

Article

Improving Classroom Teaching and Learning of Multi-Word Expressions for Conversational Use Through Action Research with Learner Feedback

Haidee Thomson 

Junior College Department, Hokusei Gakuen University, Sapporo 004-8631, Hokkaido, Japan;
h-thomson@hokusei.ac.jp

Abstract: Multi-word expressions make up a large proportion of the English language and particularly spoken language. Using multi-word expressions can assist with the impression of fluency, making them useful for language learners to know and use. However, proven methods for teaching this language phenomenon are required, so that learners can easily use multi-word expressions in their conversations. The purpose of the study was to examine the efficacy of a fluency workshop focused on multi-word expression use in conversation and to determine the most appropriate implementation for the Japanese context. An action research structure was used over three iterations of the fluency workshop, learner feedback and teacher observations were used to make improvements. Learner feedback regarding the usefulness of each activity for learning English was compared between the original cohort and subsequent cohorts. The results showed significant differences in levels of perceived usefulness for activities where improvements were made, but also for some activities where no specific improvement was made, suggesting that teaching improves through practice. Pedagogical implications include maximising the time on task via clear instructions, providing visual time constraints, and offering scaffolding to support the use of multi-word expressions when recall seems beyond a learner.

Keywords: multi-word expressions; conversational fluency; action research; role play; dictogloss; time on task



Citation: Thomson, Haidee. 2024. Improving Classroom Teaching and Learning of Multi-Word Expressions for Conversational Use Through Action Research with Learner Feedback. *Languages* 9: 336. <https://doi.org/10.3390/languages9110336>

Academic Editors: Jon Clenton, Batia Laufer and Gavin Brooks

Received: 29 April 2024
Revised: 17 October 2024
Accepted: 18 October 2024
Published: 28 October 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The ubiquitous presence of multi-word expressions in English, especially in spoken language, suggests that language learners need to learn and know these common sequences (Vilkaitė 2016). The term, “multi-word expression”, refers to word sequences that are often formulaic including “...(semi-) fixed recurrent phrases, ... multi-word verbs (put up with), ...speech formulae (What’s up), lexical bundles (in the middle of), and other types” (Syanova-Chanturia and Martinez 2015, p. 549). Multi-word expression use has been increasingly linked to speaking fluency in recent years (Boers et al. 2006; Tavakoli and Uchihara 2020), suggesting that the use of frequent multi-word expressions offers a boost to fluency for second-language and foreign-language learners (Wulff 2019). In search of an effective classroom approach to teach multi-word expressions, this paper will review the literature regarding how multi-word expressions can be learned. A classroom-based approach to learning multi-word expressions (e.g., I think I will, would you like to), is then reviewed and set up as an action research investigation, informed by learner feedback and teacher observations to suggest improvements for the learning of multi-word expressions in the EFL classroom in Japan.

Frequency of encounter is key to learning new vocabulary items (Webb 2007), and this is similar for multi-word expressions (Peters et al. 2023; Webb et al. 2013). Many English-language learners in Japan have limited exposure to the language in their daily life, and although English is taught in schools, the proficiency in Japan is low, ranking

87 out of 113 countries ([EF English Proficiency Index 2023](#)). The opportunities to speak are limited, even within the English classroom in Japan ([Koizumi et al. 2022](#)). Therefore, without intentional and repeated exposure to multi-word expressions, learners are unlikely to have enough encounters with the new language items to remember them for future use. Furthermore, the individual words that make up multi-word expressions are often already known by learners thanks to their high frequency, with many comprising multiple-function words, for example “is on the right”. As a consequence, learners may fail to notice a multi-word expression as a new item to learn due to the non-salient nature of the component words ([Jiang 2009](#)). The noticing hypothesis ([Schmidt 2001](#); [Schmidt 1990](#)) proposes that language items must be consciously noticed in order to be learnt, while the weaker version states that more noticing leads to more learning. There is also evidence to suggest that explicit noticing may not be required in the acquisition of transparent multi-word expressions (see [Szcześniak 2024](#)).

Textbooks often present exercises to raise learners’ awareness of multi-word expressions without prior modeling of correct usage. Typical exercises include gap-fill, multi-choice exercises, and matching the parts of the expression together. [Strong and Boers \(2019\)](#) found that learning phrasal verbs is more successful when correct usage examples are given, and the exercises are used for retrieval. In 25% of cases, where correct usage was not modelled, learners were repeating their errors one week later despite receiving corrective feedback. [Toomer and Elgort \(2019\)](#) found that for advanced learners, bolding collocations within a 500-word text where the collocation appears three times can enhance explicit knowledge, while reading the same text with no bolding can improve implicit knowledge of the collocations.

Previous research suggests that the active noticing of multi-word expressions in class leads to greater recycling of multi-word expressions from a test prompt ([Boers et al. 2006](#)). However, in a similar study ([Stengers et al. 2010](#)), when the test prompt was offered only in the L1, the participants did not use more multi-word expressions than the control group, suggestive that the noticing strategies used in class had not resulted in the ability to use any more multi-word expressions. In light of this, [Boers and Lindstromberg \(2012\)](#) argue that mere noticing or awareness raising is not enough if encounters are few and far between, leaving little or no memory trace. A more recent study ([Boers et al. 2023](#)) again drew learners attention to the phenomenon of multi-word expressions within classroom texts with additional guidance on how to use dictionaries and corpora to determine whether a string of words were indeed a multi-word expression, after 11-weeks the experimental group were writing essays using more accurate and wide-ranging multi-word expressions than the control group.

[Lin \(2019\)](#) argues that to learn multi-word expressions through reading alone would be to miss important features of their spoken prosody. Most of the multi-word expressions that L1 speakers know are likely learnt through hearing them in conversations. Young L1 learners will pick up the melody of expressions before the specific words and often use them with the correct prosody, even when they have yet to form the words ([Peters 1983](#)). Multi-word expressions can have specific prosodic rhythms which set them apart, so it seems important for L2 learners to hear multi-word expressions and learn their prosody when they learn them ([Lin 2019](#)). One way to learn, which involves hearing the prosody, is through watching online videos. [Majuddin et al. \(2021\)](#) found that learners in Malaysia could recall multi-word expressions better when they watched a 20 min video with captions in English twice, suggestive that incidental learning can occur if there is repetition.

