

# A Comparative Analysis of the Processing of Contextual Formulaic and Nonformulaic Sequences by Adult and High School Japanese Learners of English

Hiroko NISHIMURA

Keywords: contextual formulaic and nonformulaic sequences, processing, accuracy, speed

## 概 要

英語定型表現は、英語の運用能力、コミュニケーション能力の向上に重要とされ、近年注目されている。本研究では、大人と高校生の日本人英語学習者を対象に、英語定型・非定型表現の処理を比較検討した。その際、英語定型・非定型表現の文の中への意味的適合性を判断してもらった課題を行い、判断の正確性と速度を収集した。大人の参加者は、課題の正答率により3分割し、上位・下位群を対象として分析を行った。その結果、全ての群（大人上位・下位群、高校生）の判断速度において刺激間で有意差が確認され、定型表現が全体処理され文脈全体の意味処理を促進していることが示唆された。その一方で、判断正確性は、英語定型表現が非定型表現よりも有意に高いこと示されたのは、大人の上位群のみであった。大人下位群・高校生は、英語定型表現の知識の正確性が低く不安定であることが示された。

## 1. Introduction

In Japan, the stated objective of formal English education is to acquire communicative competence in the language (the Ministry of Education, Culture, Sports, Science and Technology [MEXT], 2018). In 2011, foreign language activities classes (English activities classes) were implemented for fifth- and sixth-grade students (MEXT, 2018). Since the beginning of the academic year in 2020, English activities classes and foreign language classes (English classes) were made mandatory for pupils in the third and fourth grades and those in the fifth and sixth grades, respectively (MEXT, 2013). Consequently, English education in Japan has entered a revisionist phase. However, due to the English as a foreign language (EFL) circumstance,

it is a challenge to accomplish this educational objective.

Recently, formulaic sequence (FS) knowledge—which is supposed to account for the large amount of English used by native English speakers and to reduce the language processing load—has been gaining attention as an important factor that can help Japan to achieve its educational objectives. Although much research has been conducted on FS processing, there is a lack of studies of FS processing in contexts (Conklin & Schmitt, 2008). Since FSs are used in contexts, there is a need for research on FS processing that is contextually appropriate. Hence, this study aims to explore the status of contextual FS knowledge by comparing it with contextual nonformulaic sequence (NFS) knowledge in the mental lexicon of Japanese EFL learners who have different English proficiency levels and educational backgrounds, as well as the implications of this knowledge for the teaching and learning of FSs.

## 2. Literature Review

### 2.1 Formulaic Sequences

#### 2.1.1 Definition of formulaic sequences

In order to establish standard approaches in the identification of FSs, there is a need for research using an enormous database of multiword units. Consequently, in 1964, the first computer-based corpus, the “Standard Corpus of President-Day Edited American English,” commonly known as the “Brown Corpus,” was published (Saito et al., 2005). This triggered the development of numerous corpora with which

scholars significantly enhanced FS research (e.g., Wray, 2002). Corpus-based research sheds light on patterns of language use and suggests that language is stored as chunks, set phrases, or collocations rather than as individual words or sounds. Meanwhile, scholarly discussions of to apply the thresholds of FSs—based on their frequency count in the corpora—are underway. Additionally, definitions of FSs vary among researchers since they have used diverse criteria to identify FSs according to the focus of their research, such as psycholinguistics, lexicology, and second language acquisition (Moon, 1997). Wray (2002) proposed that FS is (or appears to be) a prefabricated sequence—be it continuous or discontinuous—of words or other elements that is supposed to be stored and retrieved as a whole from memory in language use. Consequently, FSs are considered less susceptible to the rules of grammar in language use (p.9). Wray's definition of FS is cited by most, and is also supported by this study.

Due to different definitions of FSs, the classification of FSs also varies. For example, Moon (1997) has classified FSs into five categories: compounds, phrasal verbs, idioms, fixed expressions, and prefabs. Among them, phrasal verbs are considered particularly problematic in second language (L2) teaching and learning (Sinclair, 2004), and their commonality is often cited as a reason. Typically, phrasal verbs comprise monosyllabic verbs and adverbial or prepositional particles, which give the impression that they are very familiar words. Consequently, they are less likely to attract the attention of learners. Additionally, phrasal verbs are fixed and have meanings unpredictable from those of their component words (Sinclair, 2004). In practice, phrasal verbs and prepositional verbs—such as *carry NP out*, *look forward to*, and *get out of*—are the most common multi-word expressions (Biber et al., 1999). Therefore, this study takes such sequences—which include verbs—as target stimuli.

