TESL-EJ Teaching English as a Second or Foreign Language
The Electronic Journal for English as a Second Language

Increasing Lexical Bundles in the Learner Lexicon: Binomial Expressions and the Academic Word List

August 2020 – Volume 24, Number 2

Theresa A. Antes
University of Florida
antes@ufl.edu

Errin Beck
Portland State University
<errink@pdx.edu>

Abstract

In this article, we present the results of an eight-week in-class intervention, during which binomial expressions comprised of academic vocabulary (results and conclusions, facts and circumstances) were presented to intermediate-level learners of English. Constraints that govern the preferred order of highly fixed binomials were explained as the binomials were introduced. Through a combination of pre- and post- preference tests and a post-course repetition task, we demonstrate that learners were able to learn the preferred order of the targeted binomial expressions, as well as to generalize constraints to new binomials. Data from the repetition task provide evidence that the binomials became part of the learners' lexicon: sentences containing binomials in their preferred order were repeated with ease. Sentences with binomials in their non-preferred order, on the other hand, were either spontaneously reordered or were repeated with disfluencies of various types. Learners again displayed an ability to generalize information about constraints to new binomials. Comparisons with native speakers of English and with corpus data provide further evidence that the learners in this study have become more native-like in their control of binomial expressions following the intervention. We conclude with pedagogical implications.s

There is a wide and ever-growing body of literature emphasizing vocabulary knowledge and learning among language learners, and it is generally accepted that vocabulary range is a predictor for competency in the four core skills, reading, writing, listening, and speaking (Nation, 2001; Singleton, 1999). Exactly what should be taught and how has been the subject of much recent debate, however. In his monograph examining phraseology and academic writing, Howarth (1996) noted that advanced learners of English frequently experience difficulty at the clause level, i.e., in determining which words function as "composite units" (p. 14). Howarth argued that idiomatic expressions pose less difficulty for learners than items that are conventionally used together by native speakers (collocations, multiword units (MWUs), bundles), as the former carry a single

meaning that can be memorized by learners. The latter, on the other hand, despite their literal meaning, are often restricted in form in some way that is not immediately apparent. Howarth notes that learners use fewer of these conventional forms, instead producing more freely combining forms, which are more difficult to process (p. 262). Arnon and Christiansen (2017) concur, noting that "adults are both less likely to extract MWUs and less capable of benefiting from them in the process of learning" (p. 623)

Coxhead (2000) recognized that successful language learners wishing to enroll in college-level courses must master lexical items that frequently occur in academic settings. Through the creation of the Academic Word List (AWL), she encouraged a focus on head words and derived forms that are essential to academic success in a second language. This list, which is divided into sublists based on word frequency, has been influential in the choice of vocabulary taught to L2 learners of English (both foreign and second) since its creation; however, researchers now argue that teaching of discrete items does not go far enough. Like Howarth, Byrd and Coxhead (2010) recognized the importance of phraseology over discrete items, pointing out, for example "the importance of long, complex noun phrases" in academic writing (p. 41). They advocated an approach to teaching vocabulary that would allow presentation of both discrete words and lexical bundles, one that would allow learners to examine the bundles explicitly in class, understanding individual components and their use, and that would then put the components back together so that the whole could be used "accurately and appropriately" in speaking and writing (p. 55). They noted that explicit teaching of such sequences was likely necessary, due to a possible lack of exposure from input. Schmitt, Jiang, and Grabe (2011) corroborated this research, demonstrating a nearly linear relationship between vocabulary knowledge for a text and comprehension of that text. However, even when a learner knows the meaning of many words, a lack of experience with vocabulary beyond discrete words can cause problems at both a receptive and productive level. Arnon and Christiansen (2017) and Christiansen and Arnon (2017) further encouraged an explicit approach to teaching MWUs, noting that because L2 learners "already know that words exist, they may focus on individual words rather than multiword combinations" (Christiansen & Arnon, 2017, 547).

One area of vocabulary that has been largely neglected in the English language learner (ELL) literature is that of binomial expressions. Binomial expressions can be defined as "coordinated word pairs whose lexical elements share the same word class" (Mollin, 2014, p.1). Examples include: *give and take* or *salt and pepper*. Within academic speech or writing, these conjoined expressions become especially important: articles frequently contain headers such as *results and conclusions*; presentations may necessarily include verb phrases such as *introduce and explain*. Learners' ability to use conventionalized binomials, in the order preferred by native speakers, rather than coining new one-time expressions (hapax legomena, Mollin, 2014), is both an indication of their proficiency in the language and essential to their overall academic success.

