Classroom-Based Explicit Instruction of Formulaic Sequences: Effects on Japanese EFL Learners' Speaking Performance

Ken'ichi Yamazaki
Akita Prefectural University
ORCID ID: https://orcid.org/0009-0005-3211-2765

DOI: 10.23350/eltcrj.306

Article History
Received: 19 Feb 2026
Accepted: 26 March 2026
Published: 14 May 2026

Keywords: formulaic sequences; explicit instruction; EFL speaking; classroom-based research; learner corpus; Japanese university students

Abstract

This classroom-based study investigates whether explicit instruction in formulaic sequences (FSs) enhances Japanese EFL learners' speaking performance. Over a 15-week semester, first-year university students were divided into an experimental group (English B), which received structured FS instruction using model scripts, visual highlighting, guided rehearsal, and weekly short written tests, and a control group (English A), which followed the regular curriculum without script-based FS instruction. Pre- and post-speaking tests were administered using an individual monologue task. Quantitative analysis revealed that the experimental group demonstrated significant gains in both FS use and total word production (effect size $d = 0.95$), whereas the control group showed no significant improvement. Furthermore, a moderate positive correlation was found between written test performance and spontaneous use of learned FSs in speech ($r = .50$, $p < .01$), suggesting a transition from declarative to procedural knowledge. Comparison with a general learner corpus (ICNALE) further indicated that explicit instruction enabled learners to exceed baseline FS usage levels observed in naturalistic contexts. Qualitative observations showed that learners used FSs to organize discourse, add information, shift topics, and gain planning time, functioning as formulaic scaffolding that reduced cognitive load. These findings suggest that systematic FS instruction, combined with retrieval-based practice, can substantially enhance L2 speaking development in classroom settings. Pedagogical implications for integrating discourse-oriented FS training into regular EFL curricula are discussed.

Introduction

Achieving fluent, natural speech remains a significant challenge for many learning the English language. In particular, Japanese EFL learners often possess sufficient grammatical and lexical knowledge but struggle to produce appropriate speech in actual communication contexts. This "knowing-but-not-using" problem has long been recognized in second language acquisition research. In response, recent studies have increasingly highlighted the effectiveness of

formulaic sequences (FSs) in enhancing L2 learners' fluency and real-time speech production (Schmitt & Carter, 2004; Wood, 2015). While definitions of FSs vary, the present study adopts Wray's definition and treats FSs as multi-word expressions that are stored and used as single units in communication. Typical examples include discourse markers and other prefabricated expressions such as "by the way," "in addition," and "to be honest," which help speakers organize and connect ideas. Although the term overlaps with concepts such as lexical chunks and prefabricated expressions, the present study focuses specifically on sequences that function as ready-made units in real-time speech production rather than focusing primarily on isolated phrasal verbs or purely grammatical constructions. For Japanese learners, FSs with advanced discourse management functions are particularly difficult to acquire and use spontaneously. To address these challenges, this study investigated effects of explicit FS instruction on Japanese L2 learners' English-speaking performance using an experimental versus control group design. Furthermore, by comparison with a large-scale learner corpus, instructional effects were objectively evaluated against general usage patterns. The extent to which learning for written tests transfers into actual free speech was also investigated.

Literature Review

Formulaic sequences (FSs) are linguistic expressions that function as a single unit but comprise multiple words (Wray, 2002). For second language learners, mastering FSs is crucial. A substantial body of research has highlighted the importance of FSs in fluent L2 performance (Pawley & Syder, 1983; Schmitt & Carter, 2004; Wood, 2015). Research has also shown that acquiring FSs, particularly collocations, poses a significant learning burden for L2 learners (Peters, 2016).

Various types of FSs can be classified according to their functions within sentences. Discourse markers, as described by Schiffrin (1987), operate in relation to preceding and following utterances, organizing and guiding the flow of discourse. Although not grammatically obligatory, these expressions support smooth interaction and facilitate both speaker intention and listener comprehension. For instance, "in addition" functions as an elaborative discourse marker that signals the addition of information (Fraser, 1999).

