

JAPAN ASSOCIATION FOR LANGUAGE TEACHING JALT2018 • DIVERSITY AND INCLUSION

NOVEMBER 23-26, 2018 • SHIZUOKA, JAPAN

Formulaic Sequence Acquisition in Academic Writing: Examining Different Interventions

Lewis Murray Kanazawa Universitv

Reference Data:

Murray, L. (2019). Formulaic sequence acquisition in academic writing: Examining different interventions. In P. Clements, A. Krause, & P. Bennett (Eds.), *Diversity and inclusion*. Tokyo: JALT.

For L2 learners, successful acquisition of formulaic sequences (FSs) is recognised as being valuable for academic writing. Studies suggest that cued output exercises requiring an evaluation effort may prove beneficial. The aim of this study was to examine the value of such exercises. Four classes in a Japanese university EAP programme were each assigned a different intervention over a 4-week period. Each intervention required a different degree of involvement with selected target FSs. Writing samples collected from participants before the intervention established no significant difference in target FS use between the groups. Postintervention data, drawn from the difference in individual participant's pre- and posttest target FS use, revealed significantly increased use only from the group assigned exercises requiring the greatest involvement, suggesting that such exercises may be important for acquisition. These findings are discussed in relation to other studies concerning cued output and evaluation effort.

第二言語学習者のアカデミック・ライティング学習には、定型表現の習得が有益とされる。判断負荷のかかる手がかり提示 型課題の効果を示唆した研究もある。そこで本研究は、そうした練習課題の有効性を検証するため、日本の大学のEAPコース で4週間にわたり、4つの通常授業クラスで各々異なる介入活動を行なった。各介入は、特定の定型表現に対し異なる度合い の関与を必用とした。介入前の授業参加者によるライティング・サンプルにおいては、グループ間の有意差は認められなかっ たが、介入後のデータでは、一つのグループでのみ、定型表現の使用に大幅な増加が認められた。ここからは、このグループの 参加者が行なった練習問題に、より多くの判断作業量が含まれていたことが、定型表現の習得のために重要であった、という 可能性が示唆される。こうした調査結果について、手がかり提示型課題と判断作業の問題を扱った他の研究との関係から、考 察を行なった。 **F** or L2 learners, successful acquisition of formulaic sequences (FSs) is recognized as being valuable for academic writing (AlHassan & Wood, 2015). A number of studies have produced corpus-informed lists of useful academic FSs (e.g., Hyland, 2008; Martinez & Schmitt, 2012; Simpson-Vlach & Ellis, 2010), however, fewer studies have been carried out into how L2 learners might best acquire utility with such language. This study builds on previous research into L2 learners' written academic FS acquisition, examining acquisition exercises within the framework of Laufer and Hulstijn's (2001) task-induced involvement load hypothesis, with a particular focus on the degree to which involvement with target FSs at the intervention stage factors into subsequent use. It also explores whether the degree of control regarding target FS use within exercises at the intervention stage impacts subsequent use.

In an effort to evaluate the above elements, data was drawn from four classes, each assigned a different intervention over a 4-week period, with each intervention requiring a different degree of involvement with selected target phrases. Significant gains, in terms of target FS use, were found only in the group assigned exercises requiring greatest involvement. Prior to a more detailed description of the intervention methods and discussion of the data, an examination of the pertinent literature relating to written academic FS acquisition is presented, alongside a brief description of the task-induced involvement load hypothesis, the framework within which the current study is situated.