### 2.1.2 Significance of formulaic sequences

Pawley and Syder (1983) were the first researchers in English to identify the importance of conventionalized language. Based on a hypothesized holistic representation of FSs, four interrelated benefits can be expected: enhancement of fluency, accuracy, and communication, and reduction in cognitive load in language processing. In terms of enhancing fluency and accuracy, FSs are processed more quickly and easily than other sequences comprising the same number of words, since they are supposed to be stored holistically and retrieved as whole units from our minds. Consequently, FSs are generally spoken more fluently and with a coherent intonation contour (e.g., Schmitt & Carter, 2004). Mastery of memorized FSs frees us from the task of constructing expressions word-by-word to attend to other parts of the language activity, leading to greater fluency and accuracy (e.g., Pawley & Syder, 1983). Moreover, FSs have advantages in written language. In their examination of eye movements in native and non-native speaker of English reading short passages in English, Underwood et al. (2004) reported that FSs required less fixation and shorter duration than the same words in nonformulaic expressions. Moreover, the study by Jiang and Nekrasova (2007) reported that L2 learners processed FSs more quickly and accurately than NFSs.

In terms of enhancing communication, FSs are ubiquitous (Nattinger & DeCarrico, 1992) and make up a large proportion of any discourse (e.g., Biber et al., 1999). For example, Pawley and Syder (1983) reported that adult native English speakers retain on the order of at least several hundred thousand FSs. Erman and Warren (2000) have estimated that FSs account for 58.6% and 52.3% of spoken and written discourses, respectively. Hill (2000) has also argued that FSs permeate everything we say, hear, or write, and constitute up to 70% of our language use. Consequently, we conclude that mature native English speakers retain vast amounts of FSs and are versatile in using them in various contexts, and that non-native speakers who have a high

proficiency in English also possess a considerable knowledge of FSs (Conklin & Schmitt, 2008). In practice, most Japanese EFL learners tend to speak and write English based on their knowledge of grammatical rules and vocabulary (Yagi & Inoue, 2013). Consequently, their output often sounds unidiomatic or unnatural to native ears even when every sentence is grammatically correct (e.g., Pawley & Syder, 1983), which may hinder smooth communication. Based on these findings, an ability to retain various FS knowledge and use them appropriately can help to improve a person's communication skills.

Finally, and most importantly, a holistic representation of FSs can reduce the cognitive load in language processing, particularly for L2 learners, and is beneficial to a capacity-limited working memory (cf. Atkinson & Shiffrin, 1968). Since FSs are meant to indicate unit status semantically and phonologically, they require less effort in processing than other word sequences (i.e., NFSs). Regarding FSs' unit status, Cowan (2001) proposed that we can remember approximately four-word chunks in our capacity-limited working memory. Hence, by exploiting a holistic representation of FSs, we can increase the amount of linguistic symbols that can be processed at one time with a lower cognitive load (e.g., Levelt, 1989). A reduction in language processing load also contributes fluency and accuracy in language use. Furthermore, a certain degree of speed and automatic language use is considered important for smooth communication (Kadota, 2014). Consequently, the four benefits of FSs are interrelated and a retention of FS knowledge can improve the competency of Japanese EFL learners to communicate in English.

## 2.2 Research Question

In order to investigate the status of processing contextually appropriate FSs in the mental lexicon of Japanese EFL learners, the research question (RQ) of this study is formulated as follows:

RQ: Regarding accuracy and speed, how

differently is FS knowledge processed from NFS knowledge in contexts by Japanese EFL learners with different proficiency and educational backgrounds in English?

## 3. Methods

### 3.1 Participants

The participants included 114 adult Japanese EFL learners—who were mostly undergraduate and graduate students—and 50 second-year high school students. Participants whose mean proportion correct and reaction times (RTs) in the Phrase Appropriateness Judgment Task (PAJT) were  $\pm 3$  standard deviations (*SD*) from the mean value of the population and those who displayed an inappropriate attitude were excluded from the analysis, leaving a total of 112 adult EFL learners and 49 high school students finally included in the analysis. Prior to the experiment, all participants provided their written consent to be involved in the study. After the experiment, participants completed simple questionnaires, and the adults and students were awarded a small amount of compensation and candies, respectively, for their participation.

#### 3.1.1 Adults

The university majors of the 112 adult participants (67 females and 45 males; mean age = 22.73 years) varied and included English literature. Eighty-four participants submitted self-reported Test of English for International Communication (TOEIC) scores—which ranged from 170 to 950 (mean [*M*] = 588.67)—and 57 participants submitted Eiken test grades that ranged from grades 4 to 1 (38 participants submitted both TOEIC scores and Eiken test grades). In order to obtain a more accurate assessment of their English proficiency level, the participants were divided into three groups according to their answer proportion correct on the PAJT. Consequently, those in the middle group (*n* = 37) were excluded from the analysis, while 37 and 38 participants were allocated to the higher proficiency (upper group) and lower proficiency (lower group) groups, respectively.