Repetition is golden when it comes to language learning. However, repetition without variation becomes difficult to sustain, especially in a classroom. Some successful examples include [Wray \(2004\)](#) where one motivated TV presenter memorised the script to present a cooking programme on TV in the Welsh language (which was relatively new to her). Another relatively successful example was [Wray and Fitzpatrick \(2010\)](#) where learners chose utterances that they anticipated would be useful and had to memorise them. They then tried having conversations where they relied on these utterances and, depending on the learner, found

them confidence instilling and helpful for sustaining their conversations while also feeling the constraints of not knowing more. It therefore seems reasonable to expect that learning will be expediated by deliberate noticing accompanied by practice using the multi-word expression (to create and consolidate strong memory traces), particularly when learners have limited exposure to the foreign language, as is the case for learners of English in rural Japan.

In light of the potential for improved speaking fluency, classroom models for teaching multi-word expressions to low proficiency EFL learners are needed. Wood (2009), reported a promising model, when he reported a “fluency workshop” which gave learners multiple and varied activities to notice and repeatedly use multi-word expressions in the classroom. The intervention was reported to increase fluency, however only at a case-study level in Canada with a Japanese learner of English. The fluency workshop activities were well-known activities staged from input to automatization. The workshop included activities such as shadowing, dictogloss (Wajnryb 1990) (text reconstruction from audio), and mingle jigsaw (Wood 1998) (which involves memorising and sharing multi-word expressions). After these activities came practice and production, with activities such as 4/3/2 decreasing the time speaking on a topic with alternating partners (Arevart and Nation 1991; Boers and Thai 2017; Maurice 1983), and a free stage of speaking using the newly learnt expressions. The logical flow from input to practice building up to freer production with a variety of activities using the same multi-word expressions appealed to me as it is very reminiscent of Nation’s (2007) four strands approach to fluency. The four strands start with meaning-focused input, moving to meaning-focused output, language-focused learning, and finally fluency activities. The progression provides learners with the opportunity to speak at a higher level of fluency due to the progression from preceding activities where they have built familiarity with the language. This progression, therefore, is likely to tap into the learners’ proximal zone of development (Vygotsky 1978), a concept where tasks that are just outside of a learner’s independent abilities are made possible through support. The interactive nature of the activities also align with the socio-cultural theory that we learn language better when we use it to interact with each other (Vygotsky 1978).

The fluency workshop “floods” the learner with multi-word expressions much like the graded readers in the research by Webb et al. (2013) and Peters et al. (2023) where graded readers were “flooded” with particular multi-word expressions, and more encounters were associated with greater incidental learning. What is different here, however, is that learners are given multiple activities flooded with prompts to use the multi-word expressions (find them, repeat them, say them, recall them, roleplay with them etc.) rather than simply hear them while reading. From a teacher’s point of view the fluency workshop seemed practical and versatile. It also appears able to be adapted to any conversation context that is desired because once there is a conversation with the multi-word expressions, the activities can easily be created to go with it. Therefore, I adapted the fluency workshop to use in my EFL classes in Japan, with a focus on conversational interaction in travel-themed scenarios. The travel theme was selected since previous cohorts of students had responded to surveys about what they thought would be useful to learn in class, and using English for travel was a prominent response. The adaptation into a dialogue version for English learners in Japan was new. I therefore sought to collect evidence through action research regarding the efficacy and appropriateness of the intervention in this form and whether it could be improved for the specific learning context in Japan.

In action research, the researcher is often also the teacher, who identifies a classroom problem and then experiments with interventions to solve the problem (Burns 2010). In this case, I was the researcher and teacher, and I wanted to develop an intervention to address the lack of speaking fluency that my university-level learners of English displayed. The standard teaching practice was to use a textbook and work through each unit with limited opportunity for repeated practice with multi-word expressions. However, due to the links made in the research between the use of multi-word expressions and fluency, I felt that helping learners to use common multi-word expressions might help boost their speaking fluency. Action research follows an iterative cycle of question, intervention, and reflection

leading to the next cycle. Following this cycle is known to lead to improved teaching practice and is used by many research-orientated teaching practitioners to improve their teaching. The research questions were therefore as follows:

Research question 1: To what extent is a dialogue version of the fluency workshop effective for learning multi-word expressions?

Research question 2: What aspects of the classroom experience of the fluency workshop can be improved/made more appropriate for the Japanese context?

2. Materials and Methods

The specific action research cycle deployed here was to determine whether the fluency workshop intervention was effective and appropriate, and whether it could be improved for the Japanese context. The fluency workshop was run with three different cohorts of students over three semesters. I collected feedback from the participants regarding their experience of the activities via surveys and focus group interviews. I also took note of my teacher/researcher observations and reflections about whether improvements could be made to the classroom intervention. This feedback was then used to make improvements to the fluency workshop activities, and then the fluency workshop would be run with the second cohort of students. Feedback would then again be collected and reflected on, resulting in further improvements if required before running the fluency workshop again with the third cohort of students. The multi-word expressions, intervention activities, tests, and feedback surveys used will be described next.

In order to determine whether the fluency workshop intervention was effective, learner knowledge, use, and fluency with 30 multi-word expressions were measured using tests before and after the intervention. To measure knowledge, a cued recall test with the first letters of the words for the multi-word expressions included in a context sentence was administered via Moodle on the computers in the classroom. Then, in partners, the participants roleplayed a conversation which was recorded on the computers using Audacity in order to measure their use of the multi-word expressions in conversations along with fluency. The interactions during the activities in the fluency workshop were observed by the teacher/researcher and also reflected on by the student participants via surveys and focus group interviews after completing the fluency workshop. Feedback from the participants and teacher observations were used to make improvements to the classroom activities during a second and third iteration of the intervention with fresh participants/students experiencing the workshop each time.