For L2 learners, acquiring FSs—particularly discourse markers—is challenging. Nonnative speakers tend to use them less frequently than native speakers, and especially underuse advanced discourse-organizing markers (Müller, 2005). A contrastive interlanguage analysis of discourse markers used by nonnative and native English speakers showed that Japanese learners relied heavily on a limited range of basic markers and underused more advanced expressions (Shimada, 2014). These findings suggest that targeted instruction on discourse-related FSs may be particularly beneficial for Japanese learners.

However, several studies have highlighted the challenges learners face in acquiring and using formulaic expressions (Boers et al., 2006; Li & Schmitt, 2009; Wang & Christiansen, 2024). The effectiveness of FS instruction has also been supported by empirical research. For example, Wood (2009) reported improvements in learners' fluency following focused instruction in narrative discourse among Japanese learners. Recent empirical evidence suggests that the acquisition and active use of FSs directly contribute to various dimensions of oral fluency (François & Albakry, 2021; Nergis, 2021; Yu, 2022; Yucedal & Kara, 2023). For instance, Nergis (2021) showed that explicit instruction in academic FSs significantly enhanced pruned speech rate, while François and Albakry (2021) found that the frequency of FS use

predicted mean length of fluent run in standardized speaking tasks. Yucedal and Kara (2023) similarly reported that collocation- and idiom-enriched instruction led to higher speaking scores than a regular curriculum.

Nevertheless, two important gaps remain in the current body of pedagogical research. First, few studies have documented which specific FSs learners actually acquire and how far their use of these expressions can develop over the course of regular classes. Second, there is still limited classroom-based evidence from semester-long courses that systematically track pre- and post-instruction changes in speaking fluency. Further research is therefore needed to determine the extent to which intensive FS instruction yields robust and generalizable effects on learners' spoken performance. These gaps are particularly relevant in light of two theoretical frameworks that underpin the present study. Skill Acquisition Theory (DeKeyser, 2007) predicts that declarative knowledge of FSs can be transformed into procedural knowledge through deliberate practice, while Usage-Based Theory (Tomasello, 2003) emphasizes that repeated use drives language acquisition. Together, these frameworks suggest that explicit, practice-oriented FS instruction in classroom settings should yield measurable gains in spontaneous speech.

To address these gaps with greater objectivity, corpus-based methods were employed. A learner corpus is a systematically collected database of L2 learner language use (Granger et al., 2015). In recent decades, corpus research has clarified patterns of actual language use through frequency-based and register-sensitive analyses (Biber et al., 1998). Research on Japanese learner speech has also expanded, demonstrating how corpus data can inform language pedagogy beyond intuition-based approaches (Tono et al., 2014). For the purpose of ensuring objectivity, this study employed the ICNALE Spoken Monologue corpus (Ishikawa, 2014) to provide a statistical baseline for comparing the learners' FS usage with that of a general learner population.

Research Question

To address the issues mentioned above, the following research question was formulated:

> **RQ:** Does explicit classroom instruction influence Japanese learners' spontaneous use of FSs in speaking tasks?

Participants

Participants were 60 first-year university science students from a Japanese public university located in a regional city, with an enrollment of approximately 2,000 students, specializing in engineering and agriculture. They were drawn from two intact general English classes (English A and English B, $n = 30$ each). Their English proficiency was approximately equivalent to a TOEIC Listening and Reading (L&R) score of 350 (CEFR A2), reflecting approximately six years of English study in junior high school and high school. Speech samples were recorded from all initial enrollees. Following the exclusion of recordings containing unclear audio, 30 participants were randomly selected from each class, yielding a final sample of 60. The study received ethical approval from the university ethics committee, and all participants provided written informed consent before data collection.

The control group (n = 30) attended a general English class, English A, weekly for 15 weeks, using the same textbook as the experimental group. Each 90-minute class consisted of approximately 30 minutes of pair-work conversation practice based on textbook passages, followed by grammar and vocabulary exercises, along with reading and listening activities unrelated to conversation practice. The control group received no explicit instruction directed at FSs and took no written tests on FSs.

The experimental group (n = 30) attended another general English class, English B, with the same textbook and schedule, including the same 30 minutes of pair-work conversation practice as the control group, but received explicit instruction on FSs and took 14 short written tests weekly from Week 2 to 15. Each test required students to write 10 English FSs corresponding to Japanese prompts.