Literature Review

Studies that have explored FS acquisition in the field of academic writing seem to have come to a consensus that providing learners with focused instruction is beneficial. Jones and Haywood (2004), for example, found that such instruction led to participants showing improved awareness of FSs and performing better in gap-fill style tests. However, they also noted that participants exhibited very little use of the target FSs in their essays. Lindstromberg, Eyckmans, and Connabeer (2016) used a modified dictogloss in which target FSs were highlighted, finding that participants who received the modified



dictogloss showed better use of target FSs when reconstructing a text than did those in the control group. Like Jones and Haywood, however, participants only displayed greater use under controlled conditions, leaving the authors speculating over the extent to which such an exercise would benefit participants' freer output (Lindstromberg et al., 2016, p.18). AlHassan and Wood (2015) found significantly increased use of target FSs among participants after a 10-week period of focused instruction, which also led to higher ratings by teachers of participants' writing. The focused instruction they provided included various activities aimed at both awareness raising and productive use, ranging from very controlled matching exercises to freer use-in-a-sentence exercises. However, as their study focused more on the benefits using FSs may bring, the authors provide no indication as to which of the instructional methods employed may have been more beneficial in terms of acquisition.

Other researchers have investigated what type of activities may be most beneficial. In their survey of FS acquisition, Boers and Lindstromberg (2012) suggested viewing the studies discussed within the framework of Laufer and Hulstijn's (2001) task-induced involvement load hypothesis. This hypothesis was originally proposed in relation to single-word vocabulary acquisition, defining tasks in terms of need, search, and evaluation, in which each element (if present) is determined to be either moderate or strong, depending on the degree to which each is required to successfully complete a given task. Alongside the evaluation element, these measures are key within the present study. As Boers and Lindstromberg (2012) noted, the evaluation element, which defines the selection process learners make regarding the suitability of a given lexical item for a particular context, has been shown to play a major role in single-word vocabulary acquisition. They concluded that, similar to single-word vocabulary acquisition, the evaluation element appears to play an important role in relation to FS acquisition, stating that "the few studies that did stimulate an evaluation effort in the course of the pedagogical intervention . . . tend to show relatively encouraging learning rates" (p. 100). They also noted that incidental acquisition appears to give little benefit, but that "intentional learning of FS, instigated . . . through the process of doing vocabularyfocused exercises, seems to generate more robust learning rates" (p. 99)

Peters and Pauwels (2015) explored the value of different treatment types in relation to written academic FS acquisition, concluding that, of recognition and cued output activities, the latter appeared to be more beneficial. The authors also noted that their findings echo Boers and Lindstromberg's (2012) conclusion regarding the limited value of incidental acquisition. The cued output activities provided in their treatment, requiring participants to use target FSs to complete gap-fill exercises, rephrase exercises, and use-in-a-sentence activities can all be seen to provide an evaluation element (Peters & Pauwels, 2015, p. 32), with that required by the use-in-a-sentence type activities seen to be strong (as defined by Laufer & Hulstijn, 2001). However, the relatively restricted nature of these activities, in terms of the target FS being determined either by a gap or specified for use in a sentence, arguably positions these exercises at the more controlled end of cued output.

Murray (2017) examined the effectiveness of a less controlled output exercise than those employed by Peters and Pauwels (2015). Participants were required to complete gap-fill exercises utilizing target FSs, constituting a relatively controlled activity requiring a moderate evaluation effort. On top of this, they were required to edit target FSs into paragraphs, an activity that is both less controlled and has a strong evaluation element. Murray found significantly increased use of target FSs among the intervention group in postintervention written tests. However, the effect of each intervention activity (the more controlled gap-fill and the less controlled editing) could not be determined as the intervention groups were assigned both activities. In contrast, the current study was designed to enable greater insight into the usefulness of each activity type in terms of promoting acquisition by allowing for a comparison of more controlled, moderate evaluation exercises with less controlled, strong evaluation exercises.

The above points are addressed in this study through the following research questions:

- RQ1. Does providing vocabulary-focused cued output activities significantly affect participants' subsequent use of target FSs?
- RQ2. Does the degree of control and evaluation effort required of the output exercises significantly affect participants' subsequent use of target FSs?

Methods

Participants in the study were taken from four intact classes in an 8-week academic writing programme at a Japanese university: Group A (n = 26); Group B (n = 28); Group C (n = 27); and Group D (n = 29). All participants were 1st-year students with an approximate level across classes of CEFR level B1 (determined by the university entrance exam). A total of 12 target FSs were selected for use in the study (see Figure 1), with selection based on three criteria. The researcher having taught on the writing course for the previous 2 years, the first of these criteria drew on his knowledge of the set assignments and awareness of phrases that would be useful for students, but that they typically do not use. Relying on intuition alone, however, has been found to provide phrases that a corpus analysis would determine to be less than useful (Martinez, 2013).