Ethics approvals were applied for and granted at the research and teaching institutions. The students were informed about the nature of the research and asked to complete consent forms at the beginning of each study. The data reported are only from those who gave their informed consent. The participants were recruited from intact EFL classes (convenience samples) in an engineering university in Japan. The participants were only required to take two English courses as part of their degree. Self-reported TOEIC scores ranged from 255 to 600, equivalent to a false beginner to low intermediate A2-B1 CEFR range, there was no significant difference in this proficiency measure between the three cohorts of experimental groups according to the non-parametric Kruskal–Wallis test: $H(2) = 1.573, p = 0.455$.

Each fluency workshop was conducted over six 90 min class sessions using three units, intended to raise awareness, knowledge, use, and fluency with 30 multi-word expressions of four words. The materials for the experimental condition had travel-related themes that the students might experience and therefore consider useful (ordering food at a cafe, asking for directions on the street, and booking a hotel room). Each unit was built around a model dialogue developed by the researcher containing ten common multi-word expressions (transparent in nature). A list of the specific MWEs used along with frequency information can be seen in Appendix A. The multi-word expressions were checked against the spoken sub-corpus of COCA (Davies 2013) to confirm their common usage. The learners were familiarised with one round of the fluency workshop activities in class using different multi-word expressions before the study began.

The experimental group participated in the following fluency workshop activities based on the Wood (2009) model but adapted to a dialogue model. The activities started with input and progressively built knowledge and practice with support, leading to fluency activities where the multi-word expressions were used in roleplay conversations. The concept behind this progression is that after noticing the multi-word expressions in the context of the model dialogue, practising the pronunciation and rhythm, trying to recall the multi-word expressions from memory and using them in roleplay, that the learners would become familiar with the multi-word expressions and then try to use them in their own conversations. The activities were as follows:

- (1) Listening to the dialogue as a class and answering gist questions in pairs;
- (2) Marking pauses—the students listened to the audio as a class and marked pauses on the transcript (this activity was removed for the second and third cohorts);
- (3) Phrase instruction—the teacher showed the multi-word expression on the class screen and explained the meaning and function, and the students then found and marked the multi-word expression in their transcript and were encouraged to take notes;
- (4) Shadowing—the students listened to the audio on computer headsets repeating the words just after the speaker four times to learn the pronunciation and prosody/rhythm;
- (5) Dictogloss (Wajnryb 1990) (text reconstruction from audio)—the students covered their transcripts and listened to the audio once as a class without taking notes, then listened a second time while taking notes and sharing with a partner, after which they listened again and shared, and listened a fourth time before sharing notes with a small group and attempting to reconstruct the dialogue together;
- (6) Mingle jigsaw (Wood 1998)—the students were each given a multi-word expression on a piece of paper to memorise; they had to write it down and cover it so others could not see it. Papers were collected and then the students exchanged their multi-word expressions orally with others; they walked around sharing their multi-word expressions and returned to their desks to write down each multi-word expression. They were not permitted to write while listening to a speaker, and the goal was to find all 10 multi-word expressions;
- (7) Roleplay with notes—the students lined up facing a partner and roleplayed the dialogue while referring to notes; they rotated with partners three times and then switched roles and rotated with another three partners;
- (8) Decreasing-time roleplay without notes (based on 4/3/2 (Arevart and Nation 1991; Boers and Thai 2017; Maurice 1983)—the students roleplayed with a time limit that decreased with each partner change. They roleplayed with three different partners and then switched roles and rotated with another three partners;
- (9) Role-play recording—the students sat with a partner in front of a computer and audio recorded their roleplay without looking at notes. The students listened back to their recording and reflected with a worksheet on their speaking fluency (added from the second cohort onwards);
- (10) A related situation roleplay—the students chose a situation from several options that were similar to the model and then roleplayed ad lib with the encouragement to use the multi-word expressions if they could.

Participant feedback from the surveys and focus groups along with the teacher observations were considered and adjustments/improvements were made to the activities for the second and third cohorts as shown in Table 1.

One unit took about 160 min to complete. All three units along with pre- and post-testing were completed within six class sessions. During the final unit, the students were surveyed about their experience with each activity after they completed it. A control group which received no manipulation was included with the first and third cohorts in order to provide a baseline to consider whether any improvement within the experimental groups was due to the treatment or some other reason (see Loewen and Plonsky (2016, p. 34) for the definition of a control group). The control group was not deliberately exposed to the multi-word expressions (other than during the pre-test). They learned English for

the same number and length of class sessions; however, they continued with their usual focus on engineering topics through linked skill activities intended to improve their fluency progressively through listening, reading, writing, and finally speaking on each topic.

Table 1. Adjustments made to the fluency workshop activities after the first cohort.

Activity	Improvement
Shadowing	Added an explicit instruction to pause/replay as desired
Dictogloss	Changed to a small-group activity from the start with a visible countdown timer and a simplified worksheet
Mingle jigsaw	Added a visible countdown timer
Decreasing-time roleplay with no notes	Added a display of multi-word expressions, adjusted the timings, and added scaffolding for the hotel unit
Recording the roleplay	Added time and structure to listen to and reflect on the recording
Related situation roleplay	Reduced the scenario choices, displayed images, and added planning time in L1

In order to collect feedback from the participants about their experience with the classroom activities, a Likert scale survey was used. The survey asked the participants about how useful they perceived the activity to be for learning. The Likert scale for the first cohort had five response options: strongly disagree, disagree, neither disagree nor agree, agree, and strongly agree. For the second and third cohorts, the Likert scale was extended from five to six response options, replacing the neutral [neither agree nor disagree] with a choice of [slightly disagree] or [slightly agree]. In the first cohort, the tendency to choose the neutral response was strong, so extending to a six-point scale was intended to push the participants to show a preference. In order to compare their responses with those of the first cohort, the responses were analysed as a binary of either agree or disagree, and neutral responses were not included. The participants could also write comments about their experience with the activity. Furthermore, focus group interviews were conducted with several self-selected participants; these were audio recorded, and informative comments were noted to help further understand the learner experience of the activities. Additionally, the researcher/teacher made observations about how effective the learning activities seemed to be during the class sessions and how they could be improved going forward. The notes from both the focus groups and the teacher observations were used to complement the activity feedback survey results with more in-depth qualitative explanation.