Using the same procedures described above, after further stratification, only a total of 36 participants remained in each group. Using the guidelines provided by the Common European Framework of Reference for Languages (CEFR), the English proficiency level in the upper group (23 females and 13 males; mean age = 25.83 years) was estimated at B1 based on the self-reported TOEIC scores of 29 participants ( $M = 735.86$ ). In the lower group (20 females and 16 males; mean age = 20.26 years), the CEFR level was estimated at A2, based on the self-reported TOEIC scores of 21 participants ( $M = 450.48$ ).

### 3.1.2 High school students

The students (23 females and 26 males) were from a private high school who were attending the third term in the second grade in the academic year of 2017. They had completed their English module taught by the school and had probably not begun their preparations for university entrance examinations. After English activity classes were implemented for pupils who were attending the fifth grade and sixth grades of elementary school in 2011, they became part of the pioneer student cohort who attended these classes in their fifth grade, a fact that was verified from their response to the questionnaire which inquired about their background in English learning. None of them was a member of the English-speaking club activities in high school. Their CEFR level was estimated at A1 to A2 based on the self-reported Global Test of English Communication (GTEC) for Students (Benesse Corporation) scores of 47 participants ( $M = 484.75$ ).

### 3.2 Tasks

For this experiment, participants were asked to complete the PAJT, which was designed to measure and compare how accurately and quickly they assessed the semantic appropriateness of FSs and NFSs to the context. The task was undertaken individually.

### 3.3 Materials

The PAJT stimuli comprised two types of sentences: those with appropriate and inappropriate contexts for target FSs and NFSs (see Appendix). A total of 30 and 20 target FSs and NFSs, respectively, were chosen from the stimulus list of Isobe (2014), and 10 NFSs were modified for this experiment. In order to measure the processing accuracy and speed of participants, continuous sequences that did not require any grammatical processing were used as stimuli. Since processing speed can be influenced by many factors—such as word familiarity, number of words, and word length (Gernsbacher, 1984)—stimulus sequences and contexts were cautiously controlled to detect the processing of target word sequences based on two criteria.

First, target FSs must be used with a certain degree of frequency, while target NFSs must employ word sequences that are used infrequently in practical settings. Moreover, the target FSs would involve sequences that the participants had likely encountered when they were learning English in junior high and high schools. Control NFSs comprised 30-word sequences, and a word that was considered central to the paired FS was replaced. For example, a NFS that was paired with the FS for *take part in* was changed to *take work in*. The British National Corpus (BNC) was used to examine how frequently these sequences occurred in the general usage of English. It was found that the numbers of occurrences of target FSs and NFSs in the BNC were over 300 and less than 90, respectively. An unpaired *t*-test was conducted to confirm that the overall number of occurrences of FSs in the BNC was significantly different from that of NFSs,  $t(29.00) = 7.90$ ,  $p < .001$ ,  $r = .83$ . As the result of the Levene's test indicated unequal variances ( $F = 42.90$ ,  $p < .001$ ), the degrees of freedom were adjusted from 58 to 29. The familiarity of the replaced words in the NFSs was controlled so that they did not differ significantly from the words in the FSs. Data from the "Word Familiarity List for Japanese English Learners" (Yokokawa, 2006) were

used to estimate the familiarity of every word. In terms of word familiarity, the result of an unpaired *t*-test did not show a significant difference between replaced words in the NFSs and the original central word in the FSs,  $t(58) = -1.15$ ,  $p = .26$ ,  $r = .15$ . The numbers of syllables of paired FSs and NFSs were equally matched. Finally, the forms of the verbs in each paired sequence—such as tense, aspect, and voice—were thoroughly identical.

Second, since FSs are used in contexts, contexts in which the stimulus sequences were to be embedded were created such that the components of the stimulus sentences—other than FSs and NFSs—were carefully controlled based on the following criteria: word familiarity, number of words, readability, and grade level. In PAJT, the stimuli were sentences that were embedded with parentheses where the target sequences would be placed. The Flesch Reading Ease and Flesch-Kincaid Grade Level (bundled with Microsoft Word) were used to estimate readability, and mean word familiarity was calculated using data from the “Word Familiarity List for Japanese English Learners” (Yokokawa, 2006). An unpaired *t*-test was performed to confirm that there were no significant differences in the number of words, word familiarity, readability, or grade level between the stimulus sentences of FSs and NFSs.

The results of an unpaired *t*-test demonstrated that the two types of stimulus sentences did not significantly differ (see Table 1; in tables, *df* denotes degree of freedom in this study). As another controlling factor, and in order not to interfere with the true reactions of participants in the PAJT (i.e., their semantic judgment concerning the target sequences in context), the level of English that was used in the stimulus sentences—excluding the target FSs and NFSs—was designed to be easily understood. For example, most of the words that made up the stimulus sentences were chosen from words whose degree of word familiarity (Yokokawa, 2006) was 6 or higher (maximum [*Max*]: 7; see Table 1). Furthermore, since the stimuli consisted of sentences with parentheses, the participants were obliged to understand and memorize two parts of the sentence: the parts before and after the parentheses. Therefore, it was appropriate to use between one and four words for each part. The total number of words in the stimulus sentences—including FSs and NFSs—ranged from 6 to 9 ( $M = 7.3$ ,  $SD = 0.83$ ). Additionally, the influence of preceding stimulus sentences on the processing of following stimuli—the priming effect—was considered. Consequently, none of the stimulus sentences that were used in this task were contextually interrelated.