Therefore, in order to provide participants with useful and frequently occurring FSs, the corpus of British Academic Written English (BAWE; University of Oxford Text Archive, 2014), containing approximately 6.5 million words, was also consulted to ascertain the frequency of any given phrase. As is clear in Figure 1, of the 12 target FSs, the minimum frequency can be seen to be 14 times per million. In addition to BAWE, the Academic Formulas List (AFL; Simpson-Vlach & Ellis, 2010) was also consulted, with all but two of the target FSs present therein. The two phrases not present in the AFL were deemed to be frequent enough within BAWE to merit inclusion in the study.

	BAWE*	AFL**
as well as	366	0
in terms of	266	О
can be seen	191	0
a number of	178	0
is/ are likely to	102	0
it is clear	64	0
with respect to	49	0
in contrast to	35	Х
wide range of	32	0
to some extent	25	0
in spite of	21	Х
give rise to	14	0

Figure 1. Target formulaic sequences. * = The frequency per million words of target FS within BAWE; ** = the presence of each within AFL; O = present; X = not present.

Familiarisation

In the first 3 weeks of the course, prior to the intervention, all four groups were assigned the same worksheets designed to familiarize participants with the target FSs (see Appendix A for an example). Each week, a different worksheet was distributed showing two paragraphs in which target FSs were gapped, with the target FS required to complete the gap-fill listed above the paragraphs. In addition to completing these gap-fill exercises, participants were also required to match the target items with their function. Preintervention use of the target phrases in an uncontrolled activity was ascertained from the midterm paper (140-170 words), written at home, which all participants were required to submit in Week 4 as part of the writing course.

Intervention

During class time in Weeks 4 to 7, all four groups were again assigned to read a paragraph each week requiring a different degree of involvement with selected target phrases for completion, as follows:

- Group A: Target phrases were shown underlined, requiring no involvement or output, thus constituting a weak evaluation.
- Group B: Target phrases were gapped, requiring participants to select the appropriate item for completion. This can be seen to fulfil a moderate evaluation, as the words required to complete the tasks were provided.
- Group C: Target phrases were replaced with words/phrases serving the same function, requiring participants to edit in the appropriate item. This can also be seen to fulfil a moderate evaluation, as the FSs required to complete the tasks were provided.
- Group D: Assigned the same exercise as Group C, requiring a moderate evaluation. Additionally, participants were instructed to edit target FSs into paragraphs they had written themselves, an exercise for which they were allowed approximately five minutes of class time. While the editing activity typically made use of paragraphs participants had written for homework, the same as participants in the other groups were required to write, the editing itself was carried out in class. The activity can be seen to be considerably less controlled than the other exercises assigned in that participants were free to choose which of the target FSs to use (or at least were limited only by the context of the paragraph). By requiring participants to select appropriate FSs and decide how best to incorporate them within their paragraph, the evaluation effort is strong. Aside from the implication inherent in the editing activity itself that the target FSs would constitute a positive feature in their writing, Group D participants were given no additional explicit verbal or written instruction encouraging them to utilize the target FSs in their writing.

Postintervention data was drawn from the writing programme's final assessment (220-250 words) that participants were required to write in class. While language use did form



one part of the grading criteria for this assessment, use (or otherwise) of the target FSs did not factor into the grading. Preintervention and postintervention assignments were entered into AntConc software (Anthony, 2018), and occurrences of target FSs in each participant's papers were identified. A target FS frequency count per 200 words was then calculated for each of the two data sets for each participant. Individual gains between the pre- and postintervention data were then calculated. As the distributions within each group were found to be not normal, nonparametric Kruskal-Wallis tests were run to determine any significant difference between the groups in terms of preintervention use and gains made between the pre- and postintervention. Post hoc Mann-Whitney tests were then run to compare such gains between each of the groups to determine where the significance lay. As six Mann-Whitney tests were run, comparing each group with the others, a Bonferroni correction was made, adjusting the critical value for significance to .0083 (.05 divided by 6, the number of tests conducted).

