

Teaching Formulaic Sequences in the Classroom: Effects on Spoken Fluency

Michael McGuire & Jenifer Larson-Hall

Formulaic sequences (FS) are frequently used by native speakers and have been found to help non-native speakers sound more fluent as well. We hypothesized that explicitly teaching FS to classroom ESL learners would increase the use of such language, which could further result in increased second language (L2) fluency. We report on a 5-week study where students in a control group (n = 8) heard authentic English and practiced speaking and listening using a task-based approach, while students in a treatment group (n = 11) did the same but also focused on noticing and using FS found in weekly topic transcripts. Measures of speech rate (syllables per minute) and mean length of run (number of syllables found in the longest stretch with no pauses) served as objective measures of fluency. Sixteen native-speaker judges assessed excerpts from pre- and posttests for subjective fluency. The number of syllables of FS (expressed as a ratio of the total number of syllables) was counted by two judges. Results found large effect sizes for group membership in all measures, with the treatment group increasing FS use and fluency to a large extent and statistically outperforming the control group on most measures. We conclude that explicitly teaching formulaic sequences may lead to increased use of such phrases and also increased fluency.

Les locuteurs natifs emploient souvent des formules; celles-ci font également en sorte que les locuteurs non natifs semblent parler la langue avec plus de fluidité. Nous avons émis l'hypothèse selon laquelle l'enseignement explicite de formules en classe à des étudiants en ALS augmenterait l'emploi de formules, ce qui mènerait à une meilleure compétence en L2. Nous rendons compte d'une étude de 5 semaines pendant lesquelles des étudiants dans un groupe témoin (n = 8) ont écouté de l'anglais authentique et se sont pratiqués à parler et à écouter selon une approche basée sur les tâches. Les étudiants du groupe expérimental (n = 11) ont fait la même chose mais se sont également penchés sur le repérage et l'emploi de formules dans des transcriptions thématiques chaque semaine. Les mesures du débit de parole (syllabes par minute) et de la longueur moyenne des tronçons (nombre de syllabes dans le plus long tronçon sans pause) ont servi de mesures objectives de la fluidité. Seize juges-locuteurs natifs- ont évalué la fluidité subjective à partir d'extraits tirés de pré-tests et post-tests. Deux juges ont compté le nombre de syllabes des formules (exprimé sous forme de ratio du nombre total de syllabes). Les résultats ont révélé une importante ampleur des effets pour l'appartenance au

groupe avec toutes les mesures. L'augmentation de l'emploi des formules et de la fluidité chez le groupe expérimental a été notable et le rendement de ce groupe a été supérieur de façon significative à celui du groupe témoin pour la plupart des mesures. Nous concluons que l'enseignement explicite des formules pourrait mener à un emploi accru de ces expressions ainsi qu'à une meilleure fluidité.

KEYWORDS: fluency; formulaic sequences; second language instruction; second language acquisition; speech rate; explicit instruction

Formulaic sequences (FS) have long been recognized as an essential part of native speaker language (Pawley & Syder, 1983), and Conklin and Schmitt (2008) report that one-third to one-half of native speaker speech is formulaic. The application of such an important linguistic topic to second language learners is a research area that seems to be gaining in interest in recent decades, with enough research to enable a recent overview of experimental and intervention studies by Boers and Lindstromberg (2012). The general consensus from that overview is that although use of formulaic sequences by second language users is associated with better fluency, good methods of teaching language users to incorporate formulaic sequences in their own language have not yet been identified; implementing awareness-raising techniques that rely on incidental acquisition of such language chunks has overall not been shown to be effective. Boers and Lindstromberg (2012, p. 102) note that these findings may flow from the fact that “the instruments used to measure changes in participants’ repertoires of formulaic sequences were too blunt, or perhaps their statistical power was too low for detection of any but quite large effects.”

More research on FS is needed, but we think the available evidence suggests that it should be more oriented toward deliberate practice of FS. We believe that relying on incidental acquisition may itself be the problem, and that skill acquisition theory (DeKeyser & Sokalski, 1996; Muranoi, 2007) predicts that more success in speaking tasks may be obtained when participants practice actual speech production of FS.

The case is also being made persuasively that a greater use of formulaic sequences corresponds to higher ratings for fluency in a language, although the amount of empirical research providing evidence that use of formulaic sequences improves second language oral fluency is still small. Studies that have looked at the naturalistic correlation between proficiency and formulaic sequences have found that use of such sequences increases with higher proficiency and experience with the language. Wood (2007) examined four Japanese learners of English over a 6-month period of study abroad and found that the learners’ mean length of run (the number of syllables uttered between pauses) increased from 3.3 syllables at the first sampling to 4.1 by

the last (sixth) sampling. Concurrently, their ratio of number of FS used to number of runs increased from 30 to 37. Thus, fluency and use of FS increased somewhat over the study abroad period without any special intervention. Stengers, Boers, Housen, and Eyckmans (2011) also found modest correlations between number of FS and higher language learning proficiency in L1 Dutch and L2 English or Spanish. Forsberg (2010) looked at the association between proficiency and use of FS in L2 French and found that the higher the level of proficiency the learners had, the more conventional sequences they used. Research has shown that even if formulaic sequences are not explicitly addressed, students' use of them will naturally increase over time as they gain more experience with the language (MacKinnon, Lockwood, & Williams, 2004).

Incidental Acquisition of Formulaic Sequences

The next step in the path of studying how non-native speakers can take advantage of the power of FS for their own fluency is to ascertain whether the relationship between FS and fluency is causal, that is, whether helping students to use more FS leads to increased fluency and whether there are any effective steps teachers can take in the classroom to improve fluency and use of FS. Currently, the number of intervention studies that explicitly test whether instruction regarding formulaic sequences can deliberately increase the number of fixed expressions used in spontaneous oral productions is extremely limited, and, as noted by Boers and Lindstromberg (2012), mostly concern incidental acquisition as they focus on awareness-raising and awareness-directing. However, incidental acquisition of vocabulary has been shown to be extremely limited and much slower than explicit learning (Horst, 2005; Hulstijn, 2001), so from the outset one would not expect that any particular formulaic sequence would be learned and later incorporated into a language user's speech after simply coming across it in the course of reading or listening to language. In what follows we will review the literature on incidental FS acquisition and see that simply calling students' attention to FS without deliberate practice using them has not produced strong results.

Boers, Eyckmans, Kappel, Stengers, and Demecheleer (2006) tested the efficacy of teaching students to simply notice formulaic sequences as a means of improving their oral proficiency levels. Upper-intermediate to advanced level Dutch L1 students majoring in English participated in the study. Those in the experimental group ($n = 17$) were introduced to collocations and fixed expressions found in their academic reading materials and practiced identifying language chunks, while students in the control group ($n = 15$) used the same materials but never had their attention directed to FS over the course of 22 hours of instruction. However, neither group actually practiced using the FS in their oral productions. When students were tested on their oral English proficiency to speak about an article they had read in the course of the test,

the recordings of the experimental group were judged by two non-native EFL teachers, one of whom had conducted the actual interviews where participants spoke about the article. The experimental participants were asked to talk about the topic of the article that they read, and in doing so were observed to recycle many formulaic sequences that they encountered in the article as they produced their monologues, and they did so more frequently than those who had been in the control group. The experimental group was judged to subjectively sound more proficient and fluent than those of the control group, to the extent that the correlation between oral proficiency scores and the number of formulaic sequences used was $r = 0.43$, a medium-strength correlation accounting for 16% of the variation in scores (this is calculated by averaging the correlations from fluency scores by two judges and FS counts by two different judges found in Table 1 of Boers et al., 2006).

