Non-native and Native Acceptability of Variable Slots in Fixed Sequences:

A Corpus Driven Approach

Amanda Black

Northern Arizona University

Abstract

The present study examined how variable slots within a fixed formulaic construction are accepted by second language learners and native speakers of English via an acceptability survey. Of particular interest was whether frequency of occurrence of a variable construction played a determining role in acceptance. In addition, proficiency was considered. A corpus was used to determine frequency of the selected variable constructions and a comparison of the acceptability judgments was made. Results indicated that overall, acceptability of variability within formulaic sequences closely matched the frequency of use as demonstrated in a corpus independent of status as a native or second language speaker of English. However, proficiency seemed to impact results as acceptability was more closely related to the frequency of use in a corpus for participants of higher proficiency levels. These results support hypotheses made by usage-based theorists by indicating frequency is the driving force behind acceptance of a variable construction.

Keywords: acceptability, formulaic sequences, frequency, second language acquisition, usage-based theory, variability

Non-native and Native Acceptability of Variable Slots in Fixed Sequences:

A Corpus Driven Approach

Background

The interests of the present study are motivated by several pieces of relevant research. First, relevant findings have been provided by corpus linguistic research which has demonstrated that natural language does indeed make considerable use of recurrent multiword patterns that are "distributionally regular" and highly predictable in terms of their lexical concordance (Sinclair, 1991; Hopper, 1998). Corpora have also been used to determine common lexical bundles in different English registers and have been used to classify and categorize a list of common formulaic sequences comparable to the Academic Word List (AWL) (Biber & Conrad, 1999). Due to the fact that corpora have provided substantial evidence that FS occur far more commonly than happenstance would allow, the present study also made use of corpus investigation.

Specifically, two corpus-driven studies were conducted by Vivian Cortes (2002) and David Oakey (2002). Cortes (2002) examined the use of lexical bundles in freshmen academic writing and identified the most common four word sequences used. In addition, she identified shared common syntactic frames within the lexical bundles, many of which were fixed with occasional variable slots. For example, she identified the syntactic frames: Passive verb + PP fragment (e.g., *is based on),* and Preposition + noun phrase fragment (e.g., *in an effort to).* Cortes was also able to classify the bundles into several functional categories an example of which is inference or deduction as demonstrated by the bundle *it is obvious that.* The findings by Cortes in conjunction with the operationalization of FS and fixedness have motivated the selection of several of the FS to be tested in the present study including: *is *on the* and *in a(n) * to.*

Oakey (2002) looked at the use of formulaic phrases in English academic writing. He created a corpus to determine the formal and functional variations of a lexical phrases across academic disciplines. Oakey conducted his study by looking at the phrase *it is * that* and determining how the wildcard space changed in both form and function within and between disciplines. He concluded that, in addition to "topic priming", the phrase was used for attributed and non-attributed support, negatively evaluated statements, and reference within the text demonstrating that function varied more than form. While these results are interesting, the present study aims to look at the acceptance of variable slots within FS. To do so, the researcher will search for frequency counts of the above phrases as well as Oakey's *it is * that* to determine whether acceptability judgements made by learners have any connection to frequency of use in an academic corpus. The motivation behind the selection of academic texts comes from access to participants as well as from Ellis, Simpson-vlach, & Maynard (2008) whose analyses of academic corpora demonstrated that academic discourse contains a high frequency of formulaic sequences that have important functions in academic writing.

In addition to these two studies, Staples, Egbert, Biber, & McClair (2013) stated, "formulaic sequences are widely used in academic writing and are known to be an important aspect of EAP writing development. However, little research has investigated the frequency, function and degree of fixedness of their use by ESL writers across proficiency levels" (p. 214). This quote demonstrates a gap in the research about FS in terms of fixedness and their use across proficiency levels. While the present study was not looking at the productive use of FS, it attempted to fill the aforementioned gap by examining acceptability of FS with variable slots, comparing this to frequency in a corpus, and checking for differences in acceptability across proficiency levels.

To examine acceptability, the present study relied on an acceptability survey. This was motivated by a study conducted by Ellis *et al.* (2008) that examined the representation and processing of formulaic sequences in two grammaticality judgments, one with formulaic sequences and the other with comparably structured non-formulaic sequences. Participants were asked to judge whether visually presented word strings were likely to be found in English. The present study will follow a similar pattern except all sentences will ask participants to judge the variable slots in the following formulaic sequences: *it is * that, in a(n) * to, and is * on*. The wildcard (or variable) slots was filled with items based on frequency of occurrence as the study conducted by Ellis *et al.* (2008) who determined that frequency had the biggest impact on processing FS for language learners. In addition, Ellis *et al*. (2008) determined that there was a difference in processing of FS by native and non-native speakers. To determine if there is a difference in acceptability of variable slots within a FS based on frequency, the present study compared results of both language learners and native speakers of English.

