned speech rate*, the L2 oral fluency score used in this study. The effect of FS instruction was not more effective than academic vocabulary instruction for improving the accuracy or complexity of speech. AcaVoc instruction, on the other hand, significantly promoted two oral complexity scores as compared to FS instruction: *syntactic variety* and *mean segmental type-token ratio*. S*yntactic variety* was operationalized as the number of different grammatical verb forms in this study. As the academic vocabulary

instruction implemented in this study contained a great deal of new verbs, it is not surprising that the AcaVoc group outperformed the FS group in terms of *syntactic variety*. Acquisition of a large repertoire of academic vocabulary helped the AcaVoc group outperform the FS group in terms of capability to add lexical variety (i.e., complexity) to their speech.

L2 oral accuracy was the only aspect of speech that was found to be not affected by instruction type. This finding can be explained by trade-off effect. According to Foster and Skehan (1996), human beings have limited attention capacity and L2 learners do not have full control of L2; for this reason, a trade-off is expected to occur among accurate, complex and fluent performance of L2 learners especially in instructional settings. Indeed, Robinson (2011) is of the opinion that fluency contrasts with accuracy and complexity. Similarly, Segalo-witz (2010) asserted that a trade-off between accuracy and fluency should be expected generally in all L2 speakers due to cognitive demand of self-monitor-ing. Kormos (2006) also notes that some speakers are fluent because they choose to compensate for their weaknesses in other areas of speech such as their accent and oral accuracy. According to Skehan (1996), overachievement in one aspect of speech results in a situation he calls *undesirable fluency* in which speakers fluently produce output for the sake of pretending to speak faster, by ignoring the accuracy, lexicalization and communicative appropriateness of their utterances. For effective fluency, a balance between all aspects should be main-tained; additionally, a large repertoire of FS should be acquired for active use. In this study, this was achieved by targeting at promoting the participants' produc-tive vocabulary knowledge. As described by Fitzpatrick and Clenton (2017) pro-ductive vocabulary knowledge refers to active knowledge and recall of vocabu-lary in productive tasks. Helping learners build this knowledge is a mammoth task for educational practitioners in EAP and EMI settings who can only work with their students for a few hours a week. Pedagogical implications of this find-ing will be discussed in the following sections.

Another significant finding was that analyses addressing the second re-search question support the main findings of this study. The count of FS and *pruned speech rate* of the FS group in the posttest were found to be in a mod-erate correlation with each other (r = .353), showing that the main effects of FS instruction on L2 oral fluency were supported by the participants' use of the target FS practiced in the intervention. In the AcaVoc group, moderate correla-tions between AcaVoc count and oral complexity scores (*mean segmental type-token ratio* r = .375 and *syntactic variety* r = .325) were observed, indicating that academic vocabulary they learned in the experiment assisted them to produce complex utterances in the posttest. Previous correlation studies found similar results. For example, Uchihara and Saito (2019) found moderate correlations (*r*

= .342) between vocabulary scores and speech rate of Japanese learners of English at a Japanese university. Tavakoli and Uchihara (2019) also found a moderate correlation (r = .325 and r = .396) between objective fluency measures and multiword units in speech with a group of university students in a British university. McGuire and Larson-Hall (2017) reported that similarly earlier studies found positive effects of instruction on use of target vocabulary items in speech elicitation tasks. It is important to note that, this finding in the current investigation should be interpreted with caution because the count of FS and AcaVoc contained only the items covered in the intervention; generative use of FS and academic vocabulary was not taken into consideration. A similar act of caution was also exercised in previous research (e.g., McGuire & Larson-Hall, 2017).

6. Pedagogical implications

Several pedagogical implications emerge for vocabulary instruction in EAP. First of all, learning of AcaVoc and FS through concordances could help EAP learners build a large repertoire of vocabulary for active use in L2 speaking tasks. It is also clear that it is possible to integrate vocabulary intervention into EAP syllabus without interrupting regular coursework. Therefore, it can be suggested that EAP practitioners can apply concordance-based vocabulary instruction in their courses according to their students' learning needs.