While Boers et al. (2006) showed that language users could be trained to be more aware of FS, a study by Stengers, Boers, Housen, and Eyckmans (2010) that replicated the Boers et al. (2006) study showed the limits of this type of incidental acquisition using Dutch L1 students ($N = 60$) learning either English ($N = 26$) or Spanish ($N = 34$). In this study students read an article that they were asked to retell for the end of term test, but instead of reading an article written in the L2 as was done in the original Boers et al. (2006) study, students read an article written in the L1. Boers and Lindstromberg (2012) report that in this case, there were no differences in “uptake of FS between the groups which had regularly engaged in text chunking and those which had not” (p. 89). The Stengers et al. (2010) study seemed to show that without immediate models, language users would not automatically generate FS in their speech even if they were aware of the importance of FS in general.

In sum, the two studies reviewed here seem to show that training language learners to notice FS in writing will help them pay attention to FS more than those who do not receive such training. However, such noticing will not result in actual increased use of FS in spontaneous conversation.

Deliberate Attention to Formulaic Sequences

Thus there appears to be a need to move beyond incidental acquisition to a more deliberate focus on FS, but as of yet, only a very small number of studies have introduced formulaic sequences to language users to try to raise awareness and then examine the results after language users practiced *using* the FS over an extended period of time. Taguchi (2007) conducted one such study. Taguchi asked beginning college learners of Japanese ($N = 22$) to memorize 37 simple grammatical constructions (such as *kuruma de*, “by car”) and practice using them in dialogues and highly controlled communicative drills. In describing the way classes were taught, Taguchi (2007, pp. 438–439) says, “The activities were tightly controlled with little room for varied responses, and free conversations or interviews were rarely incorporated.” Students were

recorded twice in a spontaneous oral conversation during the course of a semester, once after presentation of the Lesson 1 dialogue and then 5 weeks later, after the presentation of the Lesson 2 dialogue. Taguchi counted the number of chunks of language that the participants used in these conversations, and found that the learners used more chunks in the second spontaneous oral conversation ($M = 79, sd = 37$) than in the first one ($M = 33, sd = 15$). Students also used about double the range of chunks *accurately* in the second conversation ($M = 18, sd = 3$) compared to the first. Taguchi (2007) provides evidence that a very tightly scripted focus on a limited number of FS at the initial stages of language learning can result in gains if such chunks are repeatedly drilled and practiced in the classroom, but at such a beginning level one would expect language gains by any group. The gains by the participants are impressive, especially for a foreign language setting, but this study suffers from not including a control group.

Another study that introduced deliberate practice of FS was Taguchi and Iwasaki (2008), a replication of Taguchi (2007) that improved upon the original study by pretesting for students who already knew the chunks and excluding any who knew more than 5 of the 75 intended sequences. There are two groups in this study and one is labelled the control group, but it is not a control group in the sense of not working with FS; both groups, again beginning learners of Japanese, did the same type of practice with chunks as Taguchi (2007), with the difference being that the treatment group ($N = 22$) did a 5-minute structured conversation in class four times a week while the control group ($N = 20$) did not do any kind of conversation, but practiced the dialogue longer. As in Taguchi (2007), the participants were recorded in a spontaneous conversation once about halfway through the semester and then again about 7 weeks later, although the treatment that incorporated conversation practice only continued for 3 weeks. Thus, the focus of this study is not whether deliberate teaching of FS can improve their frequency of usage and thus increase speaker fluency, but rather on whether incorporating communicative practice with the FS for very short periods of time works better than repeating a dialogue. The quick answer is that it does for the use of FS but not for measures of fluency. The longer answer is that in the posttest, which was an unstructured conversation pretending to meet someone for the first time, results were measured in terms of number of chunks used (tokens), range of chunks used (types), fluency (measured by morae per minute with false starts, repetitions, self-corrected words, and English words excluded), and pause length (the total time of filled and unfilled pauses, longer than one second, found in the first 3 minutes of the student's conversation). For frequency of FS, the treatment group's mean improved substantially, but with much higher standard deviations in the second conversation 7 weeks later (Time 1: $M = 29, sd = 10$; Time 2: $M = 88, sd = 59$). The Cohen's d effect size for the comparison between groups, a standardized measure that indicates the size of the difference between groups in terms of standard deviations (calcu-

lated without any adjustment for correlation), was $d = 1.7$. This is a very large difference. The control group also improved by a large amount, but also with a much larger standard deviation at Time 2 (Time 1: $M = 28$, $sd = 19$; Time 2: $M = 58$, $sd = 48$). The effect size for the improvement of the control group was quite large at $d = 0.9$, meaning the improvement was of the magnitude of almost one standard deviation. The effect size for the difference between groups at Time 2 was Cohen's $d = 0.6$, a medium-size effect. An analysis of the gains between groups for speech rate found neither group improved statistically, but the experimental group improved statistically on pause length, which shortened by 5.95 seconds.

The Taguchi and Iwasaki (2008) study showed several things. First, it replicated the results of Taguchi (2007) in showing that focus on language chunks in the classroom where language learners are speaking and producing these FS resulted in an increase in the use of the FS in an unrelated spontaneous conversation. Second, it showed that adding in even a very small amount (just 5 minutes per class) of communicative interaction using a focus on FS resulted in moderate gains in the use of FS in an unrelated conversation. Third, it did not find that groups who practiced FS in speech spoke more rapidly. However, the group with conversation practice did reduce their use of pauses. This might have been simply due to practice with freer forms of conversation that the control group did not get, but it could have also been due to the increased use of FS. Erman, Forsberg Lundell, and Lewis (2016), in surveying two studies of fluency gains after study abroad, noted that "fluency gains are not due to speed but to longer and more complex utterances" (p. 129), and that may also be the case here in this situation of foreign language learning. This study adds valuable knowledge about the value of direct practice of FS in the classroom but, as with Taguchi (2007), suffers from not including a real control group that was not taught FS.

Up to this point we have not reviewed any studies that directly address the question of whether teaching students FS, in the sense of raising awareness of language chunks along with practicing using them in communicative interaction, increases the number of FS that students use in an unrelated conversation and also increases their fluency in that conversation over time. The Boers et al. (2006) study and its replication in Stengers et al. (2010) showed that raising awareness helped students to use more FS when they had access to written support for an oral conversation but not without it, and showed a link between use of FS and fluency, but fluency was measured only impressionistically. Taguchi's (2007) study showed that beginning language students could benefit from memorizing FS but did not provide a control group nor measure fluency, and used a very tightly controlled set of FS. Taguchi and Iwasaki (2008) did not test the claim that teaching FS would improve oral fluency and increase the use of FS, since both groups in the study learned FS, but this did seem to indicate that even very small amounts of practice using FS could help improve results on oral testing of FS.

Wood (2009) is the only study we are aware of that tests the hypothesis that teaching students FS and practicing them in communicative interaction will increase the number of FS that students use in an unrelated conversation and also increase their fluency, but this study was a case study of one L1 Japanese student studying abroad in Canada. The language learner participated in 6 weekly fluency workshops. Results showed that the participant made strong gains in speech rate and mean length of run (MLR) over 6 weeks, with speech rate increasing from 123 syllables per minute to 140 (a 14% change) and MLR increasing from 5.1 to 6.4 (a 26% increase). It is highly relevant to our own study, but cannot make more general claims about the effectiveness of this procedure. Thus we aimed to replicate this study's procedures in the current article.