Table 1  
*Mean Values of Factors of Two Types of Stimuli and Results of the t-Test*

	FSs ( <i>n</i> = 30)		NFSs ( <i>n</i> = 30)		Results of <i>t</i> -Test ( <i>df</i> = 58)		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	<i>r</i>
<b>Word Familiarity</b>							
words before sequences	6.43	0.20	6.48	0.22	-0.96	.343	.13
words after sequences	6.36	0.23	6.42	0.18	-1.10	.280	.14
<b>Number of Words</b>	7.40	0.93	7.27	0.79	0.60	.551	.08
<b>Readability</b>	73.33	14.81	74.55	16.03	-0.31	.760	.04
<b>Grade Level</b>	4.77	2.11	4.56	2.27	0.37	.716	.05

### 3.4 Procedures

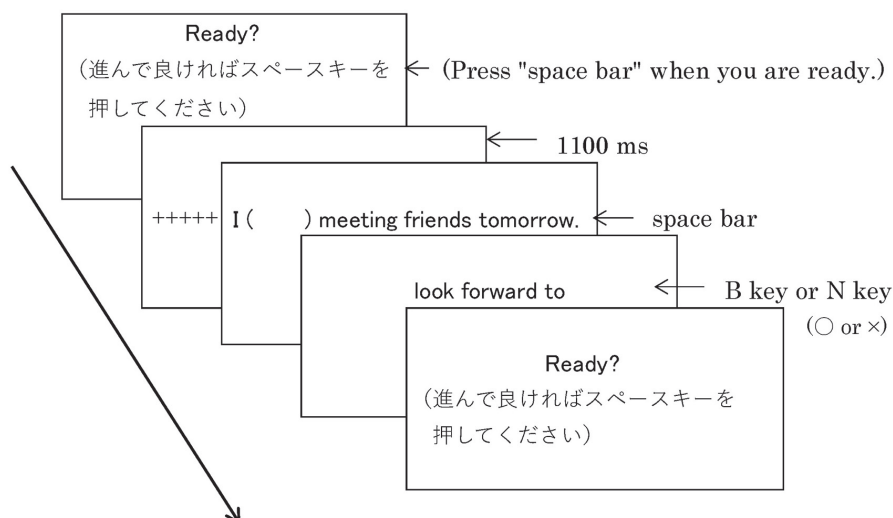
The experiment was administered individually in a quiet room. The time allotted to complete PAJT was approximately 15 minutes. A laptop computer was provided to participants during the experiment, and they were not allowed to return to the previously displayed screens. The SuperLab Stimulus Presentation Software (SuperLab Pro, Version 5) was installed on the laptop computer (SONY VAIO; SVE15114FJP, the width and height of screen size were 34.6cm and 19.4cm, respectively) to measure the participants' RTs to every stimulus and their answers during PAJT. All instructions on PAJT were provided in Japanese on the computer screen, and additional explanations were provided by the experimenter when necessary. Five exercises were prepared prior to the commencement of 60 trials. After the exercises were completed, participants pressed the space bar to launch the trial as soon as they were ready to do so. The order of the stimuli was randomized every time participants engaged in the task. In every trial, a fixation marker "+++++" appeared at the center of the computer screen for 1,100 milliseconds (ms) before a stimulus sentence with parentheses—where a word string would be grammatically suitable—was presented until participants pressed the space bar. The letters were

displayed in lowercase MS P Gothic typeface in a 30-point font size. Participants were instructed to read and comprehend the stimulus sentences and press the space bar as quickly and precisely as possible. After the keypress, a stimulus word string appeared on the screen and participants decided whether it was semantically appropriate to fill the parentheses in the previous sentence by depressing the "B" key on the keyboard for "YES" or the "N" key for "NO" as quickly and precisely as possible. SuperLab (SuperLab Pro, Version 5) aided the measurement of two factors: whether the answers entered by the participants were correct and the time in milliseconds taken by the participants to enter their response (i.e., RTs) after a stimulus was presented. After depressing the "B" or "N" key, participants were prompted on whether they were ready to proceed to the next trial, which they did after they had pressed the space bar.

After participants had completed 12 trials, they were greeted by an instruction on the screen informing them that they were allowed to take a short rest. Participants engaged in the task at their own pace, and it lasted for approximately 15 minutes. During the task, no instructions on word sequences were given to the participants. The task flow of the PAJT is shown in Figure 1.