Instructional Procedures

Both classes were taught by the same instructor (the author) on the same weekday afternoon. The instructional procedures differed substantially between the control group (English A) and the experimental group (English B), although both classes used the same textbook and followed the same weekly schedule.

Control Group: English A

Students in English A followed the standard curriculum without explicit instruction on FSs. Each 90-minute lesson consisted of approximately 30 minutes of pair-work conversation practice based on textbook passages, followed by grammar and vocabulary exercises, as well as reading and listening activities.

Students were encouraged to express their ideas freely during speaking activities, but no model speech scripts were provided, and FSs were neither highlighted nor systematically practiced. No written tests targeting FSs were administered.

Experimental Group: English B

In contrast, English B incorporated systematic and explicit instruction of discourse-oriented FSs over 15 weeks. A total of 99 FSs were taught throughout the course, and a complete list of the FSs is provided in Appendix A. Each week, the instructor prepared a short model speech script related to the lesson topic and uploaded it to the Manaba online platform (Example 1). Students accessed the script on their smartphones during class.

Instruction followed a consistent four-step cycle:

1. Noticing Stage

Target FSs were embedded in the model script and visually underlined to draw attention. The instructor briefly explained their communicative functions (e.g., addition: “in addition”; contrast: “on the other hand”; stance: “to be honest”; conclusion: “all in all”).

2. Guided Reading and Rehearsal

Students read the model speech script (shown in Example 1) silently and practiced orally, completing parentheses and selecting options provided in brackets to personalize the speech.

3. Structured Production

The 30-minute conversation practice session followed the same procedure in each class:

(1) All students were paired for the activity (the instructor paired with any student without a partner).
(2) One student spoke for two minutes referring to the model script on their smartphone, then the partner did the same.
(3) The pair then had a conversation about the same topic using the script.
(4) Students randomly changed partners and repeated the process.

Students were encouraged to prioritize inclusion of target FSs (highlighted in the script) rather than fluency during practice.

Example 1. ***An extract from a model script used in English B.***

I think part-time jobs have some advantages.

First of all, we can ( ) by doing part-time job.

It's hard / rare for students to ( ) at college,

so part-time jobs lead to a great experience.

In addition, at the workplace, we have to

[take care of / pay attention to / listen to] ( ).

It is really [tough / interesting / challenging] for me.

This will be of great value.

On the other hand, part-time job has some disadvantages.

Above all, the worst point is that we ( ) by doing part-time job.

It is bad because we cannot ( ).

If college students cannot ( ), their college life will be ( ).

Note. Students used this model speech script following the four-step instructional cycle described in this section.

4. Recycling and Retrieval Practice

From Week 2 to Week 15, students completed weekly short written tests requiring them to produce 10 previously taught FSs in response to Japanese prompts within five minutes. These tests aimed to reinforce form-function connections and promote retrieval-based learning. An example of the weekly written test is provided in Appendix B.

Through repeated exposure, guided production, and weekly testing, the experimental group students were systematically encouraged to incorporate the target FSs into their spontaneous speech.

Data Collection Procedures and Corpus Comparison Data

In a quiet computer room, each participant recorded a one-minute monologue speech sample on the topic "My College Life" at the semester's beginning (Week 1) and end (Week 15). Before recording, participants had one minute to prepare. All recordings were transcribed verbatim by the author. FSs were identified and coded based on criteria proposed by Wray (2002). FSs were operationally defined as multi-word units explicitly taught as fixed expressions and treated as single lexical items in instruction, regardless of their degree of compositionality, reflecting their holistic processing and storage as proposed in the same framework. For example, expressions such as "in addition" and "to be honest" were coded as FSs, whereas compositional phrases not taught as fixed expressions were excluded. This definition is consistent with widely accepted approaches in previous research on formulaic language and ensures construct validity in identifying FSs in learner speech. The coded items were reviewed and discussed with a native English-speaking instructor who was familiar with formulaic language research, and any disputed items were excluded or revised until consensus was reached. This procedure was adopted to enhance the reliability of FS identification and coding. As measures of fluency and formulaic competence, the use of FS frequency and word count is supported by previous research (Ellis, 2012; Wood, 2015). The facilitative role of FSs in fluent speech production has also been widely discussed (Wray, 2002).