The last major piece of motivation for the current study comes from studies conducted about frequency. Many theories of second language acquisition consider frequency as a major factor of successful language learning. For example, Usage-Based theories of language suggest that frequency and repetition affect form in language, and psycholinguistic research demonstrates that language users are sensitive to the frequency of use of a wide variety of constructions (Ellis *et al*., 2008). Furthermore, research in this area has produced evidence that language processing is sensitive to formulaicity and collocation (Wood, 2006). Given this, many studies of second language acquisition have looked at the effect of frequency on formulaic sequences. Research by Doug Biber and Susan Conrad (1999) discussed and categorized the use of lexical phrases based on frequency in a corpus to determine patterns of actual use in conversation in academic prose.

Later research by Ellis *et al.* (2008) demonstrated that non-native speakers of English were sensitive to the frequency of formulaic sequences and determined that it is predominantly the frequency of the formula that affects processing. Ellis further claims that both native and non-native speakers become attuned to more frequent word sequences, and suggests that pedagogy take this into consideration. Wood (2006) also suggests that frequency plays an important role in the noticing, processing, and production of formulaic sequences; he states, "Consciousness, awareness, and noticing of formulas in input (Schmidt, 1990) might establish a pattern-recognition unit, which is then strengthened by frequent input" (p. 17).

The presents study's consideration of frequency comes from a theory of second-language acquisition, Usage-Based Theory. This theory claims that frequency is a driving force behind acquisition and that frequent exposure to input, particularly prototypical exemplars, will help learners induce rules of a language (Ellis & Wulff, 2015). This suggests that leaners should be more accepting of FS that occur at a higher frequency as the most frequent variable slot can be argued to serve as a prototype. Thus the present study aims to test whether the variable slots of FS will be accepted at a higher rate based on frequency of occurrence and how that acceptability compares to that of native speakers because if it can be shown that learners are sensitive to frequency differences in a specific type of FS, then this would lend serious support to the claims made by Usage-Based Theory.

Research Questions

Despite the amount of research that has been done on formulaic sequences and language learning, some facets of formulaic sequences remain largely unexplored. Specifically, preference for formulaic sequences with variable slots and whether the acceptability of leaners matches the preferred use of variable sequences by native speakers. Therefore, as has already been stated, the

aim of the present study is to explore the acceptance of formulaic sequences of English language learners via an acceptability judgement to determine what variable slots learner finds acceptable (Mackey & Gass, 2005). If learners are attuned to the frequency of formulaic sequences as suggested by Ellis, then they should, theoretically, demonstrate a preference for the variable form with the highest frequency of occurrence. To test this assumption, the results from an acceptability judgement will be analyzed and compared to frequency counts in a corpus and to results of the same acceptability judgement given to native speakers. In order to conduct this study the following research questions will be used:

1) Do language learners demonstrate the same preference for variable slots in FS as native speakers?

2) Is there a difference in response to acceptability judgements based on proficiency level?

Methods

Participants

To answer the research questions above, participants were selected from the English department at Northern Arizona University. Specifically, participants were selected from the English 105 composition course and the Program Intensive English department. A total of 64 participants were surveyed including 10 participants of intermediate proficiency comparable to the Common European Framework of Reference level (CEFR) B1; 24 participants were classified as upper intermediate and advanced, respectively, with levels comparable to CEFR levels B2 and C1; 8 students were second language learners enrolled in English MA programs at NAU with highly advanced proficiency levels; 17 students were native speaking freshmen from English 105, a course that is the equivalent of the courses taught to upper intermediate and

advanced courses of the PIE; and 5 participants were native speaking students enrolled in the same English MA programs at NAU.

The demographics of the surveyed participants were as follows: 31 male, 32 female and 1 declined to state. In addition the L1 breakdown of surveyed participants were 31 Chinese, 22 English, 3 Arabic, 2 Persian, 2 Russian, 2 Spanish, 1 Korean, and 1 Turkish.

Instruments

The corpus. To determine the frequency of the variable word in the selected formulaic phrases *it is * that, in a(n) * to, and is * on*, a corpus was used. These phrases were selected based on previously mentioned literature and based on the fact that they fit the definition of FS as defined in the introduction. Each chosen sequence occurs frequently in a corpus, demonstrates a pattern, contains a variable slot, and has a functional application. The frequency of each FS and corresponding variable slots were determined using COCA, the Corpus of Contemporary American English, which has approximately 520 million words and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. Variables were selected based on frequency distribution so that the two highest, the two medial, and the two least frequent variables of each construction were used as tokens in the survey. This frequency distribution combined with semantic similarity of the variables, dubbed as semantic groups by Hunston (2001), were considered during the selection process to reduce the amount of possible constructions and to control for unanticipated variables related to context.