Research Questions

The previous research leads us to ask the following questions:

1. Do students who are explicitly taught formulaic sequences in the context of a task-based lesson improve more than controls in two objective measures of fluency: speech rate and mean length of run?
2. Do students who are explicitly taught formulaic sequences in the context of a task-based lesson improve more than controls in subjective ratings of fluency by native English-speaking judges?
3. Do students who are explicitly taught formulaic sequences in the context of a task-based lesson make greater use of formulaic sequences than controls in their spontaneous production?
4. Are increases in measures of fluency and subjective ratings of fluency related to use of formulaic sequences?

Method

Participants

Students were recruited from an Intensive English program at a university in the United States. The program is designed as a university preparation course, and students who successfully complete Level 6 in a 0–6 level system can then enter the university at the undergraduate level without having to take the TOEFL test normally required for the admission of non-native English speakers. For this program, these students attended four to six hours of academic English classes per day, focused on writing and note-taking (little speaking was done in class). The students had no other classes outside of this program, but they were assigned several hours of homework every night. Forty students of mid-intermediate to advanced level (from Levels 3 and above in the 0–6 level system) initially joined the study; however, due to the intensive nature of their regular studies and the voluntary nature of our experiment, attrition was high. A total of 19 students completed the course.

There were 8 students in the control group, with an average age of 29 ($sd = 6$); 5 of them were males. There were 11 students in the treatment group with an average age of 26 ($sd = 8$); 3 of them were males. All students were from Thailand, Japan, or China, and both groups had a mixture of students from these countries. For individual participant data, see Appendix A.

Instruction

For this study, students attended a voluntary class that met for 30 minutes three times a week. Six weeks had been planned; however, two class sessions were cancelled due to an instructor being ill. Students could therefore have attended a maximum of eight hours of class. Students' information was retained if they attended at least four hours of voluntary classes altogether. The average amount of time that students in the control class attended class was not statistically different from the average amount of time students in the treatment class attended.

Each week of the course focused on a different recorded conversation about conversational topics (see Appendix B for a list of the topics; transcripts of the conversations are available upon request). An excerpt of the Week 3 topic, "Living Places," is given below (formulaic sequences are underlined).

- J: so when do you think you'll go back? Er, do you have definite plans for when you want to?
- K: um ... I don't really have definite plans
- J: yeah
- K: I have another year before I graduate ... so ... nothing's happening for a year
- J: yeah
- K: um..... I don't know, uh, I'm thinking about going back to either Japan or maybe Korea or Taiwan ... to teach
- J: that'd be cool
- K: but I'm also thinking of moving out to Portland ... so ... I don't know

The control group and treatment group were each taught by different instructors (the first author taught the treatment group and the second author taught the control group; both were aware of the purpose of the study), but both classes were based on a task-based approach to teaching (Willis, 1996). Lessons were task-based, with an introduction to the topic by means of a recording of two native speakers doing a task similar to the one we wanted the participants to do. The teacher highlighted useful vocabulary for students and let them do a series of tasks that would lead them to be able to do the task in question by the end of the 3 class periods. The students performed tasks in class in front of their peers for a presentation portion. Although having different teachers could have possibly made some difference to outcomes, both teachers were experienced and well-liked by students, and followed the

agreed-upon lesson plans closely. The first class session each week focused on listening to the sample conversations and was essentially the same for both control and treatment groups. The main differences were how these recordings were examined and expanded upon in the two other class sessions each week (explained below).

During the first session, students heard the conversation two to three times and answered some listening comprehension questions such as *When will K graduate?* New or unfamiliar words were introduced by the instructor, and the class reviewed the answers to comprehension questions together. This was followed by a speaking task or activity on the same topic as the example conversation. Students were not given a transcript of the conversation until the second session. The differences between the control and treatment group will be explained below.

Control Group Instruction

The control group received instruction based on the traditional vocabulary/grammar split focusing on isolated vocabulary and grammar. The students first read through the transcript with the teacher and reviewed isolated vocabulary and then listened to the conversation two more times. After this, the students were given an activity focusing on vocabulary and grammar while listening to the conversation, for example, noticing verb tenses and aspect. An example of this is when students were given contrastive examples of grammar and discussed the meaning of each one (*Where do you live?* vs. *Where have you lived?*). Students then used the topic in a speaking task (generally spending 10–15 minutes speaking).

The third 30-minute session each week focused on reviewing the topic with much paired partner speaking. Speaking tasks were supplemented by language analysis, in which the teacher would lead the students in reflecting on grammar points, vocabulary items, and mistakes made by the students, although this generally only took 5–10 minutes of class time. Students spent some time focusing on increasing fluency using the power-speaking task (see Appendix C).

Treatment Group Instruction

The treatment group received instruction with attention paid primarily to formulaic sequences found in the materials rather than the isolated vocabulary and grammar. This instruction was supplemented by additional formulaic sequences.

The first 30-minute session of each week was nearly identical to that of the control group, except in the handling of the speaking activity that concluded the first session each week. For this activity, the participants were given a few key formulaic sequences to utilize in their subsequent conversations, such as *I can't stand it when....* Speaking practice usually lasted about 10 minutes for this first class.

During the second 30-minute session of the week, the students were given the transcript for the conversation. The students read through the transcript, and the teacher reviewed the formulaic sequences discussed in the previous class. The students then listened to the conversation two more times while looking at the transcript. The instructor led the students in finding formulaic sequences that appeared in the conversation and exploring their usage. Further contextual examples of the formulaic sequences in use were pulled from the *Corpus of Contemporary American English* (Davies, 2008) and given to the students in handouts. The instructor supplemented the formulaic sequences found in the conversation with others that might be found in a similar conversation (such as ... *really gets to me* in conjunction with *I can't stand it when ...*). Overall, focus was placed on using the formulaic sequences as wholes, rather than examining the grammatical structure within. In fact, explicit grammar was not brought up at all during the entire course. About half of the course time was given over to speaking practice.

The third session each week focused on production. A variety of exercises and activities were used to reinforce the formulaic sequences learned in the previous class sessions. Activities to practice speaking, including some focus on improving fluency such as the power-speaking task, were used. Some examples of these activities can be seen in the lesson plans in Appendix C. The main difference from the control group in this third session was that students were expected to be producing the formulaic sequences they had learned as they did the speaking activities. Students generally spent 15–25 minutes speaking during this session.

Fluency Analysis Procedures

Students were given a pretest prior to instruction, and a posttest following the conclusion of the course. Both tests were done in pairs in a quiet classroom. In our study we decided to focus on speaking in a dialogic task, a much more common communicative event than the monologues used by other studies of fluency. A dialogue, however, can introduce extra variables into the equation where fluency is concerned, so this point should be kept in mind when looking at results. In preparing the data, the turns of each speaker were compiled into separate recordings, and a pause of 0.3 to 0.6 seconds was then placed between each turn to emulate natural pausing. The data were not trimmed inside the response of the students, meaning natural pauses were left intact. This editing process resulted in 38 recordings: pretest and posttest recordings for each of the 19 students. Each participant's recording therefore consisted of every utterance they had made in the original pretest or posttest, but without the rejoinder of their partner, although small noises of agreement or very short interruptions of the turn were left intact. The length of these recordings ranged from 1 minute and 9 seconds to 3 minutes and 29 seconds. The average length of the recording for the control group was 145 sec. ($sd = 39$) for the pretest and 165 sec. ($sd = 37$) for the posttest. The average length

for the treatment group was 164 sec. ($sd = 17$) for the pretest and 181 sec. ($sd = 16$) for the posttest.