Results

The median and range of target FS use (per 200 words) for each group within both the pre- and postintervention data are shown in Table 1. The same data are also given for the gains made within each group. More detail on individual participants' use of the target FSs in both the midterm and final assignments is shown in Appendix B. The Kruskal-Wallis test showed no significant difference between each group's use of target FS in the preintervention data, H(3) = .19, p > .05. However, with regard to pre- to postintervention gains, the Kruskal-Wallis test revealed a significant difference between the groups, H(3) = 25.75, p < .05. As noted above, Mann-Whitney tests were run subsequent to this, with a Bonferroni correction applied to give a significance level of .0083. Comparisons of Group A with Group B (U = 301, p = .146, r = .20), Group A with Group C (U = 350, p = .983, r = -.00) and Group B with Group C (U = 323, p = .226, r = -.16) showed no significant difference in target FS use. However, comparisons of Group D with Group A (U = 172, p = .000, r = -.49), Group B (U = 150, p = .000, r = -.58) and Group C (U = 210, p = .002, r = -.42) all showed significantly greater use in Group D. This would suggest then that of the various interventions applied, having participants in Group D edit target FSs into their own work was the only activity to produce significantly increased use of the target FSs between the pre- and postinterventions.

ind Range for Pre- and Postintervention Data and Gains
Group

Table 1 Median and Dange for Dre, and Destintant antian Date and Caine

	I			
Data category	А	В	С	D
Preintervention				
Median	0	0	0	0
Range	0 - 1.24	0 - 2.68	0 - 1.26	0 - 1.36
Postintervention				
Median	0	0	0	0.94
Range	0 - 1.66	0 - 0.86	0 - 3.46	0 - 3.42
Gains				
Median	0	0	0	0.94
Range	-1.24 - 1.66	-1.84 - 0.84	-1.26 - 3.46	-1.36 - 3.42

Note. The median and range of occurrences of target FS per 200 words for each group. Group A (n = 26), FSs underlined, weak evaluation; Group B (n = 28), FSs gapped, moderate evaluation; Group C (n = 27), FSs replaced, moderate evaluation; and Group D (n = 29), FSs replaced, moderate evaluation, FSs edited into original paragraphs.

Discussion

It is clear from the data that there was no significant difference in gains in target FS use between Group A, Group B, and Group C. The intervention assigned to Group A required no output or evaluation effort, and the data shows no gain in Group A's use of target FSs subsequent to the intervention. This echoes both Peters and Pauwels (2015) and Boers and Lindstromberg (2012), indicating that incidental acquisition has little impact on learner production of FSs. In contrast, the interventions assigned to Group B and Group C provided controlled cued output requiring a moderate evaluation. However, again it is clear from the data that there was no difference between the three groups' gains in target FS use, with the great majority of participants in all three groups exhibiting no use in both the pre- and postintervention tests. This would suggest that providing the type of vocabulary-focused cued output activities given in this study, requiring only a moderate evaluation effort, has almost no impact on most students.

The above appears to contradict Peters and Pauwels (2015), who conclude in favour of providing cued output exercises. However, the participants assigned cued output



activities in the study received all three types (gap-fill exercises, rephrase exercises, and use-in-a-sentence activities), which, as noted previously, required different levels of evaluation. It was therefore not possible to say which of the exercise types may have been more effective. The current study suggests, however, that gap-fill exercises alone, one type of cued output exercise requiring a moderate evaluation effort, may not be so effective in encouraging learners to use target FSs in their own writing. Therefore, in answer to the first research question about whether providing vocabulary-focused cued output activities significantly affected participants' subsequent use of target FSs, the data drawn from this study would suggest that it does not.