Given the criteria above, the frequency count of the variable forms were as follows: *It is * that* revealed: it is clear that-2511 times, it is true that -1961 times, it is obvious that-455 times, it is certain that-199 times, it is observed that-26, and it is plain that-26 times. Next, *Is * on the* revealed: is based on the-2139 times, is dependent on the-216 times, is predicated on the-103, is

founded on the-69 times, is premised on the-48 times, and is relying on the-19 times. Lastly, *In a(n) * to* revealed: in an effort to-4906, in a position to- 2340 times, in a place to- 24 times, in a situation to-20, in an area to-6 times, and in an environment to-1 time. Table 1 below shows the frequency of occurrence for each variable slot as demonstrated by COCA.

Table 1

Frequency of Occurrence of Variable Slots

	High	**High**	**Mid**	**Mid**	**Low**	**Low**
*It is * that*	Clear	True	Obvious	Certain	Observed	Plain
	2511	1961	455	199	26	26
*Is * on*	Based	Dependent	Predicated	Founded	Premised	Relying
	2139	216	103	69	48	19
*In a(n) **	Effort	Position	Place	Situation	Area	Environment
to	4906	2340	24	20	6	1

    Acceptability survey. Based on the frequency counts from the corpus, an acceptability judgement survey was created (see appendix). The survey asked participants to read 26 sentences, each with a different formulaic sequence, and mark the sentence as acceptable, somewhat acceptable, somewhat unacceptable or unacceptable. Based on results from a pilot study, each formulaic frame was provided in six different variable forms within the same context to avoid conflation of the results due to context. As previously stated, each sequence was chosen based on its frequency in the corpus; the two highest frequencies, the two middlemost frequencies, and the two lowest frequency counts were selected. Participants were also asked to rank all six sequences in order from the construction they would most likely use in academic

writing to the one they would least likely use. Table 2 below provides an example of the test

tokens. In addition to questions about acceptability, participants were asked their L1, gender and

major to determine whether alternative variables play a role during the analysis of results.

Table 2

Sample Test Tokens

Sample number	Test token
1	It is true that students need explicit spelling instruction. I find the underlined phrase _____ to use in writing. a. Acceptable b. Somewhat acceptable c. Somewhat unacceptable d. Unacceptable
2	It is clear that students need explicit spelling instruction. I find the underlined phrase _____ to use in writing. a. Acceptable b. Somewhat acceptable c. Somewhat unacceptable d. Unacceptable
3	Please rank the phrases below in the order that you would most likely use them in writing. 1 being the one you would ***most likely*** use and 6 being the one you would ***least likely*** use. ____ It is clear that ____ It is certain that ____ It is true that ____ It is plain that ____It is obvious that ____ It is observed that

Procedures

      Before giving participants the survey, a practice session in which the researcher discussed

directions, had participants practice answering, and allowed time for questions and clarification

was conducted. This helped preclude convulsion of the results due to confusion or

misunderstandings of the assignment directions.

      Based on results from a pilot study, participants were given 15 minutes to complete the

survey. Once participants completed the surveys, results were collected and analyzed statistically

to determine the acceptability of the formulaic sequences. If participants selected acceptable for a

construction a score of 4 was given, and if participants selected somewhat acceptable for a

construction a score of 3 was given; likewise, if participants selected somewhat unacceptable for

a construction a score of 2 was given, and if participants selected unacceptable for a construction

a score of 1 was given. All results were scored and descriptive statistics were retrieved using

SPSS. Results were divided and statistically analyzed based on proficiency level, major, and

gender to answer the research questions above; findings will be discussed below.

Results

Results from the survey indicated that acceptance of variability in a formulaic sequence

was comparable to the frequency count as demonstrated by the corpus. The first searched

construction revealed *it is clear that* at the highest frequency with over 2,500 occurrences in

COCA; results from the present study demonstrate a mean acceptance rate of *it is clear that* at

3.7 which is .04 higher than the nearest accepted variable form. Likewise, 53% of participants

ranked *it is clear that* as the most likely construction they would use in their own academic

writing. Figure 1 below visually represents acceptance of the variable constructions for *it is **

*that* based on rankings in the present study.