The paired students were given a topic (likes and dislikes about the town where they are studying) about which they were asked to have a short conversation (around 5 minutes). The topic was the same for the posttest as for the pretest, which may have meant the students would naturally improve in their ability to speak about the topic, but this should not have benefited one group more than the other as neither received any particular speaking training or feedback about the topic. Of course, using the same topic might have encouraged gains in fluency in general from pretest to posttest for both groups. Participants spoke with different partners for the pretest and posttest, chosen randomly for the pretest pairing and then matched with a participant from the opposite group for the posttest pairing. In the case of the posttest, since there were more participants in the treatment group than the control group, sometimes teachers or volunteer students were asked to be speaking partners. The participants were left alone in a classroom, and the conversations were audio-recorded.

Fluency Analysis

Objective measures of spoken fluency—speech rate and mean length of run—were examined for changes from pretest to posttest. Finding an increase in either measure over time suggests an improvement in spoken fluency. Speech rate (SR) is measured as the number of audible syllables uttered per minute. Run length refers to how much a speaker is able to say before pausing. Here, MLR was calculated as the number of syllables spoken between pauses. Longer runs suggest that a speaker speaks more fluently than a speaker who pauses after shorter runs (Freed, 1995; Lennon, 1990; Möhle, 1984; Raupach, 1980). Following similar studies, any pause of 0.3 seconds or greater marks a break in a run (Wood, 2006).

We decided to add to the validity of the study by adding a subjective rating for fluency (perceived fluency) by 16 native-speaker judges, college undergraduates with no hearing challenges. Judges assigned a rating from 1 to 7 for each speaker (1 = *very disfluent*, 7 = *very fluent*). Before completing this task, the judges were given a 20-minute normative training session. The session consisted of instruction about criteria that are typically used to rate fluency in speech, such as rate of speech, pausing phenomena, and length of runs. The judges were then presented with examples of less fluent and more fluent participants whose speech was not included in the actual rating task. Judges were asked to rate these participants, and their answers were then discussed.

Formulaic Sequence Analysis

What constitutes a formulaic sequence and what criteria can be used to recognize one has been a subject of research and is somewhat elusive to define, but the overarching idea is that they are “semi-preconstructed phrases that con-

stitute single choices" made by speakers and writers (Sinclair, 1991, p. 110). Five particular criteria to identify formulaic sequences have proven useful and been used in previous studies (MacKinnon et al., 2004; Nattinger & DeCarrico, 1992; Wray & Namba, 2003): phonological coherence, recognizable collocations (*I don't know*), complexity or length of utterance, idiomaticity, and syntactically irregular phrases (which are nonetheless used, such as *by and large*). Wood (2006) emphasizes that formulaic sequences do not necessarily meet any particular combination of these criteria, and that they might feature one, multiple, or all of these criteria.

For the present study, each recording was analyzed by the first and second author, both native speakers of English, for formulaic sequences using the five criteria mentioned above. We relied more heavily on the phonological criterion than others, so that a phrase such as *the University Hub* could have been identified as a FS even though as a place name it may be seen as an open choice. Nevertheless, if it was a phrase that our participants were likely to utter often, we believed it could be a FS for them. Formulaic sequences in the model dialogues utilized in the classes were identified in the same way as those in the speech of participants. Identification of formulaic sequences was done separately by both authors, and agreement between the authors was high (95%). In order to take into account the fact that formulaic sequences are of varying lengths, we used a ratio of syllables found in formulaic sequences to total syllables in the recording (FS/S ratio), a measure of formulaic sequences that was also used in Wood (2009) and indicates what percentage of each speaker's speech is formulaic in nature.

Boers et al. (2006) note that since we cannot look inside a speaker's mind to determine what parts of an utterance are used as holistic chunks, in actual practice we will have to settle for a somewhat subjective measurement for what constitutes a FS. We believe this process may not be totally replicable from one study to another, given that judges may place different levels of emphasis on the five criteria listed above. One characteristic of the use of FS is that because they are retrieved and used as wholes, speakers who use them appear to speak more fluently (Boers et al., 2006). This raises the possible question of circularity in measuring increases in fluency for a group that has been trained to use more FS versus one without such training. If raters mark any hesitation-free word string as a FS, then there may be a spurious association between use of FS and fluency (since in this case FS and fluency essentially become one and the same). While it was not the case that every hesitation-free word string or run between pauses was counted as a FS, we admit that there may be an overlapping of these categories in our study, so the reader should keep this fact in mind.

Previous studies have used a variety of methods to try to deal with the problem of subjectivity in FS measurements. Boers et al. (2006) used a procedure with two judges not involved with the experiment who had read the literature on FS. These judges were asked to identify only correctly formed chunks of language, and only those they agreed on were used. Taguchi (2007)

transcribed the oral data and counted only the use of specific chunks of language that had been taught during the class, and Taguchi and Iwasaki (2008) used the same procedure. In Wood (2009), the author (presumably, as no particular judges are identified) identified FS in the speech of the participant using the five criteria listed previously in a “holistic manner” (p. 48) and non-grammatically-correct sequences were accepted as FS. Wood (2009) listed all of the FS that were used. This is not feasible for our study, but Appendix D lists a representative sample of FS. There appears to be no universally accepted manner of identifying FS in studies such as this, as so few exist currently, but for the future we suggest that three or more judges be used and only the FS they agree on be counted.

Following the approach taken in Wood (2009), grammatical accuracy was not considered an important variable to judging fluency, as the difficulty of a speech production task can lead to mistakes in articulation (Prodromou, 2007). For example, if a sequence of words such as “I have three roommate” was spoken with a continuous intonational contour, with no pauses between words and phonological reduction, we considered it a formulaic sequence even though it contains a grammar error and is not an idiom, collocation, or metaphor.

Results

Results from both the control and treatment groups for speech rate, mean length of run, perceived fluency by native speaker judges, and the ratio of formulaic syllables to total syllables were calculated. Statistical analyses were then conducted, with paired-sample *t*-tests conducted to examine improvement within groups over time, and independent-sample *t*-tests performed on gain scores to compare the improvement of the two different groups. In all cases data were examined for normality and no problems were found. An alpha level of .05 was adopted for statistical significance, although we also report confidence intervals and effect sizes, asserting that effect sizes are more important than statistical significance, as they can tell the reader how important the difference is rather than just that there is a difference (Larson-Hall & Plonsky, 2015). Descriptive statistics for all measures are found in Table 1.

For speech rate it is clear that the treatment group improved, while the control group’s average score actually decreased slightly, but notice that the control group’s speech rate started out higher. Students were randomly assigned to groups, although an eye was kept out for balancing levels of the students in an English-language program before they completed the pretest, and it was not anticipated that there would be any difference in proficiency levels between groups. However, it is possible that the treatment group started out weaker and thus had more room to improve than the control group. On the other hand, the differences in fluency measures are not so large in the context of the normal variation among speakers, even L2 speakers. For example, in Towell, Hawkins, and Bazergui (1996), university speakers of French as a

second language (with English as L1) were recorded describing a film they had seen, and their speech rates ranged from a low of 72 to a high of 198 syllables per minute. The same students also performed the same task in their native English and had speech rates ranging from a low of 134 to a high of 219 syllables per minute.