The second cohort had fewer participants and lacked a control group; it served as the testing ground for the changes to the intervention based on the feedback from the first cohort using the action research feedback cycle. The third cohort of participants essentially experienced the same as the second cohort with minor improvements. Therefore, the survey responses regarding the intervention from the second and third cohorts were grouped together for the analysis resulting in two groups: first cohort, and second and third cohorts. I hypothesised that the responses from the participants from the latter cohorts would be more favourable for the activities where the changes had been made as part of the action research because in theory, the changes would improve the learning experience.

3. Results

In order to answer the first research question regarding whether a dialogue version of the fluency workshop is effective for learning multi-word expressions, the learning results from the intervention, such as multi-word expressions knowledge, use, and fluency were measured with pre- and post-intervention testing. When the results are compared between the first and third cohorts, it is clear that the results were better for the third cohort, which is likely due to the improvements made to the implementation as part of the action research cycle. The results from the first and third cohorts have previously been published (see Thomson 2018; Thomson et al. 2023); therefore comparisons will be made

between these two cohorts using Cohen's *d* effect size measure. For the first cohort, the quantitative results showed "...productive knowledge of target expressions had increased significantly more for the experimental group ($n = 44$) $M = 7.63$, $SD = 4.12$, than for the control group ($n = 29$) $M = 2.55$, $SD = 3.01$, $t(71) = -5.71$, $p = 0.000$, $d = 1.41$ " (Thomson 2018, p. 39). However, there was no significant increase in the use of the multi-word expressions in conversations or fluency when compared with the control group. The results for the third cohort were reported in Thomson et al. (2023) and showed a stronger increase in multi-word expression knowledge for the experimental group ($n = 65$) $M = 8.71$, $SD = 3.64$, than for the control group ($n = 51$) $M = 2.50$, $SD = 2.92$, $t(114) = 9.93$, $p < 0.001$, $d = 1.88$. The third cohort also showed a significant increase in the use of the multi-word expressions when compared to the control group, which was not present for the first cohort (Thomson et al. 2023). These results suggest that the changes made to the activities in response to the student feedback and teacher observations improved the learning outcomes for the multi-word expressions between the first and third cohorts. These comparisons with the control group suggest that the dialogue version of the fluency workshop is effective for learning multi-word expressions.

To answer the second research question about what aspects of the fluency workshop can be improved/made more appropriate for the Japanese context, the learners' feedback via the activity feedback surveys and focus group interviews, along with the teacher/researcher observations are considered. The learners' response to the introduction activity of listening to the audio and answering the gist questions were not surveyed as it was a fairly standard classroom activity setting the scene for the following activities. Additionally, marking pauses was removed after the first cohort because the dialogue short turn format meant that pauses were few and hard to notice. The focus group participants from the first cohort reported marking pauses where they thought that they would go since they could not distinguish many. Pearson chi-square and Fishers exact test (when numbers were below the expected limits for chi-square) were used to compare the learner activity feedback responses between the first cohort and latter cohorts (second and third cohorts). The participants responded in the fifth and sixth class sessions, after completing three units, so they had experienced each activity type three times before completing the survey (the numbers vary due to the variations in attendance and completion of the surveys). The comparisons between the responses from the first cohort and the latter cohorts regarding the remaining activities are reported below. Reference is made to the participant comments where relevant from the focus group interviews and comment sections of the survey.

The phrase instruction was not altered between the first and latter cohorts as 91.2% of the participants from the first cohort responded that it was a useful activity. Table 2 shows the response numbers and percentages for the phrase instruction. As expected, Fisher's exact test showed no significant difference between the first and latter cohort responses with both groups generally agreeing that the phrase instruction was useful for learning English (two-tailed $p = 0.755$).

Table 2. Comparison of the participant responses for the phrase instruction.

Phrase Instruction	Not Useful	Useful
First cohort	3 (8.8%)	31 (91.2%)
Second and third cohorts	11 (12.6%)	76 (87.4%)

For the shadowing activity, some focus group interview participants shared that they had not realised that they could stop and replay the audio as they liked; therefore, an explicit instruction was added for the latter cohorts to stop, start, and replay the audio as they needed. Initially, I had doubted the usefulness of shadowing, but the feedback from the focus group interview participants convinced me that the activity was appreciated by the students for learning the pronunciation and rhythm of the dialogue. It is therefore possible that I gave more confident and knowledgeable instructions in the latter iterations

of the workshop, which may have also improved the activity. Table 3 shows that 85.7% of the latter cohort agreed that shadowing was useful for learning English, and this was significantly more than the original cohort according to a Pearson chi-square analysis $X^2(1, N = 122) = 4.94, p = 0.026$.

Table 3. Comparison of the participant responses for shadowing.

Shadowing	Not Useful	Useful
First cohort	14 (31.1%)	31 (68.9%)
Second and third cohorts	11 (14.3%)	66 (85.7%)

Dictogloss was changed in response to feedback and observations from the first cohort. In particular, I observed that the small groups were not sharing well, and I attributed this to the initial work performed in pairs, which seemed to lead to uncomfortableness when expanding into a group of four for the final reconstruction. There also appeared to be hesitation to be the first to start sharing. Therefore, for the latter cohorts, the learners worked in small groups of four from the beginning of the activity to compare notes and reconstruct the text, with a five-minute timer for the final reconstruction. The students appeared more willing to share openly when having this arrangement from the beginning and worked to complete the task faster under the time pressure. The dictogloss worksheet was also simplified with numbers added to the instructions and also to the dialogue lines for reconstruction, providing the learners with a short reference system for comparing and sharing in their groups. Table 4 shows that 76.4% of the latter group agreed that the dictogloss activity was useful for learning English, and this was significantly more than of the original group: $X^2(1, N = 117) = 5.39, p = 0.020$.

Table 4. Comparison of the participant responses for dictogloss.