It is * that

Label	Value
Plain	0
Certain	3%
Obvious	11%
TRUE	13%
Observed	20%
Clear	53%
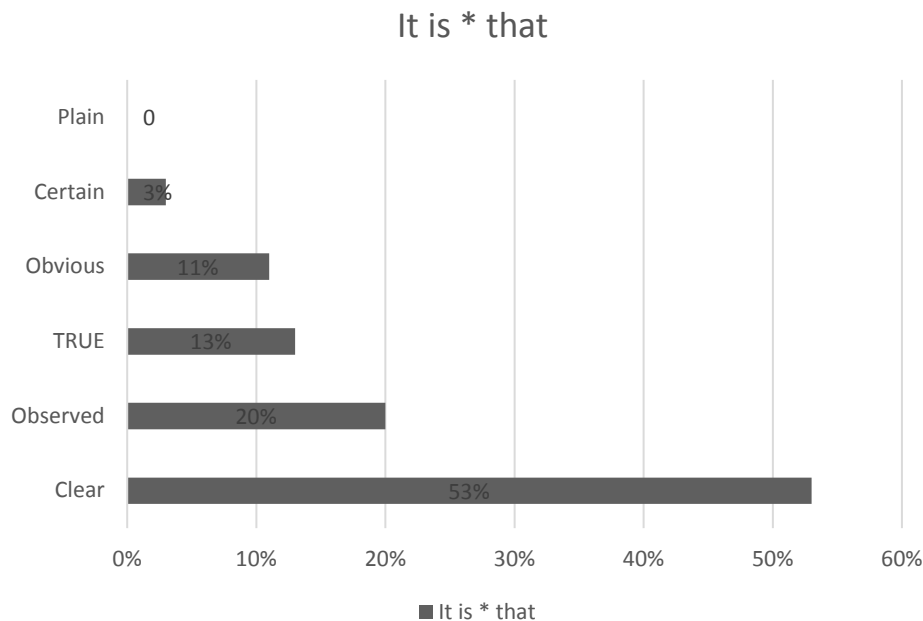
■ It is * that

Figure 1. Rate of acceptability for variability within the fixed structure *it is * that.*

The corpus also showed that the least frequent variation of this form was *it is plain that* with a frequency count of only 26. The results from the current study also show this variation as least accepted with a mean acceptance rate of 2.2. Additionally, 48% of participants ranked it as the least likely construction they would use in their own academic writing.

The next construction searched was *is * on.* The corpus showed the most frequent construction as *is based on* with 2,139 occurrences and the least frequent as *is relying on* with just 19 occurrences. As with the previous construction, the present study's findings also demonstrate a similar trend in acceptability with mean acceptability rates of 3.6 and 2.6, respectively. The majority of participants, 56%, also ranked *is based on* as the construction they were most likely to use in writing and *is relying on* as the least likely. Interestingly, the results of acceptability for this construction follow the trend of frequency distribution in COCA very closely as can be seen in Figures 2 & 3 below**.**

Is * On

Relying
Premised
Founded
Predicated
Dependent
Based

0%   10%   20%   30%   40%   50%   60%

■ The Prstent Study
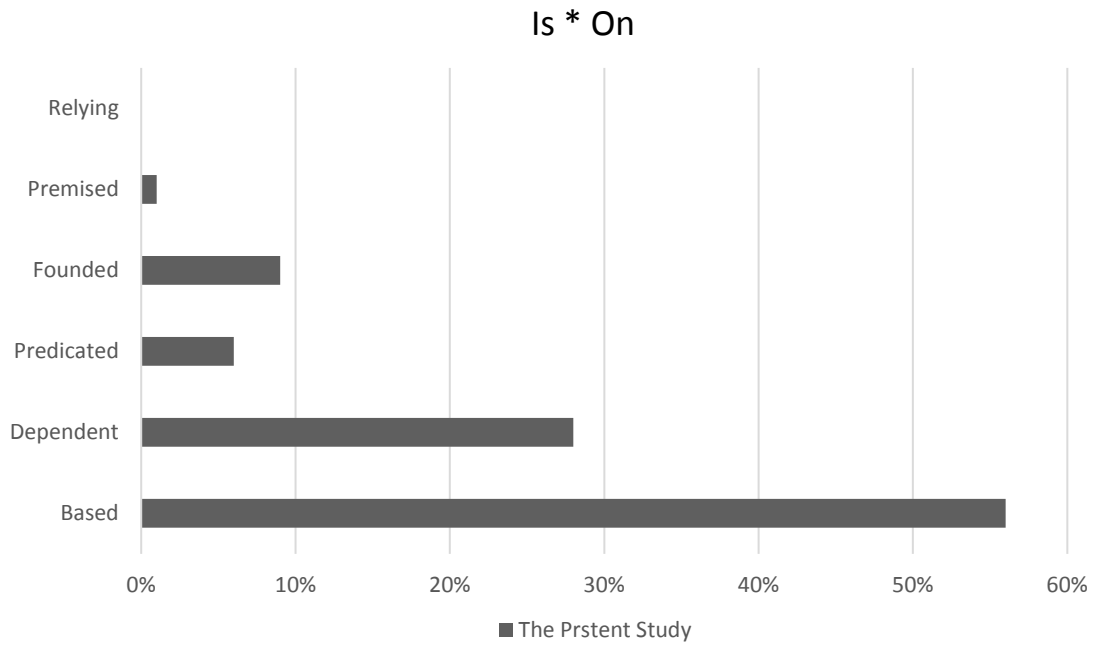
Figure 2. Rate of acceptability of *is* on* .