To provide an objective baseline for comparison, speech data from 100 Japanese EFL learners were sampled from the ICNALE Spoken Monologue corpus (Ishikawa, 2014), which contains 600 monologues by Japanese people at various proficiency levels. The ICNALE corpus task elicited opinion-based responses, resulting in frequent use of "agree with" and "disagree with." To ensure task equivalence, these expressions were excluded from the corpus baseline, yielding an adjusted mean of 1.17 FSs per speaker (50 of 100 speakers had used these expressions at least once). Comparing baseline corpus data with experimental data enabled objective evaluation of the instruction's effect.

Table 1. ***Formulaic Sequence Usage Count by Group*** **(*N* = 60)**

Group/Time	Total Words *M* (*SD*)	Total FS Count *M* (*SD*)	*t*(29)	*p*	*d*
Experimental (*n* = 30)					
1st Speech	38.53 (9.68)	0.87 (0.78)			
2nd Speech	52.30 (10.53)	1.70 (0.70)			
Change	13.77	0.83	5.22	< .01	0.95
Control (*n* = 30)					
1st Speech	40.40 (10.27)	0.70 (0.65)			
2nd Speech	47.67 (9.77)	0.87 (0.78)			
Change	7.27	0.17	0.93	.36	0.17

Note. FS=formulaic sequence. No significant between-group differences were observed at pretest for either total words ($t(58) = -0.72$, $p = .47$) or total FS count ($t(58) = 0.90$, $p = .37$).

Results

Table 1 shows a comparison of FS usage between the two groups. The experimental group demonstrated significant increase in FSs from the first to the second speech (+.83, $t(29) = 5.22$, $p < .01$, $d = 0.95$), but the control group's increase (+.17) was not statistically significant ($t(29) = 0.93$, $p = .36$, $d = 0.17$). In the second speech, the experimental group significantly outperformed the control group, with a large between-group effect size ($d = 1.13$).

As Table 2 shows, the experimental group increased both total word count and FS use from the first to the second speech. The group's total word count rose significantly ($p < .01$). Use of FSs practiced in class (planned FSs) rose from 0.27 to 1.37.

Table 2. ***Descriptive Statistics of the Experimental Group (n = 30)***

Variable	*M*	*SD*	Min	Max
1st Speech				
Total Words	38.53	9.68	19	54
Total FSs	0.87	0.78	0	2
Planned FSs	0.27	0.45	0	1
2nd Speech				
Total Words	52.30	10.53	29	71
Total FSs	1.70	0.70	0	3
Planned FSs	1.37	0.81	0	3
Written Test Average	7.97	1.34	4.57	9.71

Note. FSs=formulaic sequences, Planned FSs=FSs explicitly taught in class.

Table 3 compares sentence counts and sentences beginning with FSs between groups. The proportion of sentences starting with FSs was significantly higher in the experimental group (12.96%, $SD = 8.27$) than in the control group (2.67%, $SD = 5.50$) during the second speech ($t(58) = 5.68$, $p < .01$, $d = 1.47$). The experimental group showed a significant within-group increase (+8.50 percentage points, $p < .01$), but the control group did not.

Table 3. ***FS Sentence Counts and Percentage by Group*** **(*N* = 60)**

Group/Time	FS Sentences *M* (*SD*)	Percentage *M*
Experimental ($n = 30$)		
1st Speech	0.30 (0.47)	4.46
2nd Speech	1.17 (0.79)	12.96
Change	0.87	8.50
Control ($n = 30$)		
1st Speech	0.17 (0.38)	2.90
2nd Speech	0.20 (0.41)	2.67
Change	0.03	−0.23

Note. FS=formulaic sequence. FS sentences=sentences beginning with FSs.

Table 4 illustrates a notable increase in FS use, especially "in addition" and "to be honest," in the experimental group compared to the control group and the general learner corpus. For instance, the use of "in addition" rose from 1 to 10 in the experimental group, whereas the general corpus recorded no instance. Use of "to be honest" increased from 1 to 7, but was also never used in the control group or general corpus.

Table 4. ***Usage Frequency of Specific Formulaic Sequences***

Formulaic Sequence	Exp. Pre (n = 30)	Exp. Post (n = 30)	Cont. Pre (n = 30)	Cont. Post (n = 30)	General Corpus (N = 100)
in addition	1	10	0	2	0
to be honest	1	7	0	1	0
care about	0	5	0	1	0
all in all	0	6	1	1	0
thanks to	0	3	0	1	0
for example	4	6	4	4	20

Note. Pre=Pretest (1st speech); Post=Posttest (2nd speech).