Table 1
Descriptive Statistics

<i>Control</i>	<i>SR</i>	<i>MLR</i>	<i>NSJ</i>	<i>FS/S</i>
Pretest	148.0 (26.0)	5.7 (0.9)	4.5 (1.19)	.29 (.14)
Posttest	146.7 (16.9)	5.4 (0.8)	3.8 (.84)	.27 (.09)
<i>Treatment</i>	<i>SR</i>	<i>MLR</i>	<i>NSJ</i>	<i>FS/S</i>
Pretest	130.3 (15.3)	4.8 (.98)	4.1 (.50)	.25 (.09)
Posttest	150.9 (15.5)	5.9 (1.01)	4.6 (.87)	.35 (.07)

Note. SR = speech rate. MLR = mean length of run. NSJ = native speaker judgement. FS/S = ratio of formulaic syllables to total syllables. Standard deviations appear in parentheses.

Speech Rate

The descriptive statistics show that speech rate for the control group decreased slightly while it increased for the treatment group. A paired-samples *t*-test for the control group found no statistical difference between performance on the pretest and posttest ($t_7 = 0.2, p = .09, 95\% \text{ CI: } -15.4, 18.1, d = 0.06$). A paired-samples *t*-test for the treatment group found a statistical difference between performance on the pretest and posttest ($t_{10} = -5.3, p = .0003, 95\% \text{ CI: } -29.2, -12.0, d = 1.3$). This means that there was no change in speech rate for the control group, while there was a statistically significant increase for the treatment group. Further, the effect sizes (Cohen's *d*) show how big the difference was. For the control group, a $d = 0.06$ means there was hardly any difference between pretest and posttest scores. For the treatment group, a $d = 1.3$ is a large effect size (Oswald & Plonsky, 2010) and means that the difference between groups was somewhat more than one standard deviation.

An independent-samples *t*-test comparing the treatment group to the control group on speech rate (with equal variances not assumed) using the difference between posttest and pretest as the dependent variable found a statistical difference between groups ($t_{11.1} = -2.7, p = .02, 95\% \text{ CI: } -39.7, -4.2, d = 1.3$). Again, the $d = 1.3$ shows that the effect size is large, although the confidence interval shows that the real difference in scores could be as small as only 4.2 points or as large as 39.7 points.

Following the advice of Larson-Hall (2017), we provide graphics that show all of our data for each individual and provide a guide to the average scores as well (graphics were obtained by inputting data at langtest.jp). Figure 1 shows a parallel coordinate plot of scores for speech rate. The lines represent individual speakers, and show graphically that for the control group, some participants increased modestly in speech rate while others decreased,

but for the treatment group all of the participants increased their speech rate, with some increasing quite a bit (there is just one exception to this generalization, whose decrease was slight). Figure 1 also contains a beeswarm plot (a boxplot overlaid with the actual data points) for each group to show the distribution of scores. It can be seen that for the control group, even though the median score did not change much, the range of scores got a little smaller. For the treatment group the range of scores stayed about the same but the median score went up.

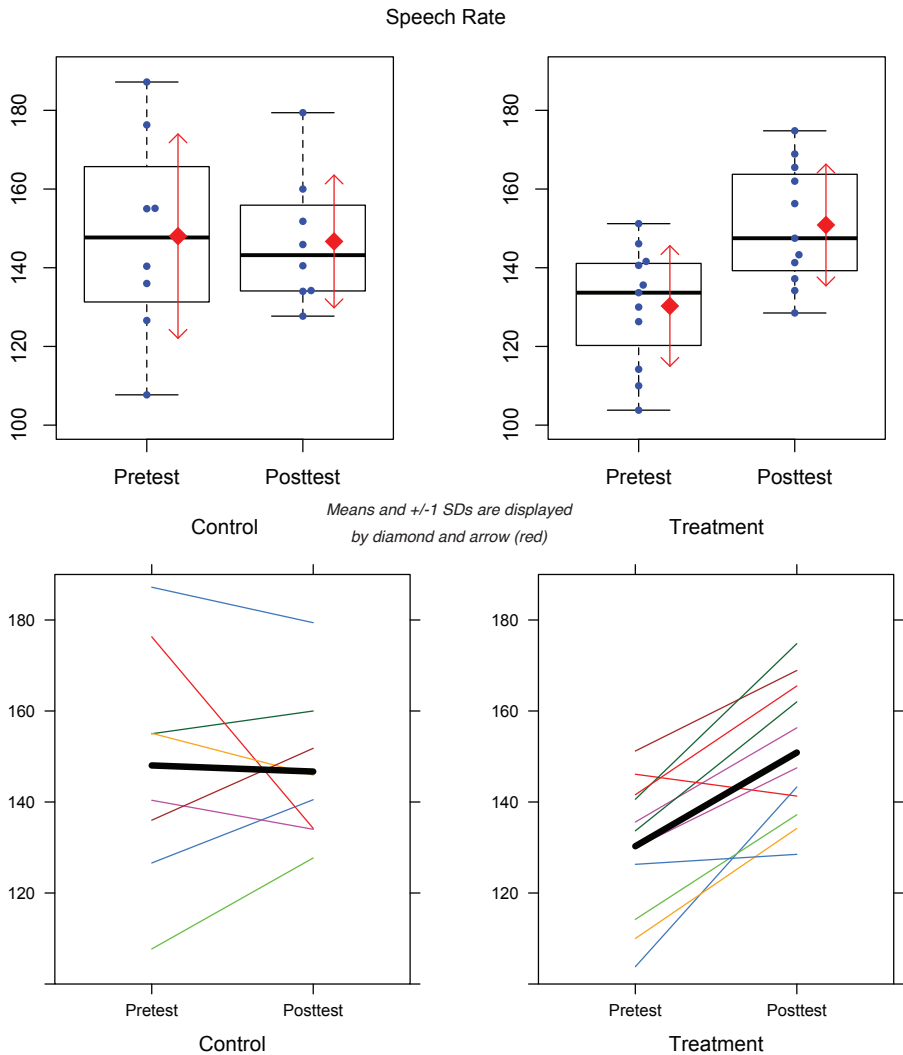


Figure 1: Top: Beeswarm plots for speech rate; Bottom: Parallel coordinate plots for speech rate.

Mean Length of Run

For the mean length of run, the treatment group overall improved by more than one syllable while the control group actually decreased. This means that participants in the treatment group were able to produce longer runs of speech between pauses, on average. A paired-samples *t*-test for the control group found no statistical difference between performance on the pretest and posttest ($t_7 = 1.04, p = .34, 95\% \text{ CI: } -0.45, 1.16, d = 0.17$). A paired-samples *t*-test for the treatment group found a statistical difference between performance on the pretest and posttest ($t_{10} = -3.5, p = .006, 95\% \text{ CI: } -1.7, -0.4, d = 1.1$). Similar to speech rate, the statistical tests here show that there was no statistically significant change in speech rate for the control group, while there was a statistically significant increase for the treatment group. The effect sizes once again show that the effect of the change was near zero for the control group, but near 1 standard deviation in improvement for the treatment group.

An independent-samples *t*-test on mean length of run (with equal variances not assumed) comparing the gains between groups from the pretest to the posttest found that the difference was statistical ($t_{15.6} = -3.1, p = .007, 95\% \text{ CI: } -2.4, -0.4, d = 1.4$). The effect size ($d = 1.4$) is large and indicates the groups differed by almost one and a half standard deviations of difference.

Looking at individual scores as seen in the parallel coordinate plot in Figure 2, it can be seen that most participants in the control group decreased or stayed the same in their run lengths, while in the treatment group almost all participants increased.