Is * On

Relying | 19
Premised | 48
Founded | 69
Predicated | 103
Dependent | 216
Based | 2139

0   500   1000   1500   2000   2500

■ Freq. COCA
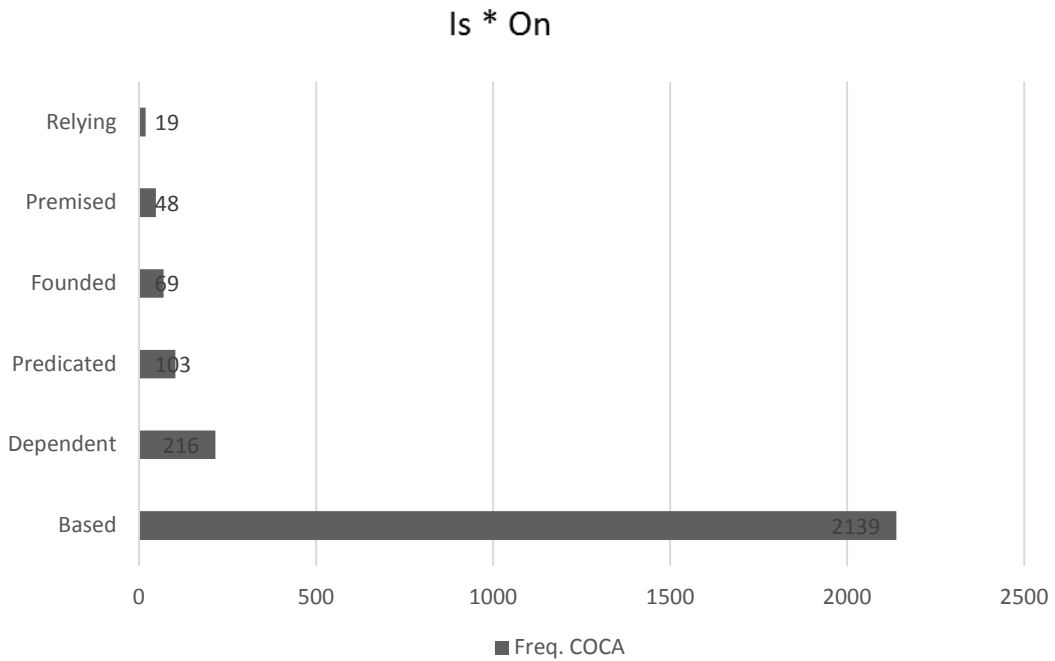
Figure 3. The frequency of *is* on* COCA.

The last construction that was searched in COCA was *In a(n) * to* which revealed the most frequent variable as in *an effort to* with 4,906 occurrences and *in an environment to* as the least frequent variable which occurs just 1 time. This construction is the only one for which the present study revealed a different trend. In terms of acceptability, *in a position to* was accepted at a higher rate than *in an effort to* demonstrating mean acceptability rates of 3.4 and 3.2, respectively. In addition, *in a position to* was ranked as the likeliest to be used in academic writing by 41% while *in an effort to* garnered a mere 28%. Results of acceptability can be seen in figure 4 below. The variable form that was ranked least acceptable was not *in an environment to* as the corpus suggests, but instead was *in an area to* with the lowest mean acceptability rate of all of the experimental constructions at 1.7; nearly 60% of participants selected it as a construction they would not likely use in academic writing. Reasons for this difference will be explored in the discussion section below.
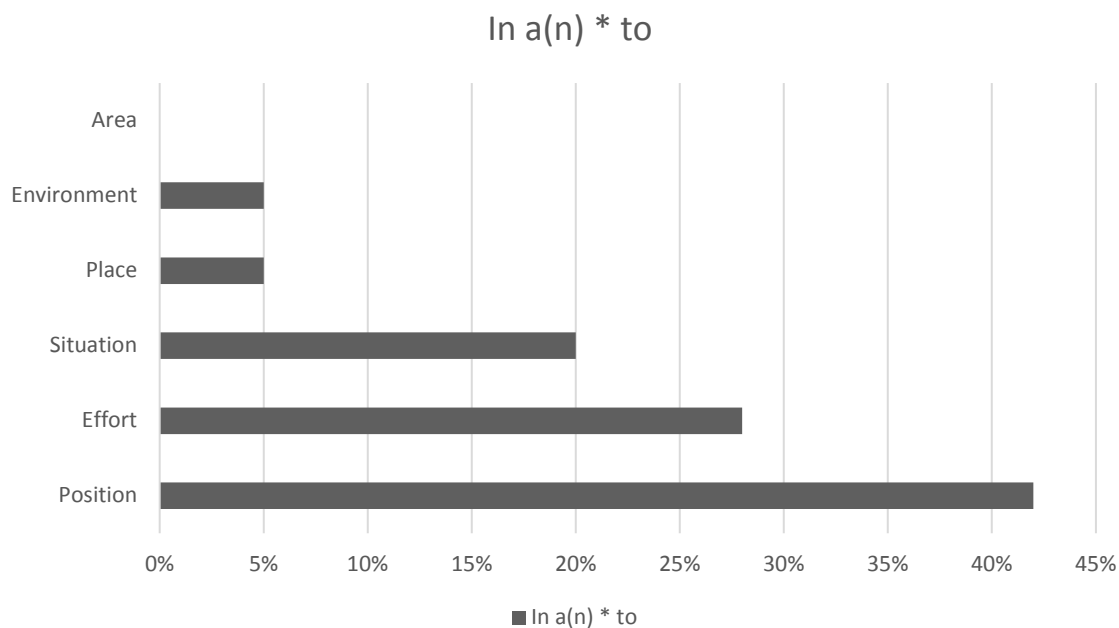