Figure 1

Flow of the PAJT (Phrase Appropriateness Judgment Task)





3.5 Data Analysis

The data collected by SuperLab (SuperLab Pro, Version 5) were then analyzed. Before the analysis, the data were scanned for outliers that could potentially skew the results. The dependent variables in PAJT were proportion correct and mean RTs; RT reflected the process of automaticity in the participants because it provides more sensitive data on their mental representations than accuracy (Jiang, 2012). The proportion correct represented the accuracy of a participant’s judgment in the PAJT (i.e., accurate judgment of the semantic appropriateness of word sequences to given contexts). Before the proportion correct was calculated, keypress errors and outliers for every participant’s RTs were discarded. The mean RT was a variable that indicated how quickly participants judged the semantic appropriacy of target word sequences to given

contexts between the presentation of the stimulus sequences and keypress to enter their response. The numbers of correct and incorrect responses were subjected to proportion-correct analysis, while only correct responses were subjected to RT analysis. A paired *t*-test was performed to examine differences in the contextual FS and NFS knowledge of Japanese EFL learners.

4. Results

4.1 Adults

4.1.1 Upper group

Table 2 presents the results of the mean proportion correct and RTs of the upper group in the PAJT. Figures 2 and 3 are graphical representations of Table 2. Error bars denote a 95% confidence interval in this paper.

The results of a paired *t*-test demonstrated

Table 2  
Mean Proportion Correct and RTs in the Judgment of Contextually Appropriate FSs and NFSs by the Upper Adult Group in the PAJT

	Proportion Correct				RTs (ms)			
	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Overall	.77	.04	.72	.86	2890	760	1400	4850
FSs	.79	.06	.69	.93	2740	730	1360	4670
NFSs	.74	.07	.55	.90	3050	850	1460	5400

Note. *n* = 36.

Figure 2

Mean Proportion Correct of the Stimulus Sentences for FSs and NFSs in the PAJT for the Upper Adult Group

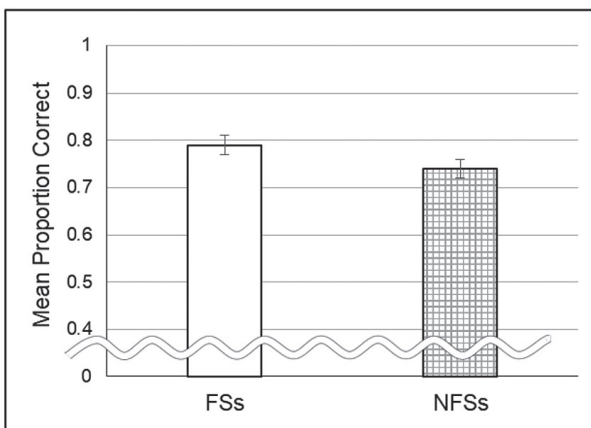
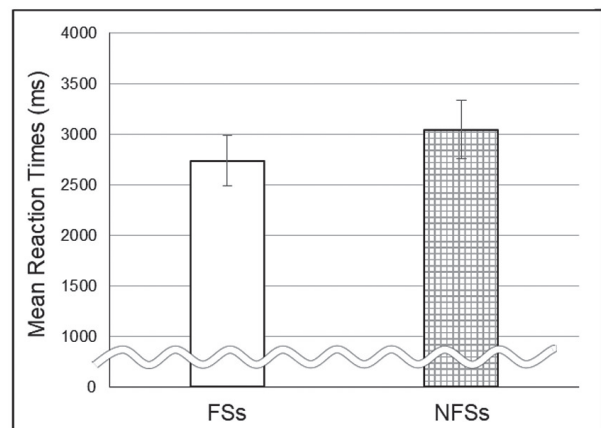


Figure 3

Mean Reaction Times of the Stimulus Sentences for FSs and NFSs in the PAJT for the Upper Adult Group



a significant difference between the stimulus sentences for FSs and those for NFSs on accuracy (i.e., proportion correct) and speed (i.e., RTs),  $t(35) = 2.93, p = .006, r = .44$ , and  $t(35) = -3.89, p < .001, r = .55$ , respectively. This finding implies that the participants in the upper group processed contextual FSs more accurately and more rapidly than they did NFSs.

4.1.2 Lower group

The results for the mean proportion correct and RTs in the lower group are shown in Table 3. Figures 4 and 5 are graphical representations of Table 3.

A paired  $t$ -test was performed to analyze judgment accuracy and speed. The  $t$ -value for

Table 3  
*Mean Proportion Correct and RTs in the Judgment of Contextually Appropriate FSs and NFSs by the Lower Adult Group in the PAJT*

	Proportion Correct				RTs (ms)			
	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Overall	.54	.05	.44	.62	2830	930	720	5240
FSs	.55	.11	.34	.82	2690	800	680	4540
NFSs	.54	.08	.41	.70	2950	1090	770	5790

*Note.*  $n = 36$ .