Native Speaker Judgement

Descriptive statistics for the native speaker judgements of fluency match the previous two sections' fluency analysis. The control group scores actually decreased in rated fluency from pretest to posttest, while scores for the treatment group increased (see Table 1). The Cronbach's alpha interrater reliability coefficient for the 16 judges involved in this study was high at 0.89.

Paired *t*-tests looking at improvement from pretest to posttest were both nonstatistical, however (Control group: $t_7 = 1.62, p = 0.15, 95\% \text{ CI: } -0.35, 1.81, d = 0.71$; Treatment group: $t_{10} = -1.41, p = 0.19, 95\% \text{ CI: } -1.19, 0.26, d = 0.26$). The effect sizes show a small effect size for the worsening scores of the control group ($d = 0.71$), but the treatment group's improvement is negligible ($d = 0.26$).

A *t*-test examining fluency rating differences between the groups (assuming unequal variances) was not statistical at the $\alpha = .05$ level ($t_{13.6} = -2.1, p = .051, 95\% \text{ CI: } -2.4, .0097, d = 1.0$), but with an effect size of $d = 1.0$, which is a medium-size effect showing that the group difference is around 1 standard deviation, we will argue that the differences between groups here are important and would be statistical with a larger group size (Larson-Hall, 2010).

Formulaic Sequence Use

The results so far have shown that the treatment group improved their fluency over the course of the 5-week period and statistically differed from the control group. However, was this due to the fact that the treatment group used more formulaic sequences in their spontaneous productions? Results of our analysis of formulaic sequences resulted in a measurement representing the number of syllables that were involved in formulaic sequences divided by the total number of syllables the participants produced (we label this as FS/S).

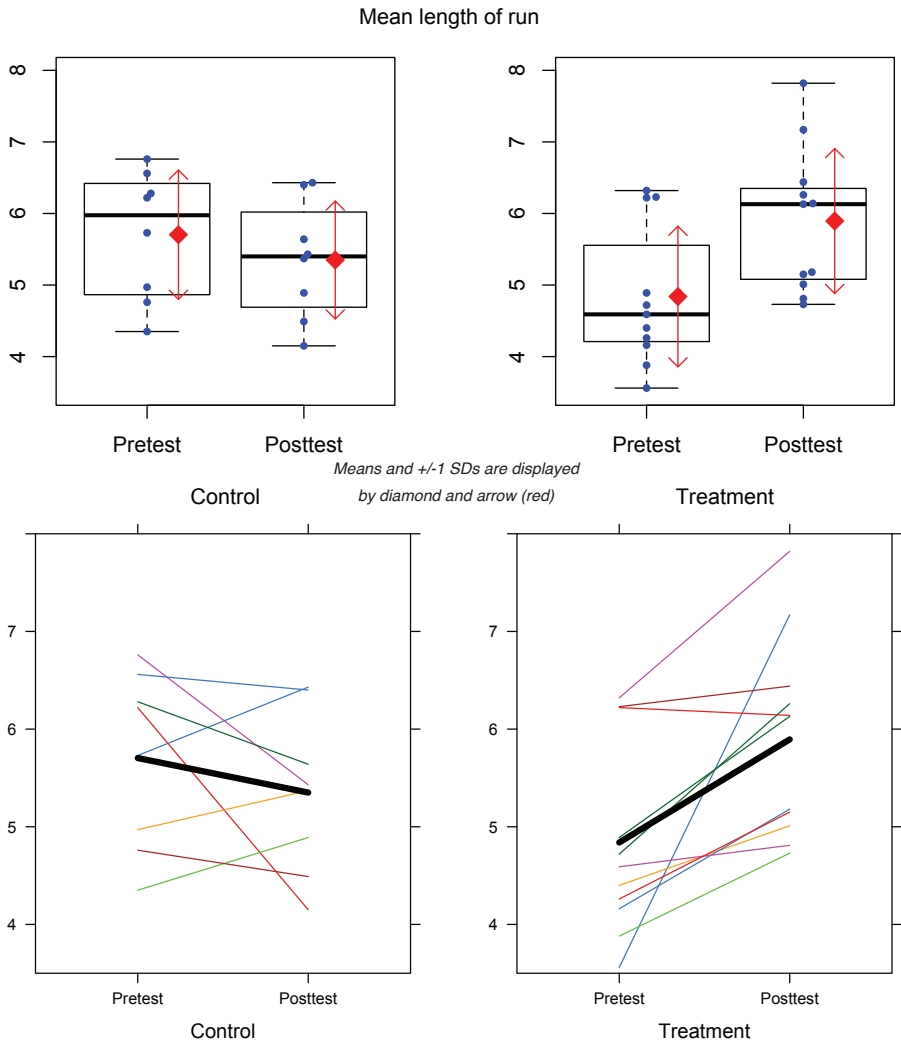


Figure 2: Top: Beeswarm plots for mean length of run; Bottom: Parallel coordinate plots for mean length of run.

The descriptive statistics in Table 1 show a slight decrease in the FS/S for the control group, and a fairly large increase in the FS/S for the treatment group.

A paired-samples t -test for the control group found no statistical difference between performance on the pretest and posttest ($t_7 = 0.25$, $p = .8$, 95% CI: -0.11, 0.13, $d = 0.2$). A paired-samples t -test for the treatment group found a statistical difference between performance on the pretest and posttest ($t_{10} = -4.6$, $p = .0009$, 95% CI: -0.15, -0.05, $d = 1.2$). This means the treatment group used more formulaic sequences (as identified by the study authors) on the posttest, with a large difference ($d = 1.2$) between pretest and posttest, while the control group essentially did not change in their use of FS.

An independent-samples t -test (with equal variances not assumed) comparing the gains between groups from the pretest to the posttest on this ratio of formulaic sequences found that the difference was not statistical ($t_{9,6} = -2.1$, $p = .07$, 95% CI: -0.01, 0.10, $d = 1.0$). Again, although this result is not statistical given an $\alpha = .05$ alpha level, the effect size (which is not affected by sample size) shows that this is an important result with a medium-size effect.

Figure 3 shows individual results for the use of formulaic sequences. Although a few students in the control group did improve significantly, some more than doubling their use of formulaic language, these improvements were not across the board, as in the treatment group. The results of the control group were very disparate, in fact, with half of the class dropping in their formulaic sequence usage in the post test.

Relationship Between Formulaic Sequence Use and Fluency Measures

To examine the question of whether increased use of formulaic sequences resulted in higher fluency, we looked at the correlation in the posttest between the use of formulaic sequences (FS/S) and the fluency measures of speech rate (SR) and mean length of run (MLR). The correlation for all participants between FS/S and MLR is $r = 0.28$, and the correlation between FS/S and SR is $r = 0.54$. The effect sizes for both correlations are respectable; generally an $r = 0.28$ would be seen as a medium-size effect and an $r = 0.54$ as a large effect (Cohen, 1992). In sum then, the use of formulaic speech in general plays a moderate role in influencing how fluently a non-native speaker speaks English as a second language.

Discussion

This study points to the possibility that explicitly teaching formulaic sequences in the classroom can be effective. This supports the observations from Wood's (2009) case study, which was the basis for the present study, and complements the results from Taguchi and Iwasaki (2008), which found that even a small amount of speaking practice (5 minutes per class) can be quite helpful in improving use of FS. While task-based teaching is a useful method for helping students receive large amounts of input and output and

gain speaking practice, this study seems to show that incorporating a healthy component of attention to formulaic sequences in the classroom might help students in a second language environment use such formulaic sequences more frequently and thus improve in their speaking fluency.