Figure 4. Rate of acceptability of *in a(n) * to.*

In order to address the research questions of the present study, results were grouped based on native or non-native English ability and proficiency, and were subsequently reanalyzed. Results demonstrated a slight difference based on proficiency level as the lower groups were slightly less in line with the corpus while higher groups demonstrated acceptability that was more closely aligned with the corpus. The group with the lowest proficiency, level four, ranked *is dependent on* as the construction they would most likely use in their own writing (over 70% of level four ranked it at number one) even though their mean acceptability rate demonstrates that they rated *is based on* as most acceptable. In addition, native speaking MA students had 100% agreement in the ranking of *in an effort to* as the most acceptable variable form which is in line with the corpus but different than the other proficiency groups.

Lastly, results were analyzed based on status as a native or second language learner of English. Findings demonstrate no difference in trend than the overall findings with *is based on, it is clear that*, and *in a position to* as the most acceptable variabilities while *is relying on, it is plain that*, and *in an area to* were selected as the least acceptable variabilities. As discussed above, this is mostly in line with the frequency distribution demonstrated in COCA except for the third construction which predicted *in an effort to* and *in an environment to* as the constructions with the most and least acceptability, respectively.

Relevance to PIE and Second Language Learning

The results from the acceptability survey were extremely useful in answering the research questions of the current study. They demonstrate that acceptability of variable slots within formulaic sequences is heavily impacted by frequency. These findings support hypotheses made by several theories of second language acquisition, but are particularly supportive of Usage-Based Theory. According to Ellis and Wulff (2015), frequency fosters acquisition and exposure

to input, particularly prototypical exemplars, can facilitate language learning. The findings of the present study certainly support this claim as the majority of the formulaic sequences with the most frequent variable slot were accepted at a higher rate than those constructions with lower frequencies.

In addition, there was no difference based on status as a native or non-native speaker, but there was a difference based on proficiency. Results indicated that lower proficiency groups, low intermediate and native speaking freshmen, were less attuned to frequency than higher proficiency groups. The group with highest proficiency overall, native speaking MA student, ranked acceptability at a rate that was nearly identical to frequency of the corpus. These results support predictions made by Usage-based Theory which suggest that exposure to input affects acquisition. Indeed, the present study demonstrates that those with higher proficiency (i.e., more exposure to input) were more attuned to frequency and vice versa.

The third construction, *in a(n) to,* was the only construction with an acceptability rating that differed from the corpus. This finding is also supportive of Usage-Based Theory as Ellis claims that non-prototypical examples, or those that are less frequent, will be acquired later than prototypical examples. If we look at the frequency of the construction *in a(n) * to* , 17, 118, and compare it to the overall frequency of the other two constructions *it is * that*, 23,530, and *is * on*, 32,032 we can see that *in a(n) * to* occurs far less frequently than the other two constructions. Therefore, it is likely that this construction demonstrated a different trend than the one revealed in the corpus as a result of limited exposure in the input. This is further supported by the fact that the construction with the highest frequency, *is * on,* is the construction with an acceptability rate most closely associated to the results from the corpus.