Figure 4  
*Mean Proportion Correct of the Stimulus Sentences for FSs and NFSs in the PAJT for the Lower Adult Group*

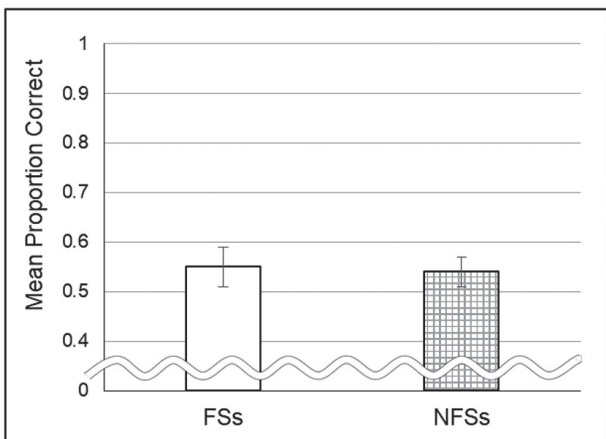
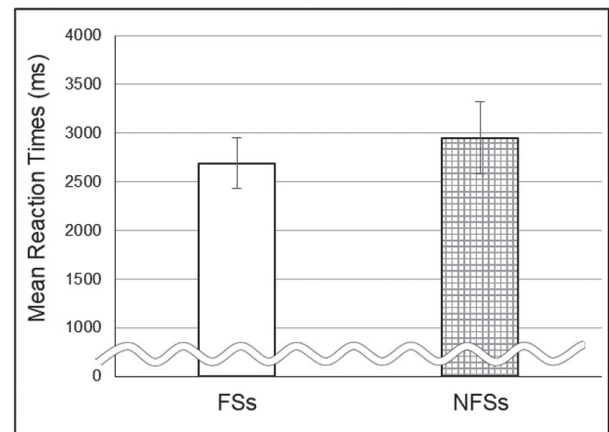


Figure 5  
*Mean Reaction Times of the Stimulus Sentences for FSs and NFSs in the PAJT for the Lower Adult Group*





accuracy did not indicate a significant difference between the judgments of FSs and NFSs:  $t(35) = 0.28, p = .783, r = .05$ . However, the  $t$ -value for speed revealed a significant difference between judgments for both types of stimuli:  $t(35) = -3.12, p = .004, r = .47$ , indicating that the participants in the lower group judged the appropriacy of FSs in their contexts more rapidly than NFSs when their

judgments were correct.

#### 4.3 High School Students

Table 4 shows the mean proportion correct and RTs in the PAJT. Figures 6 and 7 are graphic representations of Table 4.

The results of a paired  $t$ -test showed that the judgment accuracy for FSs in context was not

Table 4  
*Mean Proportion Correct and RTs in the Judgment of Contextually Appropriate FSs and NFSs by High School Students in the PAJT*

	Proportion Correct				RTs (ms)			
	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Overall	.60	.07	.41	.74	2850	830	1010	5050
FSs	.61	.09	.38	.76	2740	830	950	4830
NFSs	.60	.09	.39	.80	2970	900	1080	5590

Note.  $N = 49$ .

Figure 6  
*Mean Proportion Correct of the Stimulus Sentences for FSs and NFSs in the PAJT for High School Student Group*

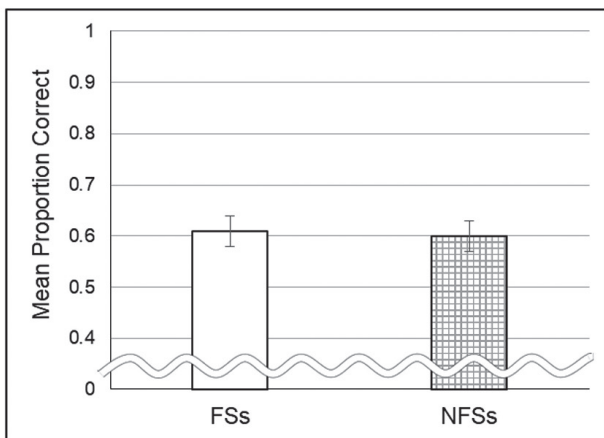
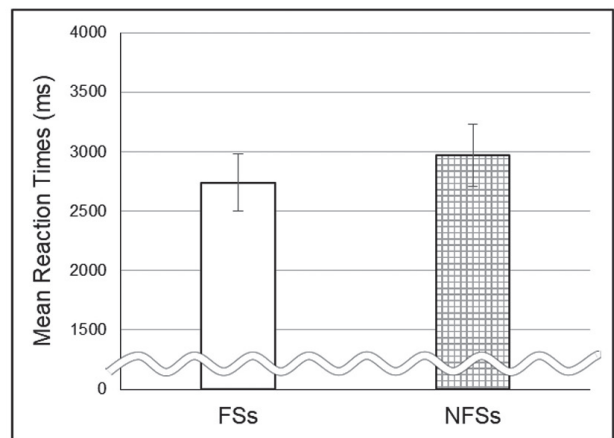