Our first research question asked whether students in the treatment group would improve on fluency measures, and whether this improvement would

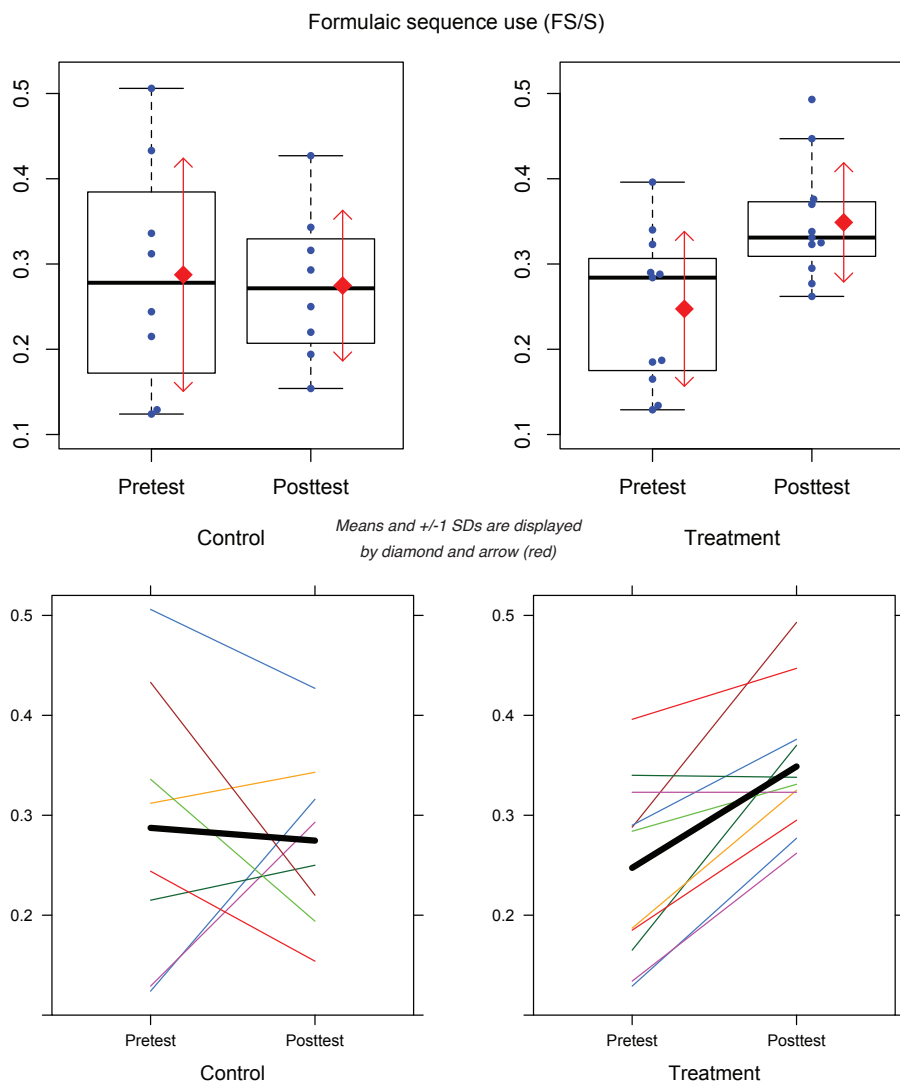


Figure 3: Top: Beeswarm plots for FS/S; Bottom: Parallel coordinate plots for FS/S (the percentage of formulaic sequences used in speech).

be greater than in the control group. Numerically and statistically, the treatment group improved from pretest to posttest and exceeded the control group in both objective measures of speech rate and mean length of run, with fairly large effect sizes in every case ($d = 1.1$ to 1.4). This could indicate that the explicit teaching of FS helped increase speech rate and mean length of run for the treatment group, although we noted that this group started out with weaker scores than the control group, which may have made improvement easier. In addition, all of the students were studying abroad, so improvements found in this study may not apply to situations where students are not getting as much input in their L2.

Our second research question concerned the judgements of native speakers as to the participants' fluency. We asked whether receiving explicit FS training would help raise fluency scores as perceived by the NS. Here, neither group made any statistical improvement over the course of the semester, with there being apparently too much variation to conclude that either group statistically improved or, conversely, got worse. It may be that seven levels of gradation for fluency was too difficult to judge; if the task had been to compare the pretest and posttest fluency and decide whether it had improved, decreased, or stayed the same, perhaps there would have been more certainty in the scores.

Our third research question asked whether participants would increase their use of FS from pretest to posttest if they were explicitly taught about FS and practiced using them in a speaking context. Here we found that the control group did not improve in their use of FS while the treatment group did, and there was a medium-size difference between the groups. It appears that the treatment group did improve in their ability to use FS in speaking.

Our fourth research question was whether increased fluency was related to the use of FS. Correlations also showed there was a medium relationship between all participants' use of formulaic sequences and their fluency measures. If this result can be trusted, it indicates that explicit teaching of FS can help students use more of them in spontaneous speech, and that this use leads to increased fluency in spontaneous speech.

However, while the results of the present study are in general positive and indicate that explicit teaching of FS may be responsible for improving speech rate and mean length of run, there are a number of limitations to our study and possible points for improvement in future research. The most important is the procedure for identifying FS. This is still an area of debate and, unfortunately, our focus on phonological coherence to judge FS may have rendered the identification of FSs and the measure of fluency somewhat circular. We suggest that future studies employ at least three judges who have studied the FS literature to identify FS, and to only accept FS that all judges agree on. Another issue is that we chose to use subjective native-speaker judgements in order to supplement the objective measures of fluency. This still seems like a good idea, but in the future we recommend giving judges a compari-

son to work with, and to narrow their choices about whether a speaker has improved or not.

The limited scope of our study should be pointed out as well. First, the sample sizes were small, so care should be taken to not generalize too far. Second, the participants were all living in the United States, and improvements in their fluency and use of FS were to be expected in the natural course of their intensive contact with the L2. Third, the format of the pre- and posttests was a conversation as opposed to a monologue, which may have influenced scores in a number of ways. Fourth, while the treatment group focused on FS, the control group may have benefited in other ways, such as increased accuracy, that were not taken into account for this study.

Acknowledgements

The authors kindly thank the participants in both of their volunteer English classes for enthusiastically participating in their experiment, and also the University of North Texas IELI for the use of classrooms and other support for the volunteer classes.

The Authors

Michael McGuire teaches English at Kansai Gaidai University. He is interested in finding ways to help students improve their language autonomously, including using parallel (bilingual) corpora as a help for writing in the English classroom.

Jenifer Larson-Hall teaches second language acquisition at Kitakyushu University in Japan. She hopes more authors will consider the value of statistical graphics such as those found in this article and refers interested readers to her 2017 article "Moving Beyond the Bar Plot and Line Graph to Create Informative and Attractive Graphics."