These findings have several implications for the field of second language acquisition and pedagogy. First, as suggested by Celce-Murcia (2008), formulaic sequences should have their own pedagogic focus in a language classroom. Previous studies have indicated that formulaic sequences are easily processed and make up a large amount of English discourse. This, in conjunction with results from the present study, provide support for Celce-Murcia's classification of formulaic competence as its own category in her teaching model. Furthermore, Biber & Cortes (2004) prove that formulaic sequences have specific functions in a text, and thus, should be relied upon in speaking and writing instruction. Schmit & Carter (2004) pose a question that is relevant to the findings of the current study which asks, "does giving attention to formulaic sequences increase the chances of their acquisition?" The results of the present study provide a clear answer to this question, yes. More frequent exposure to a desired formulaic sequence should heighten the chance of acquisition.

Though the results from the present study are relevant, there are several limitations that must be discussed. First, this study was limited by its participants in several ways. In particular, amount of participants was not evenly distributed across groups. The lower proficiency group had only 10 participants, the highest had only 5 while the mid-proficiency group had 24. Results would have been more reflective if participants had been more evenly distributed. In addition, L1 and age were fairly similar across all participants and may have yielded more interesting results if a wider array of L1s and age groups were available. In addition to limitations as a result of participants, there may also have been some limitations related to the acceptability survey. As the same survey was given to all groups, it can be argued that it was more difficult for lower levels to comprehend which may have impacted their selections. While this is certainly possible, the researcher determined that changing the survey would have resulted in more limitations.

The results from the present study only begin to fill the gap in research concerning FS, frequency, degree of fixedness, and differences across proficiency levels first mentioned by Biber *et al.* (2013). Several studies can be conducted to strengthen these findings. In particular, a study looking at second language learners enrolled in major courses and their responses to frequent formulaic sequences in corresponding disciplines could be conducted to test for effects of discipline on acceptability. Furthermore, studies examining production of common FS with variable slots could be conducted to determine if higher frequency occurrences are also produced more.

References

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university

    teaching and textbooks. *Applied Linguistics*, *25*, 371-405.

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language*

    and Computers, *26*, 181-190.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of*

    spoken and written English. London: Longman.

Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language

    teaching. In E. Alcon Soler & M. P. Safont Jorda (Eds.) *Intercultural language use and*

    language learning (pp. 41-57). Dordrecht, Netherlands: Springer.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than

    nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*, 72-

    89.

Cortes, V. (2002). Lexical bundles in freshman composition. In R. Reppen, S. M. Fitzmaurice, &

    D. Biber (Eds.), *Using corpora to explore linguistic variation* (pp. 131–145). Amsterdam:

    John Benjamins Publishing Company.

Ellis, N. C., & Wulff, S. (2015). Usage–based approaches to SLA. In *Handbook of cognitive*

    linguistics (Vol. 39, pp. 409-432). Walter de Gruyter.

Ellis, N. C., Simpson-vlach, R., & Maynard, C. (2008). Formulaic language in native and second

    language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL*

    Quarterly, *42*, 375-396

Ellis, N. C. (2008). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In M. H. Long & C. Doughty (Eds.), *Handbook of second and foreign language teaching.* Oxford, UK: Blackwell.

Ellis, N. (1996). Sequencing in SLA: phonological memory, chunking, and points of order. *Studies in Second Language Acquisition*, 18, 91–126.

Erman, B. and B. Warren (2000) The idiom principle and the open-choice principle. *Text , 20*, 29–62.

Gass, S., Mackey, A., & Ross-Feldman, L. (2005). Task-based interactions in classroom and laboratory settings. *Language Learning*, *55*, 575-611.

Hopper, P. J. (1998). Emergent grammar. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 155-176). Mahwah, NJ: Erlbaum.

Hunston, S. (2002). Pattern grammar, language teaching, and linguistic variation. *Using corpora to explore linguistic variation*, 167-183.

Moon, R. 1997. Vocabulary connections: Multi-word items in English. In *Vocabulary: Description, Acquisition and Pedagogy*, N. Schmitt and M. McCarthy (Eds.), 40–63. Cambridge, UK: Cambridge University Press.

Nattinger, J. & DeCarrico, J. . 1992. *Lexical phrases and language teaching*. Oxford, UK: Oxford University Press.

Oakey, D. (2002). Formulaic language in English academic writing. In R. Reppen, S. Fitzmaurice & D. Biber (Eds.), Using corpora to explore linguistic variation. (pp. 111-129). Philadelphia, PA: John Benjamins.

Pawley, A. & Snyder, F. (1983) Two puzzles for linguistic theory: nativelike selection

and nativelike fluency. In J. Richards and R. Schmidt (Eds.), *Language and

communication*. (pp. 191–226). London, UK: Longman.

Renouf, A., & Sinclair, J. M. (1991). Collocational frameworks in English. In K. Aijmer, & B.