To date, there has been virtually no attempt to teach academic vocabulary and corresponding binomial expressions in tandem, despite the importance of these expressions in academic settings. In this article, we present results from both intuition data and production data, in the form of a repetition task, showing that learners' ability to recognize and produce canonical forms of binomial expressions increased dramatically following an eight-week intervention. Results suggest that a targeted intervention, combining vocabulary instruction with an explanation of constraints governing binomial word order, will allow ELLs to understand better not only the expressions that are specifically focused on in class, but also others encountered, thus leading to greater understanding and, eventually, the opportunity for increased use.

Review of the Literature

Knowledge and Instruction of MWEs

Recent vocabulary-related research includes an increased focus on multi-word expressions (MWEs) and their role in effective oral and written expression of ideas, especially with regard to organizing discourse and rhetorical moves (Conrad, Biber & Cortes, 2004; Ellis & Simpson-Vlach, 2009; Meunier, 2012). MWEs are significant indicators of register and genre. Studies have shown that when those expressions are left out, academic writing can be perceived as unnatural, foreign, or not genre-appropriate (AlHassan & Wood, 2015; El Dakhs, Tasneem Prue & Ijaz, 2017; Ellis, Simpson-Vlach, & Maynard, 2008; Howarth, 1996; Li & Schmitt, 2009). Extensive literature has shown that despite the ubiquitousness of such bundles in the language, ELLs have neither the same intuitions concerning MWEs as native speakers nor the necessary accuracy and fluency to use these expressions effectively (Bestgen & Granger, 2014; Boers & Lindstromberg, 2009; Coxhead & Byrd, 2007; El-Dakhs et al., 2017; Hoang and Boers 2016; Meunier, 2012; Schmitt, 2000; Wray, 2002). El Dakhs et al. (2017) argue that L2 classes must include formulaic sequences / MWEs in order to bolster language acquisition, particularly where writing is concerned (p. 22).

It follows, therefore, that MWEs deserve dedicated attention as well as efficient and effective instruction in curricula preparing language learners for university coursework. In order for learners to take up a new feature, however they must first take notice of that feature. Numerous previous studies (AlHassan & Wood, 2015; Arnon & Christiansen, 2017; Boers & Lindstromberg, 2012; Christiansen & Arnon, 2017; Hoang & Boers, 2016) have indicated that because they are often transparent, MWEs frequently escape learners' notice. Theories related to *noticing* and the role of attention in language learning are based on the notion that our memory of information is stabilized and lengthened by conscious awareness (e.g., Schmitt, 1990). Following suggestions that explicit attention to lexical forms enhances their acquisition in a second language, researchers contrasted meaning-based instruction and form-focused instruction. Laufer and Girsai (2008) demonstrated that form-focused instruction resulted in a higher number of single vocabulary words and collocations in content-based tasks than meaning based instruction alone. Likewise, Tajeddin and Daraee (2013) compared EFL learners' retention of new vocabulary words in form-focused versus non-form-focused tasks, and found higher retention for the former, regardless of the nature of the latter (comprehension question or message orientation). The authors of both studies argued that learners' ability to notice and retain words and expressions was enhanced during the instruction phase that explicitly focused their attention on form. Recent studies have further affirmed that explicit instruction can aid in the uptake of MWEs. AlHassan and Wood (2015), for example, demonstrated that learners' ability to use MWEs in their academic writing can be enhanced through explicit instruction, regardless of proficiency level, and even in the long term. In a case study, Wood (2009) indicated a rise in the number of multi-word phrases used in spontaneous speech after instruction, resulting in increased fluency. El-Dakhs et al. (2017) showed that recognition of MWEs increased in their study, but that the MWEs were not produced in spontaneous speech, mirroring other studies that indicate that production abilities lag behind comprehension abilities. Furthering an exploratory study by Jones and Haywood (2004) that showed that learners' awareness of phrases in academic language increased, Peters and Pauwels' (2015) replication study presented a comparatively positive set of results. They showed that after explicit instruction on formulaic sequences (FS), learners recognized and produced an increased number of FS in both cued response and spontaneous written production exercises. Despite the positive results on learner uptake of MWEs during previous form-focused studies, however, and the potential ease of teaching the governing constraints of binomial expressions, this remains one area that has not yet been studied.