Dictogloss	Not Useful	Useful
First cohort	13 (46.4%)	15 (53.6%)
Second and third cohorts	21 (23.6%)	68 (76.4%)

Regarding the mingle jigsaw activity, the participants in the focus group interviews from the first cohort commented that it was hard to approach others in the class whom they were unfamiliar with, and some people did not seem interested in sharing. In the latter cohorts, a five-minute countdown timer was therefore displayed to motivate the action to collect all ten expressions within the time limit, much like a race. Table 5 shows that 79.1% of the participants from the latter studies agreed that the mingle jigsaw was useful for learning English. The Pearsons chi-square test showed that this was a statistical increase from the first cohort $X^2(1, N = 119) = 8.87, p = 0.003$.

Table 5. Comparison of the participant responses for the mingle jigsaw.

Mingle Jigsaw	Not Useful	Useful
First cohort	16 (48.5%)	17 (51.5%)
Second and third cohorts	18 (20.9%)	68 (79.1%)

There were no changes made to the roleplay activity because 72.5% of the first cohort responded that the roleplay was useful for learning English; furthermore, the focus group interview participants were satisfied with the roleplay as it was. Table 6 shows that 89.1% of the responses for the latter cohorts perceived the roleplay to be useful for learning English, which is statistically more than of the first cohort: $X^2(1, N = 132) = 5.74, p = 0.016$. Despite making no changes to the activity, it is entirely possible that, as the teacher, my instructions and confidence had improved with the experience and practice from the first cohort which therefore improved the learners' perception about how useful the roleplay was.

Table 6. Comparison of the participant responses for the roleplay.

Roleplay	Not Useful	Useful
First cohort	11 (27.5%)	29 (72.5%)
Second and third cohorts	10 (10.9%)	82 (89.1%)

The decreasing-time roleplay was intended to encourage the learners to feel time pressure and speak faster. In the first cohort, the time increments for the decreasing-time roleplay started at 120 s, decreasing to 90 s, and finally 60 s. The feedback from the focus group participants revealed that 120 s and also 90 s were longer than necessary. Therefore, the time increments were shortened to 100, 80, and 60 s for the café and directions roleplay. However, the focus group participants pointed out that the model dialogue for the hotel unit was longer and more challenging, and they struggled to finish this one within the time constraints. Therefore, I shortened the model dialogue for the hotel unit and added a realia booking form to remind the learners about the details that they had to request from their conversation partner. The time increments for the hotel unit were set longer at 120, 100, and 80 s for the third cohort. As the teacher, I also observed the learners substituting words and phrases when they could not recall the multi-word expressions, since they had no notes to refer to. In order not to lose the opportunity to use the multi-word expressions during this activity, I decided to display the multi-word expressions on the class screen for reference for the second and third cohorts. Table 7 shows that 86% of the participants in the latter cohorts (significantly more than of the first cohort) responded that the decreasing-time roleplay was useful for learning English; $\chi^2(1, N = 121) = 10.00, p = 0.002$.

Table 7. Comparison of the participant responses for the decreasing-time roleplay.

Decreasing-Time Roleplay	Not Useful	Useful
First cohort	14 (40.0%)	21 (60.0%)
Second and third cohorts	12 (14.0%)	74 (86.0%)

The purpose of recording audio is usually so that you can listen back to it. In the first cohort, the learners recorded their roleplays but were not given time in class to listen back to the complete recording. As the teacher, I assumed (incorrectly) that the students would listen back to the recordings after class. Without listening back to your recording, this activity lacks usefulness for learning. The feedback from the focus group interviews revealed that few students listened back to their recordings despite having access. Therefore, I created a worksheet for the participants in the latter cohorts to report their reflections on their fluency while listening back to their recordings during class time. Table 8 shows that 86.4% of the respondents from the latter cohorts found the roleplay recording activity useful for learning English, statistically, more than of the first cohort; $\chi^2(1, N = 125) = 13.14, p < 0.001$.

Table 8. Comparison of the participant responses for the roleplay recording activity.

Roleplay Recording	Not Useful	Useful
First cohort	16 (43.2%)	21 (56.8%)
Second and third cohorts	12 (13.6%)	76 (86.4%)

In the first cohort, for the related situation roleplay activity, I observed the participants taking a disproportionate amount of time deliberating with their partner about which scenario to roleplay, leaving almost no time to roleplay. Therefore, I reduced the scenario options from four to two, and added photos on the class screen as scaffolding to help the learners imagine the scenario and start the roleplay faster. Also, I added one minute

of planning time so that the learners could discuss their intentions using their L1 before starting the roleplay. Table 9 shows that 86% of the respondents from the latter cohorts considered the related situation roleplay activity to be useful for learning English. Again, statistically, this was significantly more than the first cohort; $\chi^2(1, N = 124) = 10.08, p = 0.002$. In the next section, these results will be discussed in relation to previous research and teaching implications.

Table 9. Comparison of the participant responses for the related situation roleplay activity.

Related Situation Roleplay	Not Useful	Useful
First cohort	15 (39.5%)	23 (60.5%)
Second and third cohorts	12 (14.0%)	74 (86.0%)

4. Discussion

The first research question asked to what extent a dialogue version of the fluency workshop (inspired by Wood 2009) would be effective for learning multi-word expressions. The adaptation and modification of the fluency workshop to a dialogue model for a Japanese EFL learning context, based on the results from this series of studies appear to increase learners' knowledge of multi-word expressions and use of multi-word expressions (Thomson 2018; Thomson et al. 2023). Another fluency-workshop-inspired study (McGuire and Larson-Hall 2018) also using a dialogue format reported fluency increases. However, in the current research, fluency did not appear to increase when compared with a control group. The difference between an increase or not in fluency could be related to the pre- and post-intervention speaking prompt; in the current set of studies, the learners were roleplaying a scenario, whereas in the McGuire and Larson-Hall (2018) study, the participants spoke about their personal likes and dislikes for the fluency measurement, which may have been easier to speak about versus an imagined scenario. Additionally, their participants had higher proficiency and were studying in the US, so they would likely have had more exposure to English use outside of the classroom than the EFL learners based in Japan in these studies. Perhaps using a monologue speaking task to measure fluency might give more insightful results in regard to fluency. Additionally, it could be argued that fluency in a conversation and fluency in a monologue are two different types of fluency as the conversational version requires some social skills to understand your interlocuter and respond appropriately while also encouraging continued dialogue. In any case, it would be interesting to test speech fluency via a single-person speaking task to ensure that increases in fluency are not missed due to the social dynamics involved in conversations.