Altenberg (Eds.), *English corpus linguistics* (pp. 128–143). London, UK: Longman.

Schmidt, R., and Carter, R. (2004). Formulaic sequences in action: An introduction. In N.

Schmitt (Ed.), *Formulaic sequences: acquisition, processing and use*. (pp. 1–22).

Amsterdam: John Benjamins.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing

development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for

Academic Purposes*, *12*, 214-225.

Read, J., & Nation, P. (2004). Measurement of formulaic sequences. In N. Schmitt (Ed.)

Formulaic sequences: Acquisition, processing and use (pp. 23-35). Amsterdam: John

Benjamins.

Reppen, R., Fitzmaurice, S. M., & Biber, D. (Eds.). (2002). *Using corpora to explore linguistic

variation* (Vol. 9). Philadelphia, PA: John Benjamins Publishing.

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*,

13-33.

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge, New York: Cambridge

University Press.

Appendix: Acceptability survey

1. Please select your gender:
   a. Male
   b. Female
   c. Do not wish to state

2. What is your major?

3. How long have you been studying English?

4. Please select your age:
   a. 18-25
   b. 26-35
   c. 36-45
   d. 46 & older

5. Please select your native language (the language you have spoken from childhood):

   a. Chinese   d. German
   b. Arabic   e. Russian
   c. Persian   f. Other: _____

For numbers 6-26 below, please circle only *one* answer. You should rely on your own opinion of English to answer the questions. There are no wrong answers. You will mark the sentences as: Acceptable ☺ ◄────────────────► ☹ Unacceptable

6. Change Theory is founded on the idea that teachers can change their behavior. I find the underlined portion of the sentence _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

7. Change Theory is premised on the idea that teachers can change their behavior. I find the underlined phrase_____ to use in writing.

   a. Acceptable   b. Somewhat acceptable
   c. Somewhat unacceptable   d. Unacceptable

8. Change Theory is based on the idea that teachers can change their behavior. I find the underlined phrase _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

9. Change Theory is predicated on the idea that teachers can change their behavior. I find the underlined portion of the sentence _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

10. Change Theory is dependent on the idea that teachers can change their behavior. I find the underlined phrase _____ to use in writing.

    a. Acceptable   b. Somewhat Acceptable

    c. Somewhat unacceptable   d. Unacceptable

11. Change Theory is relying on the idea that teachers can change their behavior. I find the underlined portion of the sentence _____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

12. Please rank the phrases below in the order that you would most likely use them in writing. **1** being the one you would ***most likely*** use and **6** being the one you would ***least likely*** use.

    _____ Is based one   _____Is predicated on

    _____Is dependent on   _____Is premised on

    _____Is founded on   _____Is relying on

13. It is clear that students need explicit spelling instruction. I find the underlined portion of the sentence _____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

14. It is true that students need explicit spelling instruction. I find the underlined phrase_____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

15. It is obvious that students need explicit spelling instruction. I find the underlined phrase _____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

16. It is certain that students need explicit spelling instruction. I find the underlined portion of the sentence _____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

17. It is plain that students need explicit spelling instruction. I find the underlined phrase _____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

18. It is observed that students need explicit spelling instruction. I find the underlined portion of the sentence _____ to use in writing.

    a. Acceptable   b. Somewhat acceptable

    c. Somewhat unacceptable   d. Unacceptable

19. Please rank the phrases below in the order that you would most likely use them in writing. **1** being the one you would ***most likely*** use and **6** being the one you would ***least likely*** use.

_____ It is clear that   _____ It is certain that

_____ It is true that   _____ It is plain that

_____It is obvious that   _____ It is observed that

20. Politicians who win elections are <u>in a position to</u> implement their ideas. I find the underlined portion of the sentence _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

21. Politicians who win elections are <u>in a place to</u> implement their ideas. I find the underlined phrase _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

22. Politicians win elections <u>in an effort to</u> implement their ideas. I find the underlined phrase_____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

23. Politicians who win elections are <u>in an environment to</u> implement their ideas. I find the underlined phrase _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

24. Politicians who win elections are <u>in a situation to</u> implement their ideas. I find the underlined portion of the sentence _____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

25. Politicians who win elections are <u>in an area to</u> implement their ideas. I find the underlined phrase_____ to use in writing.

   a. Acceptable   b. Somewhat acceptable

   c. Somewhat unacceptable   d. Unacceptable

26. Please rank the phrases below in the order that you would most likely use them in writing. **1** being the one you would ***most likely*** use and **6** being the one you would ***least likely*** use.

_____In a position to   _____In a situation to

_____In an effort to   _____In an environment to

_____In a place to   _____In an area to

Thank You ☺