Secondly, learners can benefit from both AcaVoc and FS instruction. For example, EAP instructors who use FS instruction in their classes can observe oral fluency gains in their students' speaking tasks if they teach the target items in concordances and integrate their use in communicative EAP tasks. However, they should use tasks that encourage the use of FS (Gatbonton & Segalowitz, 2005; Segalowitz, 2010). On the other hand, EAP students learning single academic words, as opposed to students learning FS, can have greater gains in L2 oral complexity, which indicates that instructors who aim to promote their students' L2 oral complexity can focus more on teaching single academic words from available spoken academic corpora. L2 oral accuracy, in comparison, might not be significantly affected by instruction type. However, instructors should consider trade-off effects between aspects of speech while interpreting their students' performance in L2 after receiving vocabulary intervention.

7. Limitations

There are a number of limitations of this study that need to be addressed in future research. A major drawback in research design was small sample size. Future research can consider replicating this study with larger groups and, if possible, with

different L1 backgrounds. It is also clear that instructional studies with samples from different levels of proficiency can shed light onto our understanding of how academic vocabulary instruction shapes L2 oral proficiency development. Future research can aim to fill this gap. As a final remark, this study used only transactional tasks to elicit spoken data. It would be intriguing to explore to what extent academic vocabulary instruction affects L2 speakers' quality of speech in interactional language tasks.

8. Conclusion

The current study was motivated to expand the literature about the effects of vocabulary instruction on L2 speech. It was found in this study that FS instruction, as has also been consistently discussed in previous studies, could be more effective than academic vocabulary instruction in increasing L2 speakers' oral fluency in an EAP setting. Academic vocabulary instruction, on the other hand, could be more effective in increasing the number of different verb forms and words used in speech. Although vocabulary learning in concordances cannot enhance all aspects of speech, it can help EAP learners to gain a large repertoire of vocabulary.

References

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35-59.

AlHassan, L., & Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes, 17*, 51-62.

Bardovi-Harlig, K., & Vellenga, H. E. (2012). The effect of instruction on conventional expressions in L2 pragmatics. *System, 40*, 77-89.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* Cambridge University Press.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*(6), 245-261.

Brumfit, C. (1984). *Communicative methodology in language teaching: The roles of fluency and accuracy.* Cambridge University Press.

Cai, L. J. (2016). An exploratory study on an integrated genre-based approach for the instruction of academic lexical phrases. *Journal of English for Academic Purposes, 24*, 58-74.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213-238.

Coxhead, A., Dang, T. N. Y., & Mukai, S. (2017). Single and multi-word unit vocabulary in university tutorials and laboratories: Evidence from corpora and textbooks. *Journal of English for Academic Purposes, 30*, 66-78.

De Jong, N. H. (2016). Fluency in second language assessment. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 203-218). Mouton de Gruyter.

De Jong, N. H., Steinel, M., Florijn, A., Schoonen, R., & Hulstijn, J. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition, 34*(1), 5-34.

De Jong, N. H., Steinel, M. P., Florijn A., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics, 34*, 893-916.

De Jong, N. H., & Wempe, T. (2008). *Praat script speech rate.* http://sites.google.com/site/speechrate/.

De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods, 41*(2), 385-390.

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics and TESOL. *TESOL Quarterly, 42*(3), 375-396.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language.* Oxford University Press.

Field, A. (2009). *Discovering statistics using SPSS*. SAGE.

Fitzpatrick, T., & Clenton, J. (2017). Making sense of learner performance on tests of productive vocabulary knowledge. *TESOL Quarterly, 51*, 844-867.

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition, 18*, 299-324.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354-375.

Fulcher, G. (2003). *Testing second language speaking*. Longman/Pearson Education.

Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: a focus on access to fluency. *Canadian Modern Language Review, 61*, 325-353.

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. John Benjamins.

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *Modern Language Journal, 91*(3), 433-445.

Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research, 4*, 900-913.

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum.

Kormos, J., & Dörnyei, Z. (2004). The interaction of linguistic and motivational variables in second language task performance. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 9(2), 1-19.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387-417.

Leonard, K.R., & Shea, C.E. (2017). L2 speaking development during study abroad: Fluency, accuracy, complexity and underlying cognitive factors. *Modern Language Journal, 101*, 179-193.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum.