Table 5 compares the mean number of FSs used per participant across groups. The experimental group's FS use (1.70) exceeded the general learner corpus baseline (ratio = 1.45), and the control group (0.87) remained below the baseline (ratio = 0.74).

Table 5. ***Comparison with the General Corpus Baseline***

Group	Total FS Count *M*	Ratio to the General Corpus
General Corpus (N = 100)	1.17	1.00 (baseline)
Exp. Post (n = 30)	1.70	1.45
Cont. Post (n = 30)	0.87	0.74

Note. Post=Posttest (2nd speech). Task-dependent FSs ("agree with" and "disagree with") were excluded from the corpus count to ensure task equivalence (50 of 100 corpus speakers used these expressions).

Discussion

One significant finding of this study concerns the dramatic increase in FS usage among the experimental group. Over 15 weeks, the experimental group showed significant gains in FS use in their speaking (effect size d = 0.95, which Cohen classifies as "large"). As shown in Table 4, FSs such as "in addition" (from 1 to 10 instances) and "to be honest" (from 1 to 7 instances) showed remarkable growth, but these expressions were virtually absent from both the control group and the 100-speaker general corpus. This pattern suggests that, without explicit instruction, Japanese EFL learners may need a long time to acquire and use certain FSs in spontaneous speech. Additionally, students in the experimental group produced many more words in the second test. In the experimental class, explicit instruction of FSs—repeatedly highlighted in model scripts, explained in terms of their communicative function, practiced through structured pair-work activities and weekly written tests, and actively used in their speeches—significantly impacted Japanese learners' English-speaking performance.

As is often pointed out, in Japan, a large amount of secondary school English lessons is still spent on grammar explanation and translation with limited opportunities for speaking

activities (Butler & Iino, 2005). This study has revealed that even Japanese EFL learners can enhance their speaking skills and use FSs in spoken production with enough instruction and practice in FSs.

These results support the findings of Wood (2009) and Yucedal and Kara (2023) regarding fluency improvements. Moreover, the findings of the present study converge with, yet also extend, a growing body of instructional research. McGuire and Larson-Hall (2018) found that the FS instruction group demonstrated gains in speed fluency and used a larger number of FSs at posttest. The present study corroborates these results in a larger sample (n = 30) within a semester-long regular classroom, suggesting that such effects are replicable under more ecologically valid conditions. Nergis (2021) showed that explicit instruction enhanced pruned speech rate in academic EFL contexts, while François and Albakry (2021) found that FS frequency predicted mean length of fluent runs. The present results align with these findings in demonstrating that FS instruction produces measurable gains in oral production, though the present study used total word count as its primary fluency index rather than temporal measures of speech rate.

On the other hand, the control group did not show any significant change in the number of FSs or total words. Thus, the experimental intervention clearly contributed to the observed improvement in measured FS use and fluency, whereas the control group showed no statistically significant gains (Table 1). One possible explanation for the relatively limited fluency often observed among EFL learners is the need for planning time during real-time speech production. Learners may be highly conscious of accuracy, which can reduce the time available for generating subsequent content. FSs, however, allow speakers to retrieve multi-word units holistically, thereby reducing processing load and enabling longer stretches of fluent speech (Wood, 2009). In contrast, producing less frequent or more complex vocabulary requires greater online processing, which may disrupt fluency. From the perspective of the Trade-off Hypothesis (Skehan, 2009), learners may therefore prioritize fluency over lexical complexity during real-time speech production. In the present study, students in the experimental group frequently used FSs at the beginning of their utterances. This pattern may indicate that sentence-initial FSs functioned as planning devices, allowing learners additional time to organize subsequent content and produce longer, more complex utterances. These findings suggest not only quantitative gains but also qualitative shifts in how learners manage cognitive load during real-time speech production. It should be noted, however, that this study measured only specific aspects of speaking performance, and the control group may have developed in other unmeasured areas. The present findings suggest that FS instruction is quite effective in enhancing L2 learners' speaking skills through formulaic scaffolding.