Binomial Expressions as MWEs

The term multi-word expressions encompasses a variety of types, such as phrasal verbs, collocations, fixed idioms. One type of MWE that has yet to be given much attention in the L2 research is binomial expressions. To date, research on binomials expressions has mainly been conducted using corpus data or native speakers of English (Benor & Levy, 2006; Morgan & Levy, 2016; Siyanova-Chanturia & van Heuven, 2011). Mollin (2014) is one of the rare studies that has contrasted native speaker and L2 learner intuitions. Additionally, most previous studies have focused on knowledge (intuition), rather than on the acquisition of these expressions, despite the importance of binomials in helping L2 writers increase the complexity of their writing.

Morgan and Levy (2016) investigated the underlying factors for how speakers store and combine single words into binomials. They examined the roles of and relationship between abstract knowledge of ordering constraints versus direct experience with binomials, considering how both new and known expressions are processed. Their results suggested a 'trade-off' theory in which a speaker switches between sources of information in the processing of binomials depending on how novel the expression is to that speaker. Binomials with which speakers had direct experience, that is, frequent ones, were retrieved from the mental lexicon as chunks or exemplars. Less frequent binomials, however, were processed according to underlying constraints, meaning that abstract knowledge served as the source of information needed to correctly order the units of the binomial expressions. Logically, L2 speakers have neither the same mental lexicon nor the direct experience with binomials as native speakers. Yet, it may be inferred that by boosting that direct experience as well as learning the constraints that govern the order of binomials, learners may enhance their knowledge of the order of binomial expressions and be able to apply the dual-pronged 'trade-off' approach suggested by Morgan and Levy (2016).

In a highly controlled study, Mollin (2014) further investigated the question of word order within binomials, and demonstrated that non-native speaker intuitions differ from that of native speakers. To control for overall frequency, Mollin chose 32 binomials that appeared in the British National Corpus 50 or more times, and then presented them to her participants in both canonical and reversed order. Participants indicated, on a 7-point Likert scale, their preferred order. Non-native speakers were further asked to indicate if they were familiar with the expression or not. Only judgments for which participants indicated a familiarity were included. Results indicated a reluctance by native speakers to choose "only possible response" for any binomials, including those that were entirely frozen. Despite this, native speaker intuitions about which binomials were relatively fixed were far more accurate than those of non-native speakers. Mollin argued that frequency alone could not account for this distinction, as overall frequency had been controlled for. Instead, she argued that setting frequency — i.e., the number of contacts that one has with an expression — likely plays a role in acquisition. Mollin concluded that fixed binomials should be taught explicitly, while reversible binomials (those that fall in the 50-74.99% range of fixedness, in her study) should not be. The support found for exemplar learning corroborates other studies (Morgan & Levy, 2016), but does not examine the role of constraints, and whether explicit teaching of both binomials and their constraints might be combined to the benefit of non-native learners of English.

This study therefore aims to extend the existing literature on binomial expressions by comparing the intuitions of native speakers with those of non-native speakers, both before and after the latter received explicit instruction on these expressions. We also offer expansion by examining not only non-native speaker intuitions on the order of binomials but also their ability to produce binomials during a repetition task. We investigate to what extent explicit instruction can aid non-native speakers in building intuitions on the order of highly fixed binomials, whether this information is

generalizable to new binomials, and to what extent production during a repetition task will reflect intuitions shown on the preference task.

The following research questions guide our study:

1. On the preference task prior to the intervention, what evidence is there that learners are or are not familiar with binomial word order, particularly in comparison to native-speaker data?

2. Do results from the end of term preference task indicate that learner intuitions have become more native-like following the in-class interventions? If so, do the changes affect only those binomials that were explicitly taught, or is there evidence of generalization to new binomials?

3. On the production task, in what ways (if any) do ELLs manifest knowledge of binomial word order that can be attributed to the interventions?