Malmström, H., Pecorari, D., & Shaw, P. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes, 50*, 28-39.

McGuire, M. & Larson-Hall, J. (2017). Teaching formulaic sequences in the classroom: Effects on spoken fluency. *TESL Canada Journal, 34*(3), 1-25.

Nergis, A. (2021). Can explicit instruction of formulaic sequences enhance L2 oral fluency? *Lingua, 255*, 1-21.

Peters E., & Pauwels P. (2015) Learning academic formulaic sequences. *Journal of English for Academic Purposes, 20*, 28-39.

Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning, 61*, 1-36.

Saito, K. (2020). Multi- or single-word units? The role of collocation use in comprehensible and contextually appropriate second language speech. *Language Learning 70*(2), 548-588.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Simpson, R. (2004). Stylistic features of academic speech: The role of formulaic expressions. In U. Connor & T. A. Upton (Eds.), *Discourse in the professions: Perspectives from corpus linguistics* (pp. 37-64). John Benjamins.

Simpson, R. C., Briggs, S. L., Ovens, J., & Swales, J. M. (2002). *The Michigan Corpus of Academic Spoken English*. The Regents of the University of Michigan.

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

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*(1), 1-14.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 1-27.

Taguchi, N. (2007). Chunk learning and the development of spoken discourse in a Japanese as a foreign language classroom. *Language Teaching Research, 11*(4), 433-457. https://doi.org/10.1177/1362168807080962

Tavakoli, P., & Hunter, A-M. (2018). Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teacher Research, 22*(3), 330-349.

Tavakoli, P. & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Language Learning 70*(2), 506-547.

Thompson, P., & Nesi, H. (2001). The British academic spoken English (BASE) corpus project. *Language Teaching Research, 5*(3), 263-264.

Towell, R. (2002). Relative degrees of fluency: A comparative case study of advanced learners of French. *IRAL—International Review of Applied Linguistics in Language Teaching, 40*, 117-150.

Uchihara, T., & Clenton, J. (2018). Investigating the role of vocabulary size in second language speaking ability. *Language Teaching Research*, 1-18. https://doi.org/10.1177/1362168818799371.

Uchihara, T., & Saito, K. (2019). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal, 47*(1), 64-75. https://doi.org/10.1080/09571736.2016.1191527

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics, 12*(1), 39-57.

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.

APPENDIX A

Partial list of formulaic sequences practiced in the FS group (classified according to functions)

EXPLAINING FURTHER

what I mean when I say
what I mean by
what I really mean is
let me explain
by that I mean
what I'm trying to say is
I want to say

NOTING DEFINITIONS/REASONS

this is known as
the overwhelming majority
is the term for
that would be
and the reason is
for the following reasons

APPENDIX B

Partial list of academic vocabulary practiced in the AcaVoc Group

Week 1:

indicator
advocate
commitment
ultimately
annual
volume
integrity
devoted
discretion
refined
reliance
readjustment
accumulation
aggregation
alternatively
illustrative
analogous

Week 2:

overlapping
incorporation
parameters
output
prospective
brevity
commodities
incompatible
compounds
intensity
comprehensive
inconclusive
infrastructure
concurrently
non-conformist
monitoring

APPENDIX C

Prompts used in oral production tasks (pretest)

1) *Picture Description*
   Dear participant,
   In the first part of the interview, you will describe a picture. You are expected to speak without interruption. You have thirty seconds to examine the picture and plan your speech.
   You have three minutes to describe it. You can use the chronometer.

2) *Oral Narration Test*
   Dear Participant,
   Please read this article and summarize it orally. After you finish reading, you will have thirty seconds to plan your speech and three minutes to complete your oral summary. You can keep the article during your oral summary, but you are not allowed to read parts from it.

3) *Oral Argumentation Task*
   In this part, you will talk about YOUR IDEAS about some problems in the world.

   Here is a list of problems that need attention as determined by world leaders in the year 2000 in a meeting organized by the United Nations:
   - Children's health
   - Mothers' health
   - HIV/AIDS and other diseases
   - Gender equality issues
   - Extreme poverty and hunger
   - Education
   - Environmental sustainability
   - Developing partnership between the world governments

   Examine the list. Is there anything that you do not understand?