Figure 7  
*Mean Reaction Times of the Stimulus Sentences for FSs and NFSs in the PAJT for High School Student Group*



significantly higher than that for NFSs, while the judgment speed for FSs in context was significantly faster than that for NFSs when the high school participants correctly judged the appropriacy of word sequences,  $t(48) = 0.85, p = .401, r = .12$ , and  $t(48) = -3.26, p = .002, r = .43$ , respectively.

### 5. Discussion and Conclusion

This study yielded the following answers to the RQ investigated in this study:

- A) Japanese EFL learners who have completed at least the second grade in high school may process contextual FSs significantly faster than NFSs when they can properly comprehend the word sequences and contexts.
- B) Japanese EFL learners who have a high proficiency in English (with an estimated CEFR level of B1) can process contextually appropriate (inappropriate) FSs more accurately and rapidly than NFSs.
- C) There is no difference in accuracy in judging the contextual appropriateness between FSs and NFSs by adults with a low proficiency in English and high school students who are Japanese EFL learners (with an estimated CEFR levels from A1 to A2).

In PAJT, participants were instructed to judge the semantic appropriateness of the target word sequences in the given contexts as accurately and quickly as possible. Thus, the RT data reflected the semantic processing speed of the sequences in contexts where they would or would not be suitably embedded. The results revealed that when their judgments were correct, all participants judged contextually appropriate FSs significantly faster than they did NFSs. This finding suggests that FSs may be processed differently from NFSs in the mental lexicon of Japanese EFL learners, which specifically implies the hypothesized holistic representation of FSs. Although the continuous processes of the mind have made it impossible to distinguish the processing boundary between FSs and their contexts, there is however a possibility that holistic FS processing can lessen cognitive load

and increase the speed of sentence-length semantic processing, which is an essential factor for smooth communication.

However, in terms of judgment accuracy, a significant difference was only seen in the upper adult group. The appearance of a significant difference in RTs in the results, regardless of the level of English proficiency, can be attributed to the inherent tendency of RTs to provide more sensitive data on mental processing than can accuracy (Jiang, 2012).

Meanwhile, only participant in the upper adult group could judge contextually proper FSs more accurately than NFSs. Based on their relatively high proportion correct in the PAJT, they appeared to have a stable grasp of English that can be flexibly used in various contexts (cf. Conklin & Schmitt, 2008). Conversely, the judgment accuracy for appropriate FSs and NFSs either did not differ significantly or was not high in the other two groups. Consequently, there is a possibility that FSs were not clearly perceived and learned as FSs by participants in the lower adult group or by the high school students. Surprisingly, but convincingly, according to their comments in the submitted questionnaires, most participants in the lower adult group and the high school group indicated that they thought that all stimulus word sequences may have been FSs. Another common comment of participants was that they had recalled having seen some stimulus word sequences in their English learning journey, but could not remember their meanings. As mentioned in the literature review, FSs are often non-literal, and phrasal verbs in particular are not easily discernible because they comprise words that are common and familiar words. This explanation may account for the participants' detailed comments on the instability of FS knowledge and difficulty in FS learning. Moreover, proportion correct of the high school participants indicated serious problems: Although they had attended English classes since their elementary school days and continued to attend weekly English classes, they may still experience a

lack of exposure to English in FS learning.

In conclusion, a favorable finding of this study is that the RT results suggest that the participants unconsciously processed FSs holistically and enjoyed this benefit at a subliminal level (since RTs had demonstrated a significant difference between the two types of stimuli). On the other hand, this study reveals that a major challenge faced by most Japanese EFL learners in FS learning is that unless they have a certain level of English proficiency (i.e., a CEFR level of B1), their contextual FS knowledge tends to be unstable and they are unable to make semantically accurate FS judgments. Based on the characteristics of FSs, there are three educational implications for their teaching and learning. First, teachers should provide opportunities for learners to encounter FSs repeatedly in various contexts, such as through extensive reading and listening. Second, they should teach learners that combinations of familiar words may sometimes have very different meanings, and encourage them to look up any word combinations that they encounter repeatedly. Third, sometimes, FSs should be explicitly taught in class.