A key theoretical finding is the positive correlation between written test performance and the use of learned FSs in speech. Table 6 shows that written test scores correlated moderately with the number of taught FSs used in the second speech (r = .50, p < .01), indicating that students who performed better on tests tended to use more learned sequences in speech. The mean score on written tests was 7.97 of 10 (SD = 1.34), suggesting active engagement in FS learning. This result supports Skill Acquisition Theory's (DeKeyser, 2007) proposed transition from declarative to procedural knowledge. Instruction in classes, intensive written tests, and practice served as deliberate practice, promoting proceduralization. The instructional cycle—consisting of noticing through visually underlined scripts, guided rehearsal, and repeated retrieval through short written tests—may have strengthened form-function connections and facilitated later spontaneous use. This suggests that repeated written testing and speaking practice functioned as effective retrieval-based learning, strengthening knowledge and

promoting automatization. As Roediger and Karpicke (2006) pointed out, retrieval-based practice strengthens memory and facilitates later recall in production. Moreover, written test scores moderately correlated with total word count ($r = .48$, $p < .01$), suggesting that traditional written tests may contribute to increased numbers of words produced in English speech.

Table 6. ***Correlations in the Experimental Group (n = 30)***

	Total Words (2nd)	Learned FSs (2nd)	Written Test Avg.
Total Words (2nd)	—	.78**	.48*
Learned FSs (2nd)	.78**	—	.50*
Written Test Avg.	.48*	.50*	—

Note. FSs=formulaic sequences, Pearson's *r* for correlations among total words, learned FSs, and written test average, $^{**}p < .01$, $^{*}p < .05$.

Müller (2005) demonstrated that non-native speakers use discourse markers differently from native English speakers, suggesting that L1 background may influence how learners deploy discourse markers in L2 speech. The author's observations of Japanese university students' everyday L1 conversation suggest that equivalents to textbook-like FSs (e.g., "all in all") are rarely used in casual speech. Nevertheless, mastering such fixed expressions appears essential for improving L2 fluency. The present findings indicate that learners used a mixture of both relatively formal expressions (e.g., "all in all") and more conversational ones (e.g., "to be honest"). This may suggest that classroom instruction can override L1-based preferences and expand learners' functional repertoire.

Qualitative observations indicate that learners employed FSs for several purposes: (1) adding information, (2) providing examples, (3) shifting topics, and (4) gaining planning time. These uses suggest that FSs, especially discourse markers, function as scaffolding devices that reduce cognitive load (Wood, 2015) and enable longer utterances. During structured pair-work activities, students were encouraged to begin responses with target FSs, which may have contributed to the observed increase in sentence-initial usage. In the present study, students in the experimental group frequently used "in addition." This finding is partially comparable to that of François and Albakry (2021), who reported that clarifying FSs accounted for nearly 50% of all FSs used by their participants. While "in addition" can be considered a typical example of a clarifying expression, other frequently used items in the present study, such as "to be honest" and "all in all," appear to serve different discourse functions, including expressing stance and summarizing information. This suggests that learners may rely on a wider range of discourse-organizing FSs in spoken production. These findings reinforce the role of FSs as functional scaffolding devices.

Comparisons with the general learner corpus demonstrated that explicit instruction enabled learners to exceed average performance. After excluding task-dependent FSs from the ICNALE corpus (specifically, "agree with" and "disagree with," which appeared frequently due to the corpus task's opinion-eliciting nature), the adjusted baseline was 1.17 FSs per speaker. The experimental group's FS use (1.70) substantially exceeded this baseline, representing a ratio of 1.45, whereas the control group (0.87) remained below the baseline (ratio = 0.74). This pattern demonstrates that without explicit training, beginners' FS use remains below optimal levels, but that focused instruction can rapidly elevate performance beyond average levels observed in naturalistic learning contexts.

The present study combined explicit FS instruction with weekly retrieval-based written tests. While it cannot be strictly asserted which component contributed more to the observed gains, it could be suggested that the speaking practice embedded within explicit FS instruction played a central role in the experimental group's higher performance, given that it provided repeated opportunities for meaningful, contextualized use of target forms. Future research employing conditions that isolate each component would help clarify their respective contributions.