Methodology

This study consisted of a pre-test and a post-test, which flanked eight weeks of classroom instruction. The first data collection session took place in the second week of the 10-week quarter, outside of class in a computer lab. Participants read and signed the informed consent form, then completed a brief background questionnaire. A free vocabulary size test (VST) developed by Paul Nation and available online at https://my.vocabularysize.com was then administered, followed by a preference task (see below for details). Participants attended courses for eight weeks after the first data collection. In their Reading class (one of their four core classes), students received instruction on the AWL, binomials based on the AWL, and constraints governing the order of those binomials (see below). After the tenth week of the term, students returned to the computer lab to perform the post-test tasks, a second administration of the preference task, and a repetition task.

Participants

The three ELL participants in this study were enrolled full-time in a pre-academic track at an intensive English program at Portland State University. They had lived in the United States as full-time students for three to six months prior to the beginning of the study. The participants, two males and one female, came from three L1 backgrounds (Mandarin, Thai, and Japanese) and had similar proficiency levels. Courses they were enrolled in have course aims that span CEFR B2 and C1. A $10 honorarium was given to each participant after the pre-test data collection session and again after the post-test session.

To ensure that the binomials chosen for the study were appropriate for a focus with English language learners at this point in their academic career, they were also tested on a peer group of native-speaking students. It was assumed that the native-speaking students would know the individual academic words comprising the binomials, but we held no assumptions concerning their knowledge of the binomials themselves nor about their ability to produce them. Undergraduates, as emerging academics, may not exercise the same control of MWEs as expert speakers and writers commonly featured in corpus data. Previous studies (Coxhead and Byrd, 2007; Hoey, 2005; Hyland, 2012) have noted important differences between novice and expert speech and writing, even in the L1. A peer group was thus included in this study as a means of validating our materials. It is fair to say that most students in an IELP aspire to enroll at some point in college-level courses in an English-speaking country. By determining that the binomials chosen were, in fact, known both receptively and productively to a population of language users amongst whom the ELLs seek to be counted, we were better able to target specific MWEs for our study. The controls in this study were volunteers from the University of Florida. They were all undergraduate monolingual L1 English speakers. We began with 30 controls and, in order to reduce confounding variables,

narrowed this sample to nine individuals with comparable vocabulary size test scores, with no stated learning disabilities. Demographic information for these participants is given in Table 1, below.

Table 1. **Participant Demographic Information.**

Participant #	VST Score	Gender	Native Language	Age
EL1	5000	M	Chinese	21
EL2	5600	M	Thai	24
EL3	8000	F	Japanese	21
C1	14800	F	English	20
C2	14200	F	English	21
C3	14800	F	English	19
C4	15000	F	English	20
C5	14200	F	English	18
C6	14600	F	English	21
C7	12200	M	English	21
C8	12700	F	English	19
C9	14700	F	English	18

Selection of Binomial Pairs

Two main sources were used to select binomials for the data collection instruments as well as for instruction. First, the Corpus of Contemporary American English (COCA) (Davies, 2014) was examined to find binomials that incorporated a word from the AWL, at relatively high rates of fixedness. Second, materials used in class were then consulted to determine which head words from the AWL learners would encounter during the term, making introduction of binomials containing those head words a logical extension of the class discussion. We arrived at the list of ten headwords shown in Table 2, each associated with two binomials, one more frequent and one less frequent, for a total of twenty binomials. Degree of fixedness for each binomial was determined following Mollin (2014) and Morgan and Levy (2016): the number of occurrences for "A and B" divided by the total number of occurrences for "A and B" and "B and A." A score of 1.00 indicates that a binomial is totally frozen in one order; a score of .50 indicates that a binomial is reversible. Each binomial was independently determined by the two authors to follow at least one of three constraints identified by Morgan and Levy (2016): perceptual markedness, iconic/scalar sequencing, and metrical length (see below). Binomials were introduced by the classroom instructor throughout the quarter, along with an explanation of the constraint(s) that they followed.

Table 2. Targeted binomials.