   Which one of these important problems do you think needs the most immediate attention and which one needs the least immediate attention and why?

   You have thirty seconds to plan your speech and three minutes to complete the task.

APPENDIX D

Sample handout and worksheet (Formulaic Sequences Group)

Handout 3

Below are some useful phrases that you can use to express your personal response. Examine them in context and use them in speaking exercises today.

It seems to me
In my opinion
What I think is
My point is
Personally I think
That's my point

The following are some examples of how these phrases are used in real-life university classrooms.

it's the same reason that, i don't have a a content analysis on this so my number could be wrong but	it seems to me	that about eight out of ten movies that come out are centered in Los Angeles. now you could say
for limited health care dollars you can't ignore that. the economic competition is basically	in my opinion	is what is driving this attack on nurse practitioner practice. basically, organized medicine and
and you're trapped in between two worlds. one of them is not just the postmodernist world, but	what i think is	the myth of the modern world which, i think he falls for, to a certain extent.
he doesn't think you have a God. i opt for the atheist Hume, in all of this. but	my point is	simply to situate him, in this complicated game of defending
was working from was a seventy thirty split. that seventy percent was spent on law enforcement, and	personally i think	that's a great split, personally, as a member of jail.
that was quite different, right? It may change.	that's my point	it might change, I say.

APPENDIX D.2.

Sample worksheet (for Formulaic Sequences Group)

Fill in the blanks with the most appropriate sequences.

| What I think is |
| my point is |
| in my opinion |
| it seems to me |
| personally I think |
| that's my point |

1. _____ quite different from authorities in the field. And it really makes a difference in how you look at what this second novel is about.

2. Album sales, like ticket sales for movies, are not, _____, something that it has high musical quality.

3. The local governors had criticized our actions but they were mostly jealous because we were successful. At least _____.

4. _____ there are a lot of mistakes in the paper. It lacks the visual data and graphs.

5. That movie was rather weak, but _____ overall it was fine because I could fine well-hidden social facts.

6. What looks evident to you may change. But _____. You know... the Earth revolves around the Sun. That's directly observable.

APPENDIX E

Sample handout and worksheet (Academic Vocabulary Group)

Handout 3

Below are words that were cited in the texts of this week. Examine them in context and use them in speaking exercises today.

refined
reliance
pursue
accumulation
visibility
alternatively
analogous

The following are some examples of how these words are used in real-life university classrooms.

but they in a sense become, a kind of ornament if you will, although extremely subtle, extremely	refined	, extremely spare. so in contrast with this idea of, high modernism something changes, in the sixties
plants, is the fact, there's a lot of, advantages we'll talk about, but, it may help us decrease our	reliance	on oil-derived chemicals. so lots of chemicals that're, additives in foods or, additives to plants
consumption in terms of, constraints given in the environment, that you create certain houses, you	pursue	certain economic strategies, because you are limited, given your environment, on what those strategies
if it's low oxygen, it doesn't matter if it's high carbon dioxide, it doesn't matter if it's an	accumulation	of waste products, it just simply doesn't matter, the metabolic response of the animal is identical,
well what we found in our assessment our program opportunities is that, we have more visibility. more	visibility	could strengthen, the partner, relationship and program attendance that's the opportunities they can
it differently. Kelly's phrase was what he called constructive alternativism. that you could	alternatively	construct the world and your experience, differently from moment to moment. it seems to me that
great freedom in building up and modifying the form. modelling is a lot like it is really in some ways	analogous	to oil painting. you have a chance, to fuss around with it you can change it. you can, you know

APPENDIX E.2

Sample handout and worksheet (Formulaic Sequences Group)

Fill in the blanks with the most appropriate words.

refined
reliance
pursue
accumulation
visibility
alternatively
analogous

1. There is no guarantee that you will succeed in this test, although it is _____ to the previous one.

2. Good promotion can help _____ of the Project.

3. Today we are thankful for the _____ of technological evolution.

4. It has become almost impossible to find _____ resources.

5. As a species, our _____ on animals as a source of nutrition may lead to our end.

7. You can volunteer in an established organization, or if you are rich enough you can _____ establish for own organization.

8. Who would like _____ a career where there are limited promotion opportunities?