Taken together, this study extends previous classroom-based FS research by combining a controlled intervention, corpus-based benchmarks, and retrieval-focused assessment to document how specific discourse-oriented FSs develop over a semester. These findings advocate integrating retrieval-based FS training into EFL curricula to bridge the gap between controlled practice and spontaneous speech.

Limitations and Future Directions

During the classes, the instructor examined students' self-directed study practices. No participants were receiving private English conversation lessons. Both classes followed standard curriculum procedures taught by the same instructor using the identical textbook, minimizing differential motivational effects.

Several limitations must be acknowledged. First, this study's participants were first-year university science students at one institution, thus restricting generalizability to other academic majors and grade or proficiency levels. Different populations may yield different results. Second, the speaking task—an individual monologue on "My College Life"—represents only one task type, and FS use may vary across task types (e.g., interactive dialogues, presentations, debates). Future studies should include various tasks and participant demographics.

Third, the instructional intervention was implemented within a regular classroom setting using model scripts and weekly short written tests; therefore, it is difficult to isolate the relative contribution of individual components (e.g., visual highlighting, guided rehearsal, retrieval-based testing). Future research could experimentally compare different instructional elements to determine which components most strongly influence spontaneous FS use.

Fourth, while disputed FS coding items were resolved through discussion with a second rater, a formal inter-rater reliability coefficient was not calculated, which represents a methodological limitation that future studies should address.

Fifth, variations in individual students' effort and engagement may have contributed to some of the gains observed in the experimental group. The moderate positive correlation between written test scores and spoken FS use suggests that more motivated learners may have benefited disproportionately, and the degree of individual variation observed may limit the generalizability of the aggregate findings.

Future research should systematically develop discourse FS-focused instructional programs, including functional classification (addition, contrast, exemplification, conclusion), presentation of alternatives, and instruction on appropriate timing of use. Moving forward, addressing individual differences in instructional responsiveness will also be vital.

Conclusion

This study showed that explicit FS instruction significantly improved Japanese learners' English-speaking performance. After 15 weeks of instruction and written testing, the experimental group showed a large within-group gain in FS use (d = 0.95) and significantly outperformed the control group at posttest (between-group d = 1.13). Written test scores moderately correlated with FS use in speech (r = .50, p < .01), supporting a transition from declarative to procedural knowledge (DeKeyser, 2007).

One notable finding is the remarkable increase in FS usage, such as "in addition" and "to be honest," among the experimental group, indicating that such forms may take considerably longer to acquire without explicit instruction. This suggests that prioritizing explicit teaching of FSs that are difficult to acquire naturally can be especially effective for Japanese learners still developing their spoken English skills. For learners still at the stage of "thinking before speaking," this method may provide particularly strong benefits.

Comparison with a general learner corpus (adjusted baseline: 1.17 FSs) revealed that the experimental group (1.70) exceeded average performance by 45%, whereas the control group (0.87) remained 26% below average. This finding validates explicit instruction's effectiveness in elevating learner performance beyond levels typically achieved without explicit instruction.

Effective instruction requires not only exposing learners to useful expressions but also teaching them systematically, including repeated practice and usage opportunities. This approach—having teachers carefully select and model useful FSs and providing learners with repeated opportunities to retrieve and use them in meaningful tasks—proved effective in this study in enhancing Japanese EFL learners' speaking performance. Continued research is needed to further clarify which types of FSs should be prioritized and to explore optimal modes of instruction.

Acknowledgements

I sincerely would like to thank the editor and the anonymous reviewers for their valuable and helpful comments and suggestions on this paper. This study was supported by JSPS KAKENHI Grant Number JP26K04066.

References

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. 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*(3), 245–261. https://doi.org/10.1191/1362168806lr195oa

Butler, Y. G., & Iino, M. (2005). Current Japanese reforms in English language education: The 2003 "Action Plan". *Language Policy, 4*(1), 25–45. https://doi.org/10.1007/s10993-004-6563-5

DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 97–113). Lawrence Erlbaum.

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics, 32*, 17–44. https://doi.org/10.1017/S0267190512000025

François, J., & Albakry, M. (2021). Effect of formulaic sequences on fluency of English learners in standardized speaking tests. *Language Learning & Technology, 25*(2), 26–41. https://doi.org/10.64152/10125/73429

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics, 31*, 931–952. https://doi.org/10.1016/S0378-2166(98)00101-5

Granger, S., Gilquin, G., & Meunier, F. (Eds.). (2015). *The Cambridge handbook of learner corpus research*. Cambridge University Press.