The fluency workshop activities began with guided noticing of the multi-word expressions within the context of hearing and reading a model dialogue. The necessity of noticing transparent multi-word expressions has been questioned in the European context (Szcześniak 2024). While I would argue that these multi-word expressions are indeed transparent or fairly transparent, the activities in the fluency workshop preclude incidental learning and fairly explicitly draw attention to the multi-word expressions. Perhaps these particular multi-word expressions could be learnt simply through incidental encounters; however, the control group presents the result of leaving learning up to chance in the Japanese context, and the encounters with English are too few and far between to provide any guarantee of the learning of even fairly common multi-word expressions. Previous studies (Boers et al. 2006; Stengers et al. 2010) have found that noticing alone is unlikely to result in strong enough memory traces for learners to later recall the multi-word expressions without support. The fluency workshop intervention has gone beyond simply noticing and created stronger memory traces through practice as suggested by Boers and Lindstromberg (2012).

The fluency workshop activities included shadowing, which to some may seem like an unnecessary activity. However, the Japanese participants, with their L1, which sounds very different to English, insisted that shadowing was important because it helped them

to practise how to pronounce and say the multi-word expressions. The usefulness that the learners perceived for this activity supports the assertion that [Lin \(2019\)](#) makes for learning prosody when learning multi-word expressions.

The second research question asked what aspects of the classroom experience of the fluency workshop can be improved/made more appropriate for the Japanese context. The learners found the dictogloss ([Wajnryb 1990](#)) challenging but were helped by working to reconstruct the text in small groups of four people. The grouping was beneficial because if one person had not been able to take notes, someone else in the group would have taken some notes. Sharing in the group from the start worked better than changing from pairs to groups part way through. Also, the addition of a visible countdown timer for the final time to share and reconstruct the text appeared to help the learners overcome their fears about being the first to speak in their group. Similarly, the mingle jigsaw activity was difficult for the students to start as they had to decide whom to approach, and for some this resulted in holding back and not being as active as they had the potential to be. Simply displaying a visible five-minute countdown timer for the dictogloss and mingle jigsaw activities visibly sped up the initial interactions between the students as they were aware that they had limited time to complete a task. Raising the learners' awareness of the limited time available appeared to motivate the learners to interact with more vigour. It is possible that time pressure is more helpful with introverted learners, or those who are not so confident socially to approach others. I would also argue that the visual countdown keeps the task at the forefront, and conversations will be less likely to go off topic.

The time constraint also added a game-like element to the tasks, with some perhaps even seeing it as a race. This gamification of the activity appeared to speed up the students' activity, helping the students to increase their time on task (a concept which is generally associated with better learning outcomes) ([Godwin et al. 2021](#)). The visible time limit appeared to move the focus onto completing the task rather than feeling the fear of initiating a conversation. The use of a classroom timer to speed up the students' activity is not a new idea, and [Wurtele and Drabman \(1984\)](#) reported preschoolers speeding up their actions to tidy up their classroom after "beat the buzzer" was introduced. Using countdown timers has been credited with creating a sense of urgency in students to put in more effort and complete the task assigned ([Olipas and Luciano 2020](#)). Based on this experience, I would recommend using a countdown timer where teachers want to speed up student engagement and intensify engagement with the task.

Simplicity and the clarity of instructions and tasks was very important for quick transitions into each learning activity. For example, the instructions for the dictogloss were simplified from the original, and numbers were added on the worksheet to help the learners focus on what to do next. These number references likely also assisted with the communication between small-group members. During the reconstruction process, the learners could state a number reference to describe which part of the dialogue that they were talking about. Reducing the choice is another form of simplification to help the learners quickly move into the activity. For example, the related situation roleplay choices were reduced which led to faster decisions being made by students and likely resulted in more time interacting with the multi-word expressions. Simplifying the instructions or task helped the learners to spend more time using the multi-word expressions.

The provision of appropriate scaffolding helped the learners to activate their knowledge to use in interactions. For example, the related situation roleplay was not only simplified by reducing the scenario options, but photos aimed at activating the imagination, knowledge, and vocabulary for the scenario options were shown for the latter cohorts. The visual scaffolding appeared to help the learners imagine the scenario, remember the related words, and roleplay with more confidence. Both simplification and scaffolding helped to speed up the interaction and time on task, thereby increasing the time spent using the multi-word expressions for communication.

Further to the scaffolding point, an additional exposure to the multi-word expressions was added during the decreasing-time roleplay. Notes were hidden during this activity,

so if the learners were unable to recall the multi-word expressions, they would need to use other words to complete the roleplay. As the teacher, I observed the learners using alternative words during the original intervention when the multi-word expressions could not be quickly recalled. While trying to recall is part of the learning process, failure to recall and use of alternatives results in foregoing a chance to speak and communicate using the multi-word expressions. Therefore, in the latter interventions with the decreasing-time roleplay, multi-word expressions for the conversations were displayed on the class screen to help the learners recall and use them in their conversations. This additional opportunity to refer to the expressions would have increased the use of the expressions during the decreasing-time roleplay activity. The implication here is that if the goal of an activity is to use certain expressions, ensuring a reference point for those who are struggling will support greater use. The display of the expressions also raises the awareness of learners to try to use the expressions, rather than reverting to the path of least resistance, and using different words and expressions.

The perceptions of activity usefulness overall appeared to increase from the first cohort to the latter cohorts. The only activity that did not significantly change in terms of learner response to usefulness was phrase instruction. All others (including those where no substantial changes were made) showed significant improvements in the perception of usefulness between the first cohort and the latter cohorts. We can expect that, like many things in life, teaching improves with practice. In the latter iterations, I was more familiar with the activities, and with how the students were likely to respond based on the original intervention experience. The focus on learner feedback and teacher reflection resulted in greater improvements than would have been realised through simple repetition. The action research cycle provided structure to observe and actively seek feedback on my practice. Student feedback raised my awareness of the learner experience and thereby helped me to improve my instructions and the learning environment for the students.