Targeted Binomials (AWL keyword in bold)	Constraint	AWL sublist	COCA Frequency / Fixedness	Part of Speech
Fair and **accurate**	Length	6	74 / .86	ADJ
Clear and **accurate**	Length	6	22 / .92	ADJ
Large and **complex**	Length	2	80 / 1.00	ADJ
Vast and **complex**	Length	2	26 / .87	V
Identify and describe	Scalar / iconicity	1	58 / .95	V
Identify and evaluate	Scalar / iconicity	1	28 / 1.00	V
Schools and **community**	Perceptual markedness	2	105 / .96	N
Teachers and **community**	Perceptual markedness	2	24 / .96	N
Wealth and **status**	Length	4	80 / .86	N
Class and **status**	Length	4	23 / .96	N
Findings and **conclusions**	Scalar / iconicity	2	63 / .98	N
Results and **conclusions**	Scalar / iconicity	2	43 / .98	N
Individual and collective	Perceptual markedness	1	250 / .96	ADJ
Individual and communal	Perceptual markedness	8	25 / .83	ADJ
Understand and **appreciate**	Scalar / iconicity	8	135 / .83	V
Recognize and **appreciate**	Scalar / iconicity	8	25 / 1.00	V
Attract and retain	Scalar / iconicity	7	85 / .93	V
Attract and hire	Scalar / iconicity	7	20 / 1.00	V
Facts and **circumstances**	Length	3	80 / 1.00	N
Time and **circumstances**	Length	3	29 / .97	N

Intervention

The twenty binomial expressions were folded into an already robust vocabulary component of the curriculum of an academic reading course. All three participants were enrolled in the same section of the course. In that course, students focus on sublists 4 and 5 of the AWL, which are further broken down into smaller, more manageable lists. It is expected that students would already have studied words from earlier sublists (1-3), as these are higher in frequency, but not yet those from later sublists (6-8), as these are lower in frequency. We sought a balance across keywords that would already be known to students, those currently under study, and those that might be unknown yet likely to appear in authentic materials. Instructors introduce, practice, and assess them over the course of the 10-week term. Exposure to the full AWL is reinforced through both the textbooks and course materials used in the reading course and in other skills (speaking/listening, writing, grammar) courses in the program. In the reading course, AWL study is conducted both in and out of class. For example, students often use Quizlet to practice the form and meaning of the target vocabulary. AWL target vocabulary is most commonly presented in single word units, in lists. Some multi-word expressions that include AWL words and are related to certain language functions needed in academic writing (e.g., *current research indicates; according to the author*), are also taught.

To introduce the concept of binomial expressions, a power point presentation, in which binomials were referred to as *paired expressions*, was used at the beginning of the term. The presentation (see Appendix A) offered a rationale for learning these types of expressions and compared them to

collocations, a term students are familiar with. It included common, non-academic examples of binomials, and explained the rules that determine the order of words.

In line with the presentation and practice of AWL vocabulary in small batches, the twenty binomial expressions were divided into four groups of five expressions each. Practice worksheets (see Appendix B) included the meaning of the binomial expressions, contextual and rhetorical information on their usage, cloze activities, explanations of governing constraints, and deductive exercises related to those constraints. The worksheets were completed in class, with instructor guidance and peer dyad/group input and interactions. To complement these instructional materials, the instructor pointed out examples of binomials as they appeared in textbook reading texts.

Directing learners' attention to the target binomials and a planned, form-focused approach served as the basis of the instructional methodology. This approach raised the relevance and saliency of the binomials and also minimized the additional cognitive load associated with learning new expressions. In addition, the noticing, repetition, and practice of both the target binomial expressions and binomials that were encountered spontaneously was a recurring action in class. This presentation style established the binomial expressions as fixed chunks rather than simply a randomly selected pair of words linked with a conjunction. This approach also ensured that the learners would encounter the targeted expressions in multiple modes and contexts. The explanation of the constraints was also a recurring occurrence in the teaching of the target expressions.

No overt testing of the binomial expressions was conducted by the classroom instructor. Thus, any uptake of binomials or the constraints related to them can be attributed to their use and practice in class and the relevance attributed to them by the students, rather than to an attempt to memorize them for the simple purpose of passing an exam.