The second research question concerns whether providing less controlled output exercises requiring a stronger evaluation effort than those assigned to Groups A, B, and C would significantly increase participants' use of target FSs. In addition to receiving the same intervention as Group C, namely editing target phrases into paragraphs in place of words or phrases serving the same meaning or function, Group D were also instructed to edit target FSs into paragraphs they had written themselves. This latter exercise provides less control but a stronger evaluation element. The data shows that of the four study groups, only Group D showed significant gains between the pre- and postintervention tests. As the only difference between Group C and Group D was the editing exercises, it appears to indicate that this factor was indeed beneficial. In answer to the second research question then, the data indicate that providing less controlled output exercises requiring a stronger evaluation effort does indeed significantly increase participants' use of target FSs.

This supports Boers and Lindstromberg's (2012) proposal regarding the value of providing an evaluation effort within the intervention, in this instance indicating that providing a strong evaluation effort was necessary for significant gains. Furthermore, it supports the conclusion of Peters and Pauwels (2015) that controlled output appears to be beneficial but goes further by identifying less controlled output as the important factor in enhancing acquisition. It also builds on Murray (2017) by identifying the value of the combination of more and less controlled exercises over more controlled exercises alone. It is important here, however, to stress that it was a combination of output exercises that provided the results. While the findings suggest that more controlled exercises were the only factor differentiating Group C and Group D, caution should be taken in equating the latter alone with improved acquisition. Rather, the results of this study can be seen only to highlight the value of providing less controlled exercises requiring a strong evaluation effort in conjunction with more controlled exercises.

While median target FS use (per 200 words) within Group D was low (see Table 2), and some Group D participants showed no increased use of target FSs (see Appendix B), it is important to note that the intervention was limited to a 4-week period and carried out in class time, during which various other elements of academic writing had to be addressed. As such, only a very short time (approximately 10-15 minutes each week) could be allocated for participants to complete the assigned paragraph exercises and edit target FSs into their own paragraphs. With this in mind, the results may be viewed quite positively. A further positive aspect relates to the use of materials at the intervention stage. Of the different interventions groups were assigned, the more controlled exercises, which produced very little return, relied entirely on teacher-created materials. In contrast, those that produced the greatest results in terms of subsequent target FS use primarily utilized paragraphs written by the students themselves, and as such may be seen to better promote student autonomy.

One limitation of the study that may have resulted in greater use from Group D was the number of encounters with the target FSs. This factor was controlled for throughout the study by assigning the same paragraphs to each group ensuring that, regardless of the intervention activity, all participants would still encounter each target FS the same number of times. However, in allotting a short period during each class in which Group D participants were required to edit target FSs into their own writing, they were inevitably provided with extra encounters as part of the intervention. To what extent these extra encounters, as opposed to the involvement load of the editing exercises, may have impacted on Group D's use of target FSs is impossible to determine.

Conclusion

Prior research examining FS acquisition suggests that providing an evaluation effort at the intervention stage (Boers & Lindstromberg, 2012), providing cued output exercises within the intervention (Peters & Pauwels, 2015), and providing less controlled output exercises at the intervention stage (Murray, 2017) may all be beneficial for acquisition. This study indicated that assigning controlled output activities alone at the intervention stage brought no benefit in terms of subsequent use. The data showed that target FS use in Group B and Group C, both assigned controlled output activities, was no different to Group A, whose intervention activity required no involvement with the target FSs. However, data from Group D revealed a significant gain in comparison to the other three groups, suggesting that assigning less controlled output exercises requiring a strong evaluation may indeed significantly enhance acquisition.



Furthermore, as noted above, in utilizing students' own writing, the editing exercises assigned to Group D may be seen to promote greater student autonomy. This contrasts with the exercises assigned to the other groups that relied entirely on teacher-created materials yet brought minimal returns. With this in mind, a future line of research that may be of value would be to establish whether or not less controlled output interventions alone produce similar results to those achieved here. Should that prove to be the case, in reducing the reliance on teacher-created materials, the benefits to student autonomy could be considerably improved alongside greater FS acquisition.