In summary, the fluency workshop model using a dialogue format can improve learners' knowledge and use of multi-word expressions. The intense focus and time spent on learning and using multi-word expressions in this classroom model comes at the expense of other types of learning (such as wider vocabulary etc.). If learning goals are to improve use of common multi-word expressions and speaking confidence for particular situations then introducing this approach is likely to be beneficial. This research has provided evidence that action research can help teachers to improve the outcomes for learning multi-word expressions in the classroom. Implications for teaching multi-word expressions reported here may also apply to teaching English as a foreign or additional language more generally. Teaching is likely to improve with practice, especially if a conscious effort is made to observe and reflect on the classroom dynamic and feedback from learners. Overall, the time spent encountering and using the multi-word expressions can be maximised by reducing the time spent introducing or starting an activity. To this end, teachers may consider displaying a countdown timer and providing scaffolding assistance such as pictures to set a scenario or reference points to prompt the use of target multi-word expressions.

Of course, there are limitations which should be considered when interpreting the results of this report. As already mentioned, the original study had a five-point Likert scale while the improved studies had a six-point Likert scale, and the responses were classified as binary in order to make comparisons between the two scales. The original study participants were not forced to choose between [agree] and [disagree], and we cannot know how they would have chosen if they were. For example, there may have been more agreement or more disagreement, changing the comparison with the improved studies. Ideally, the studies should have had the same Likert scale to make the comparison more direct. Secondly, individual learning preferences may have influenced the responses in different ways compared to if the original respondents were to experience and respond to the improved activities.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance and approved by the Human Ethics Committee of VICTORIA UNIVERSITY OF WELLINGTON (Ethics approval: 22354, 24 November 2015).

Informed Consent Statement: Informed consent was obtained from all the subjects involved in the study.

Data Availability Statement: The original data presented in the study are openly available in [OSF] at https://osf.io/bvk8r/?view_only=c87b36efc09a412aa38e78762024ffd7 (accessed on 8 December 2023).

Acknowledgments: Appreciation is conveyed to all the participants who made these studies possible. Thank you to Averil Coxhead and Frank Boerswho supervised this Ph.D. research at Victoria University of Wellington. Thank you to the fellow researchers and anonymous reviewers who have asked pertinent questions and given helpful feedback along the way.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Frequency Profiles of the Multi-Word Expressions Used in the Fluency Workshop

Unit	Target Multi-Word Expression	Frequency Profile of Vocabulary	Occurrences per Million (COCA)	Function
Café	We have about an	K1	0.62	Information
Café	I usually only have	K1	0 (COCA), 1.2B (Google Search)	Information
Café	Have on the menu	Have on (K1), menu (K2)	0.05	Information
Café	What would you like	K1	3.30	Question
Café	But so does the	K1	0.07	Comparison
Café	What are you going	K1	11.25	Question
Café	I think I will	K1	0.71	Information
Café	I will have the	K1	0.13	Information
Café	Are you going to	K1	44.02	Question
Café	To get a drink	K1	0.19	Question
Directions	How do I get	K1	0.82	Question
Directions	Can you tell me	K1	5.03	Question
Directions	When you get to	K1	2.76	Instruction
Directions	Turn left at the	K1	0.03	Instruction
Directions	Until you get to	K1	0.35	Instruction
Directions	You should see the	K1	0.33	Instruction
Directions	Right-hand side of	K1	0.07	Information
Directions	Until I get to	K1	0.10	Information
Directions	Is on the right	K1	0.70	Information
Directions	Have a good day	K1	1.56	Greeting
Hotel	How many are in	K1	0.09	Question
Hotel	Could you repeat that	Could you (K1), repeat (K2)	0.15	Question
Hotel	How do you spell	How do you (K1), spell (K2)	0.14	Question
Hotel	How will you be	K1	0.12	Question
Hotel	That will be fine	K1	0.07	Information
Hotel	It is an extra	K1	0.03	Information
Hotel	Would you like to	K1	11.98	Question
Hotel	Let me think about	K1	0.22	Response
Hotel	The details are in	The, are, in (K1), details (K2)	0.03	Information
Hotel	Thank you very much.	K1	124.09	Response