References

- 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/10.1191/1362168806lr1950a>
- Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics*, 32, 83–110. <https://doi.org/10.1017/S0267190512000050>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, 29(1), 72–89. <https://doi.org/10.1093/applin/amm022>
- Davies, M. (2008). *The corpus of contemporary American English (COCA)*. Retrieved from <https://corpus.byu.edu/coca/>
- DeKeyser, R., & Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning*, 46(4), 613–642. <https://doi.org/10.1111/j.1467-1770.1996.tb01354.x>
- Erman, B., Forsberg Lundell, F., & Lewis, M. (2016). Formulaic language in advanced second language acquisition and use. In K. Hyltenstam (Ed.), *Advanced proficiency and exceptional ability in second languages* (pp. 111–147). Boston, MA: De Gruyter.
- Forsberg, F. (2010). Using conventional sequences in L2 French. *IRAL-International Review of Applied Linguistics in Language Teaching*, 48(1), 25–51. <https://doi.org/10.1515/iral.2010.002>
- Freed, B. F. (1995). Language learning and study abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 3–33). Philadelphia, PA: John Benjamins.

- Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *Canadian Modern Language Review*, 61(3), 355–382. <https://doi.org/10.3138/cmlr.61.3.355>
- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge, UK: Cambridge University Press.
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.
- Larson-Hall, J. (2017). Moving beyond the bar plot and the line graph to create informative graphics. *Modern Language Journal*, 101(1), 244–270. <https://doi.org/10.1111/modl.12386>
- Larson-Hall, J. & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65(51), 127–159. <https://doi.org/10.1111/lang.12115>
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387–417. <https://doi.org/10.1111/j.1467-1770.1990.tb00669.x>
- MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4
- Möhle, D. (1984). A comparison of the second language speech production of different native speakers. In H. W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp. 26–49). Tübingen, Germany: Gunter Narr Verlag.
- Muranoi, H. (2007). Output practice in the L2 classroom. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 51–84). Cambridge, UK: Cambridge University Press.
- Nattinger, J., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford, UK: Oxford University Press.
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85–110. <https://doi.org/10.1017/S0267190510000115>
- Pawley, A., & Syder, F. (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). New York, NY: Longman.
- Prodomou, L. (2007). Bumping into creative idiomaticity. *English Today*, 23(1), 14–25. <https://doi.org/10.1017/S0266078407001046>
- Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 263–270). Berlin, Germany: De Gruyter Mouton. <https://doi.org/10.1515/9783110816570.263>
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2010). Does “chunking” foster chunk-uptake? In S. De Knop, F. Boers, & A. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 99–118). Berlin, Germany: Mouton de Gruyter.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2011). Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *IRAL-International Review of Applied Linguistics in Language Teaching*, 49, 321–343. <https://doi.org/10.1515/iral.2011.017>
- Taguchi, N. (2007). Chunk learning and the development of spoken discourse in a Japanese as a foreign language classroom. *Language Teaching Research*, 11(4), 433–457. <https://doi.org/10.1177/1362168807080962>
- Taguchi, N., & Iwasaki, Y. (2008). Training effects on the development of oral fluency in L2 Japanese. *Japanese Language and Literature*, 42(2), 413–437.
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17(1), 84–119. <https://doi.org/10.1093/applin/17.1.84>
- Willis, J. (1996). *A framework for task-based learning*. Essex, UK: Longman.
- Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review*, 63(1), 13–33.

- Wood, D. (2007). Mastering the English formula: Fluency development of Japanese learners in a study abroad context. *JALT Journal*, 29(2), 209–230.
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, 12(1), 39–57.
- Wray, A., & Namba, K. (2003). Formulaic language in a Japanese-English bilingual child: a practical approach to data analysis. *Japanese Journal of Multilingualism and Multiculturalism*, 9(1), 24–51.

Appendix A. Detailed Participant Information

	Participant	L1	Age	Level in institute
Control Group	DC (M)	Chinese	28	6
	GC (M)	Chinese	40	Institute grad
	KC (F)	Chinese	25	5
	NC1 (M)	Japanese	29	3
	NC2 (M)	Thai	25	5
	PC (F)	Thai	34	6
	TC (M)	Chinese	30	5
	VC (F)	Chinese	23	4
Treatment Group	AE (F)	Chinese	24	Institute grad
	JE1 (M)	Thai	43	5
	JE2 (F)	Chinese	19	5
	KE1(F)	Chinese	20	5
	KE2 (M)	Japanese	21	4
	LE (F)	Chinese	24	4
	ME (F)	Chinese	25	5
	NE (F)	Thai	26	5
	OE (M)	Thai	29	5
	SE (F)	Japanese	19	6
	XE (F)	Chinese	39	4

Appendix B. Native-Speaker Conversation Topics

Week 1 – “So what do you do?”

Week 2 – “Where are you from?”

Week 3 - “Living places”

Week 4 – Movies, etc. (talking about personal taste in music, movies, etc.)

Week 5 – Bad jobs (talking about things one doesn’t like doing or would not be good at doing)

from <http://www.ello.org>

Appendix C. Sample Lessons Plans

Experimental Group Week 4 - "Movies"

<i>Day 1</i>	<i>Listening and comprehension</i>	<i>Materials</i>
Warm up – 5 min.	"Do you like horror movies?"	
Listening – 8 min.	Give students listening comprehension questions Listen to conversation twice Go over comprehension questions Listen third time	Handout
Movie Collocations – 7 min.	Go over movie genre collocation list	Handout
Speaking – 8 min.	Conversations in pairs or groups of 3: "Do people in your country speak like J (many fragments, speaking a lot but not actually saying much)? What are some things people might do in your language that are similar?"	
<i>Day 2</i>	<i>Formulaic language and practice</i>	<i>Materials</i>
Warm up – 7 min.	"What is the most difficult part of conversation with a native English speaker?" Discuss together as a group (native speakers talk too fast? They blend words together? Etc.)	
Listening	Listen once with transcript Go over formulaic sequences found in the conversation and practice with students	Transcript Handout
Speaking – 8 min.	Conversations in pairs or groups of 3: "What kind of movies do you like most? Are there any movie genres you hate?" Using formulaic sequences from handout	
<i>Day 3</i>	<i>Practice</i>	
Practice – 15 min.	Power speaking* – movie topics "Do you like movies from your country?" "What kind of movie would be bad for a date?" "Do you prefer to watch movies at home or in a theater?"	
Practice – 15 min.	Collocation game**	

***Power-speaking:** Students were placed in pairs and given a generic opinion topic. They were given 30 seconds to say as much as they could about the topic before switching partners and receiving a new topic. Students were encouraged to speak as quickly as they could and to not worry about making grammatical mistakes. This activity is designed to focus on increasing speech rate and reducing pauses while speaking.

****Collocation game:** Students were divided into two teams and presented with a core noun that featured a high number of collocations in the corpus (see list, “Collocation Game Key”). Teams were given two minutes to write down as many collocations as they could come up with. After this, the teacher would introduce additional strong collocations and discuss their use with the students. This activity was designed to make students consider the many different forms that collocations might take (adjective + noun, verb + noun, ...), and to increase their awareness of collocation so that they might discover such lexical items in the texts they encountered in their regular classes.

Collocation Game Key:

Key Word	Special	Top 5
Test	Bomb a test	Pass a test
	Standardized test	Big test
	Ace a test	Fail a test
	Drug test	Take a test
	Test drive	Have a test
Food	Junk food	Fast food
	Food chain	Chinese food
	Food supply	Organic food
	Food shortage	Healthy food
	Serve food	Favorite food

Appendix D. A Sample of Formulaic Sequences Identified in Participant Recordings

this is my first time to come here; a month ago; I went to the; in the square; coffee shop; a lot of people; maybe not; I remember one time; winter break; international student; I just came here; half an hour; I don't have; even though; I think it's; how 'bout you; for a trip; for example; it's a lotta; because I think; my major is business