## 6. Limitations and Further Study

This study has three limitations. First, the outcomes were confined to the stimulus word sequences that were used in the experiment. Moreover, some stimulus FSs were relatively easy because a criterion for the selection of target FSs stipulated that they should involve word sequences that the participants were expected to have learned in junior high school and high school. Second, although the various influencing factors of the stimuli were controlled as much as possible, there were other factors which might have influenced the mental processes of the participants. Furthermore, to meet the rigorous and elaborate control conditions imposed by the experimental design, the stimuli might include somewhat unnatural expressions. Third, the expressions “upper” and “lower” only refer to the participants in this study, and not to the general population of Japanese EFL

learners.

In order to address these limitations, and for English language education, further research should conduct experiments that use other stimulus FSs and focus on FS learning, as through extensive reading. It is hoped that the results of such studies would provide useful suggestions for the teaching and learning of FS by Japanese EFL learners.

## Acknowledgements

I would like to express my sincere gratitude to all participants for their generous cooperation on this study.

## References

- Abboud, H., Schultz, H., & Zeitlin, V. (2014). SuperLab Pro (Version 5) [Computer software]. San Pedro, CA: Cedrus Corporation.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory: Vol.2* (pp. 89–195). Academic Press.  
[https://doi.org/10.1016/S0079-7421\(08\)60422-3](https://doi.org/10.1016/S0079-7421(08)60422-3)
- Biber, D., Jonansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- 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>
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114.  
<https://doi.org/10.1017/S0140525X01003922>
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, *20*, 29–62.
- Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281.  
<https://doi.org/10.1037/0096-3445.113.2.256>
- Hill, J. (2000). Revising priorities. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 47–69). Language Teaching Publications.
- Isobe, Y. (2014). *Psycholinguistic reality of formulaic sequences: Evidence from Japanese EFL learners* (Unpublished doctoral dissertation). Graduate School of Language, Communication, and Culture, Kwansei Gakuin University.
- Jiang, N. (2012). Conducting reaction time research in second language studies (pp. 1–13). Routledge.

- Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3), 433–445.  
<https://doi.org/10.1111/j.1540-4781.2007.00589.x>
- Kadota, S. (2014). *Eigo jotatsu 12 no pointo: Kagakuteki riron ni motozuku gaikokugoshutoku seiko no hiketsu* [Twelve tips for successful learners of English: The tips for success in second language acquisition based on scientific theory]. CosmoPier.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.
- Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2013). *Gurobaruka ni taio shita eigokyoikukaikaku jissihikeikaku* [The practical plan for English education reform to accommodate Globalization in Japan].  
[http://www.mext.go.jp/b\\_menu/houdou/25/12/\\_\\_\\_icsFiles/afiefieldfile/2013/12/17/1342458\\_01\\_1.pdf](http://www.mext.go.jp/b_menu/houdou/25/12/___icsFiles/afiefieldfile/2013/12/17/1342458_01_1.pdf)
- Ministry of Education, Culture, Sports, Science and Technology (MEXT). (2018). *Shogakko gakushushidoryo kaisetsu: Gaikokugokatsudo, gaikokugo hen* [The explanation of the government course guidelines for elementary schools: Foreign language activities and foreign language classes]. Kairyudo.
- Moon, R. (1997). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40–63). Cambridge University Press.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford University Press.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmitt (Eds.), *Language and communication* (pp. 191–226). Longman.
- Saito, T., Nakamura, J., & Akano, I. (Eds.). (2005). *Eigo kopasugengogaku: Kiso to jissen* [English corpus linguistics: Fundamental and practice]. Kenkyusha.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1–22). John Benjamins.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. Routledge.
- Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 153–172). John Benjamins.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.
- Yagi, K., & Inoue, A. (2013). *Eigo teikei hyogen kenkyu: Rekishi, hoho, jissen* [English phraseology: History, method, practice]. Kaitakusha.
- Yokokawa, H. (Eds.). (2006). *Nihonjin eigo gakushusha no eitango shinmitsudo: Moji-hen* [English word familiarity for Japanese English learners: Written words]. Kuroshio.

#### Appendix: Examples of Stimulus Sentences for the Phrase Appropriateness Judgment Task

	Stimulus Sentences with Parentheses	Stimulus Sequences
1	I ( ) meeting friends tomorrow.	look forward to
2	My family members will ( ) spring cleaning	take part in
3	I will ( ) a credit card.	apply for
4	The driver ( ) the traffic light.	looked after
5	She can ( ) the manager.	carry out
6	Glasses ( ) our table after breakfast.	came up with
7	The people ( ) the sky.	look outside to
8	She can ( ) that system.	take work in
9	Members learn how to ( ) the rules.	behave from
10	A famous actor ( ) this town.	lived after
11	They can ( ) the mystery magazines.	happen out
12	Net shopping will ( ) high popularity.	put for

Note. Stimulus sentences for formulaic sequences: No. 1~6 (Yes-Response: 1 ~3; No-Response : 4~6).

Stimulus sentences for nonformulaic sequences: No. 7~12 (Yes-Response: 7~9; No-Response : 10~12).