References

- Arevart, Supot, and Paul Nation. 1991. Fluency improvement in a second language. *RELC Journal* 22: 84–94. [CrossRef]
- Boers, Frank, and Chau Thai. 2017. Repeating a monologue under increasing time pressure: A replication of thai and boers (2016). *The TESOLANZ Journal* 25: 1–11.
- Boers, Frank, and Seth Lindstromberg. 2012. Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics* 32: 83–110. [CrossRef]
- Boers, Frank, June Eyckmans, Jenny Kappel, H el ene Stengers, and Murielle Demecheleer. 2006. Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research* 10: 245–61. [CrossRef]
- Boers, Frank, Thuy Bui, Julie Deconinck, H el ene Stengers, and Averil Coxhead. 2023. Helping learners develop autonomy in acquiring multiword expressions. *The Modern Language Journal* 107: 222–41. [CrossRef]
- Burns, Anne. 2010. *Doing Action Research in English Language Teaching: A Guide for Practitioners*. London: Routledge.
- Davies, Mark. 2013. Corpus of Contemporary American English (COCA). Available online: <https://corpus.byu.edu/coca/> (accessed on 29 June 2022).
- EF English Proficiency Index. 2023. Signum International AG. Available online: <https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2023/ef-epi-2023-english.pdf> (accessed on 9 June 2024).
- Godwin, Karrie E., Howard Seltman, Ma Almeda, Mandi Davis Skerbetz, Shimin Kai, Ryan S. Baker, and Anna V. Fisher. 2021. The elusive relationship between time on-task and learning: Not simply an issue of measurement. *Educational Psychology* 41: 502–19. [CrossRef]
- Jiang, Jingyi. 2009. Designing pedagogic materials to improve awareness and productive use of L2 collocations. In *Researching Collocations in Another Language: Multiple Interpretations*. Edited by Andy Barfield and Henrik Gyllstad. London: Palgrave Macmillan UK, pp. 99–113.
- Koizumi, Rie, Toshie Agawa, Keiko Asano, and Yo In'nami. 2022. Skill profiles of Japanese English learners and reasons for uneven patterns. *Language Testing in Asia* 12: 53. [CrossRef]
- Lin, Phoebe M. S. 2019. Formulaic language and speech prosody. In *Understanding Formulaic Language. A Second Language Perspective*. London: Routledge, pp. 78–94.
- Loewen, Shawn, and Luke Plonsky. 2016. *An A–Z of Applied Linguistics Research Methods*. London: Palgrave Macmillan.
- Majuddin, Elvenna, Anna Siyanova-Chanturia, and Frank Boers. 2021. Incidental acquisition of multiword expressions through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition* 43: 985–1008. [CrossRef]
- Maurice, Keith. 1983. The fluency workshop. *TESOL Newsletter* 17: 29. Available online: <https://files.eric.ed.gov/fulltext/ED289347.pdf> (accessed on 17 October 2024).
- McGuire, Michael, and Jenifer Larson-Hall. 2018. Teaching formulaic sequences in the classroom: Effects on spoken fluency. *TESL Canada Journal* 34: 1–25. [CrossRef]
- Nation, Paul. 2007. The four strands. *Innovation in Language Learning & Teaching* 1: 2–13. [CrossRef]
- Olipas, Cris Norman P., and Ruth G. Luciano. 2020. Understanding the impact of using countdown timer on the academic motivation and computer programming anxiety of IT students: The case of a state university in the Philippines. *International Journal of Scientific & Technology Research* 9: 235–40.
- Peters, Ann M. 1983. *The Units of Language Acquisition*. New York: Cambridge University Press. Available online: <https://www.jstor.org/stable/413467?origin=crossref> (accessed on 23 March 2017).
- Peters, Elke, Eva Puim ege, and Paweł Szudarski. 2023. Repetition and Incidental Learning of Multiword Units: A Conceptual Multisite Replication Study of Webb, Newton, and Chang (2013). *Language Learning* 73: 1211–51. [CrossRef]
- Schmidt, Richard. 2001. Attention. In *Cognition and Second Language Instruction*. Edited by Peter Robinson. Cambridge: Cambridge University Press, pp. 3–33.
- Schmidt, Richard W. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11: 129–58. [CrossRef]
- Siyanova-Chanturia, Anna, and Ron Martinez. 2015. The idiom principle revisited. *Applied Linguistics* 36: 549–69. [CrossRef]
- Stengers, H el ene, Frank Boers, Alex Housen, and June Eyckmans. 2010. Does ‘chunking’ foster chunk-uptake? In *Fostering Language Teaching Efficiency Through Cognitive Linguistics*. Edited by Sabine De Knop, Frank Boers and Antoon De Rycker. Berlin: De Gruyter Mouton, pp. 99–117.
- Strong, Brian, and Frank Boers. 2019. The Error in Trial and Error: Exercises on Phrasal Verbs. *TESOL Quarterly* 53: 289–319. [CrossRef]
- Szcze niak, Konrad. 2024. The noticing hypothesis and formulaic language. Learnability of Non-salient language forms. *Acta Psychologica* 248: 104372. [CrossRef]
- Tavakoli, Parvaneh, and Takumi Uchihara. 2020. To what extent are multiword sequences associated with oral fluency? *Language Learning* 70: 506–47. [CrossRef]
- Thomson, Haidee. 2018. Building speaking fluency with multiword expressions. *TESL Canada Journal* 34: 26–53. [CrossRef]
- Thomson, Haidee, Averil Coxhead, Frank Boers, and Paul Warren. 2023. Increasing use of multi-word expressions in conversation through a fluency workshop. *System* 113: 102994. [CrossRef]
- Toomer, Mark, and Irina Elgort. 2019. The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning* 69: 405–39. [CrossRef]

- Vilkaitė, Laura. 2016. Formulaic language is not all the same: Comparing the frequency of idiomatic phrases, collocations, lexical bundles, and phrasal verbs. *Taikomoji Kalbotyra* 8: 28–54. [CrossRef]
- Vygotsky, Lev Semenovich. 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Wajnryb, Ruth. 1990. *Grammar Dictation*. Oxford: Oxford University Press.
- Webb, Stuart. 2007. The effects of repetition on vocabulary knowledge. *Applied Linguistics* 28: 46–65. [CrossRef]
- Webb, Stuart, Jonathan M. Newton, and Anna Chang. 2013. Incidental learning of collocation. *Language Learning* 63: 91–120. [CrossRef]
- Wood, David. 1998. *Making the Grade: An Interactive Course in English for Academic Purposes*. Toronto: Prentice Hall Allyn and Bacon.
- Wood, David. 2009. Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *The Canadian Journal of Applied Linguistics* 12: 39–57.
- Wray, Alison. 2004. 'Here's one i prepared earlier': Formulaic language learning on television. In *Formulaic Sequences Acquisition, Processing, and Use*. Edited by Norbert Schmitt. Amsterdam and Philadelphia: J. Benjamins, pp. 249–68.
- Wray, Alison, and Tess Fitzpatrick. 2010. Pushing learners to the extreme: The artificial use of prefabricated material in conversation. *Innovation in Language Learning and Teaching* 4: 37–51. [CrossRef]
- Wulff, Stefanie. 2019. Acquisition of formulaic language from a usage-based perspective. In *Understanding Formulaic Language: A Second Language Acquisition Perspective*. Edited by Anna Siyanova-Chanturia and Ana Pellicer-Sanchez. New York: Routledge, pp. 19–37.
- Wurtele, Sandy K., and Ronald S. Drabman. 1984. 'Beat the buzzer' for classroom dawdling: A one-year trial. *Behavior Therapy* 15: 403–9. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.