Further areas of investigation that may be of value in examining participants' acquisition of target FSs include the appropriateness with which participants used the target phrases and their range of use. While acquisition in the current study was measured entirely on frequency of occurrence, whether or not the target FSs were used correctly and appropriately in context was not analysed. Appropriate usage would extend to frequency, examining whether participants who displayed a higher frequency count in fact made use of the same target FSs to a degree seen to constitute overuse. In terms of range, it could be of value to examine which of the target FSs participants tended to use more frequently and which they did not use. Examination of these two factors may in turn shed light on which FSs learners typically have more trouble in acquiring utility with, further adding to this field of research.

Bio Data

Lewis Murray is an assistant professor at Kanazawa University where he teaches primarily EAP classes. <lmurray@staff.kanazawa-u.ac.jp>

References

- 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. https://doi.org/10.1016/j.jeap.2015.02.001
- Anthony, L. (2018). AntConc 3.4.4w. Retrieved from http://www.antlab.sci.waseda.ac.jp/antconc_ index.html
- Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32,* 83-110. https://doi.org/10.1017/S0267190512000050
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*, 4-21. https://doi.org/10.1016/j.esp.2007.06.001

- Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 269-292). Amsterdam, the Netherlands: John Benjamins.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Lindstromberg, S., Eyckmans, J., & Connabeer, R. (2016). A modified dictogloss for helping learners remember L2 academic English formulaic sequences for use in later writing. *English for Specific Purposes, 41,* 12-21. https://doi.org/10.1016/j.esp.2015.08.002
- Martinez, R. (2013). A framework for the inclusion of multi-word expressions in ELT. *ELT Journal*, *67*(2), 184-198. https://doi.org/10.1093/elt/ccs100
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299-320. https://doi.org/10.1093/applin/ams010
- Murray, L. (2017). A study in enhancing L2 learners' utility with written academic formulaic sequences. *TESL Canada Journal*, *34*(3), 76-92.
- Peters, E., & Pauwels, P. (2015). Learning academic formulaic sequences. *Journal of English for Academic Purposes, 20,* 28-39. https://doi.org/10.1016/j.jeap.2015.04.002
- Simpson-Vlach, R., & Ellis, N. (2010). An academic formulas list: new methods in phraseology research. *Applied Linguistics*, *31*(4), 487-512. https://doi.org/10.1093/applin/amp058
- University of Oxford Text Archive (2014). *British Academic Written English Corpus*. Retrieved from http://ota.ahds.ac.uk/desc/2539

Appendix A

Example Familiarisation Worksheet Given in Weeks 1 to 3

Which of the phrases below goes in which gap?

1.	A number of	In spite of this	In terms of
	As well as	Is likely to be	It is clear

```
There are _/ ____/ __ reasons why commuting to university by bus in Kanazawa is not very convenient. Firstly, to get from my house to the campus, I have to take a bus into the city centre then back out. __/ ____/ __ commuting time, it is faster to go by bicycle. __/ ____ this, it is also quite expensive and __/ ____/ ___ very busy in the mornings. __/ ____/ ___, if the weather is terrible, I can stay warm and dry on the bus. Overall though, __/ ____ to me that taking the bus is not the best way.
```



• Which one of the phrases in **bold** signals:

Addition of similar information?	
Quantity?	
Degree of certainty?	
Reference to evidence?	
'Framing' the topic within a new category?	
Contrasting information?	

Which of the phrases below goes in which gap?

2.In contrast toWide range ofWith respect toTo some extentCan be seenGive rise to

__/ ____/ ___ what many people believe, the differences between British and American English are really quite minor. It is true that some words and spellings are different. __/ ____/ ____, the accents are also different. However, watching a British and American person talk with each other, it ____/ ____ that these differences are minor and rarely _____/ ___ misunderstandings. Overall then, it seems that, ____/ _____ general communication, there is not such a _____/ ____ differences between the two.

• Which one of the phrases in **bold** signals:

Degree of something?	
Addition of contrasting information?	
Quantity?	
Reference to evidence?	
Cause?	
'Framing' the topic within a new category?	

Appendix B

Participants' Use of Target FS in the Midterm (140-170 words) and Final (200-250 words)