Ishikawa, S. (2014). Design of the ICNALE spoken: A new database for multi-modal contrastive interlanguage analysis. *Learner corpus studies in Asia and the world, 2*, 63–76. https://doi.org/10.24546/81006690

Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing, 18*(2), 85–102. https://doi.org/10.1016/j.jslw.2009.02.001

McGuire, M., & Larson-Hall, J. (2018). Teaching formulaic sequences in the classroom: Effects on spoken fluency. *TESL Canada Journal, 34*(3), 1–25. https://doi.org/10.18806/tesl.v34i3.1271

Müller, S. (2005). *Discourse markers in native and non-native English discourse*. John Benjamins.

Nergis, A. (2021). Can explicit instruction of formulaic sequences enhance L2 oral fluency? *Lingua, 255*, 103072. https://doi.org/10.1016/j.lingua.2021.103072

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–226). Longman.

Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research, 20*(1), 113–138. https://doi.org/10.1177/1362168814568131

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Schiffrin, D. (1987). *Discourse Markers*. Cambridge University Press.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 1–22). John Benjamins.

Shimada, K. (2014). Contrastive interlanguage analysis of discourse markers used by nonnative and native English speakers. *JALT Journal, 36*(1), 47–77. https://doi.org/10.37546/JALTJJ36.1-3

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532. https://doi.org/10.1093/applin/amp047

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Tono, Y., Satake, Y., & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL, 26*(2), 147–162. https://doi.org/10.1017/S095834401400007X

Wang, S. Y., & Christiansen, M. H. (2024). Chunking in the second language: Implications for language learning and teaching. *Language Teaching Research Quarterly, 44*, 84–106. https://doi.org/10.32038/ltrq.2024.44.09

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*, 39–57. https://journals.lib.unb.ca/index.php/CJAL/article/view/19898/

Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. Bloomsbury Academic.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Yu, Y. (2022). The role of psycholinguistics for language learning in teaching based on formulaic sequence use and oral fluency. *Frontiers in Psychology, 13*, 1012225. https://doi.org/10.3389/fpsyg.2022.1012225

Yucedal, H. M., & Kara, S. (2023). The influence of formulaic sequences on students' oral proficiency via collocations and idioms. *Problems of Education in the 21st Century, 81*(6), 825–839. https://doi.org/10.33225/pec/23.81.825

Appendix A. *List of Formulaic Sequences Taught in Class*

Discourse markers	**Verb-based formulaic sequences**	**Prepositional / other formulaic sequences**
above all	add up to	a lot
all in all	aim for	a number of
as such	be able to	according to
at least	be about to do	as long as
first of all	be addicted to	as soon as
for example	be based on	at present
frankly speaking	be better off	at the same time
generally speaking	be famous for	at work
in addition	be good for	before you know it
in comparison	be of great value	behind the scenes
in fact	be proud of	cannot help but
in practice	be tempted to do	different from
in theory	care about	each other
in the first place	cheer up	even if
in a way	depend on	for the first time
on the other hand	end up ing	from now on
rather than	fill up	hand in hand
strictly speaking	find out	hundreds of
that is why	go out	in favor of
these days	hand down	instead of
to be honest	lead to	it is said that
to put it simply	listen to	not at all
without doubt	look after	not necessarily
for the following reasons	lose touch with	on the way to
in recent years	make a mistake	on time
	make an effort	ought to
	meet deadlines	over the course of
	pay attention to	per day
	shop for	regardless of
	split the bill	such as
	spend A on B	those who
	stick with	thousands of
	study abroad	together with
	take care of	written in stone
	take responsibility	a wide variety of
	think about	in some cases
	watch one's weight	thanks to

Appendix B. ***Example of Weekly Written Test***

Score _____

English B　mini-test 3

Name: _____

１．～につながる _____

２．加えて _____

３．実際には _____

４．～にとってよい _____

５．率直に言えば _____

６．外出する _____

７．世話をする _____

８．努力する _____

９．～する限り _____

１０．今から _____