Preference Task

A preference task was created in order to determine the extent to which speakers' selection of binomial ordering mirrored that found in COCA. This task was administered twice to the English language learners (pre- and post-instruction), and once to the native-speaking controls. The task consisted of three practice items, the 20 targeted binomials and an additional 80 distractor binomials. (See Appendix C for the entire list.) The latter ranged from fully frozen to mostly reversible (51% fixed), and included binomials that were similar to the targeted binomials (same headword, but different common word, for example), in order to gauge generalization across binomials that shared headwords, as well as those that shared constraints but no vocabulary items. In building the preference task, we followed Mollin (2014), with minor modifications. In her study, Mollin found that speakers were highly reluctant to choose "only possible choice" as a response. We further found the terms *strong preference* and *mild preference* reflective of British English, and modified these terms slightly for speakers of American English. Participants thus saw a binomial on the left and its opposing form on the right with the following choices beneath them: *best, highly preferred, somewhat preferred, neutral, somewhat preferred, highly preferred, best*. Only one binomial pair appeared per screen; participants made a choice and then moved to the next screen. Options alternated randomly between degrees of fixedness (frozen, highly fixed, moderately fixed, reversible), and whether the more canonical form of the binomial appeared on the left or the right of the screen. Participants were allowed to skip an expression; few chose to do so, but one ELL did skip several expressions on the pre-test. Separate versions of the preference task were created for the pre- and post-tests, with different orders of presentation of the binomials, and position changes (left versus right) for the canonical and reversed binomials. Unfortunately, the creation of two separate versions resulted in an error. We had chosen the binomials pairs *understand and*

appreciate / recognize and appreciate and *hire and retain / attract and retain* to be targeted in class. On the post-test and on the repetition task, they appeared as such. On the pretest, however, we mistakenly asked participants to rate *understand and recognize* instead of *understand and appreciate*, and *hire and attract* instead of *hire and retain*. Because we cannot compare pre- to post-test results for these expressions for the ELLs, and because the controls were assigned randomly to either the pre- or post-test version, we have eliminated these four binomials from the preference task analysis, resulting in 96 binomials that will be examined.

For the purposes of data analysis, scores were assigned to each of the remaining 96 binomials on a numerical basis (1-7) according to the option that participants chose on the Likert scale. Questions that were skipped were assigned an 8, as it was observed that participants actively chose not to answer the question, rather than to opt for 4 (neutral). Because of the rarity of such responses among most participants, and the change of behavior by one ELL from pre- to post-test, it seemed more appropriate to score these skipped responses as a non-response (8) than to omit them. Binomials were then regularized, so that all figures represent a participant's response as if the canonical order had appeared on the left-hand side of the screen. A number close to one therefore indicates that participants preferred the canonical ordering for a given binomial or group of binomials, four indicates neutrality, and a score close to seven indicates that the reverse order was preferred.

Repetition task

To further measure participants' knowledge of binomials, we incorporated a repetition task in this study. Native-speaking controls completed this task during the same session as the vocabulary size test and the preference task, ELLs at the end of the quarter, following the post-instruction preference task. The repetition task consisted of 76 items: four practice sentences, 32 distractor sentences that included a variety of general vocabulary difficulties for ELLs (idiomatic expressions such as *come up with* vs. *come back to*), 20 sentences with the targeted binomials, and 20 sentences with distractor binomials. (See Appendix D for the entire list of sentences.) Sentences ranged from 15-20 syllables each, lasting 3.5-4 seconds each. All were spoken by the same speaker, a middle-aged Caucasian woman with extensive teaching and broadcasting experience. Participants heard each sentence once. It was followed by approximately 250 ms of silence, and then a beep, signaling the speakers to begin. They were given ten seconds to repeat what they had heard; they then again heard a beep before the next sentence started. Sentences were created to form pairs: the binomials appeared in sentences that were exactly alike for each pair, with the binomial in the middle of the sentence. In one sentence, the binomial appeared in its canonical order and in the other, in its reversed order. Examples of both distractor sentences and binomials are given in 1-4, below:

1. Our boss is going to do away with these books and reorganize the office.
2. Our boss is going to do us all a favor and reorganize the office.
3. News reports must be clear and accurate to avoid bias.
4. News reports must be accurate and fair to avoid bias.

All binomials had previously been seen in the preference task. Distractor sentences were also paired, maintaining sentence structure to the extent possible within pairs, and providing a similar repetition within distractors and binomials. One sentence from each pair was distributed in the first half of the repetition task and the other in the second half, in an effort to provide for maximum memory decay before the second encounter. The task lasted a total of 22 minutes, with instructions. Participants were allowed to pause the task and take a break whenever they wanted. A few did, but never for more than 30-60 seconds at a time. While cognitively challenging, the task was successfully completed by all participants.

Results

Preference Task

Data from the preference task were first analyzed to determine if participant responses mirrored usage of binomials as found in COCA. To gauge the extent to which participants were sensitive to degrees of variations in word order among the binomials, we divided the 96 binomials into bands of fixedness, moving from most fixed (entirely frozen) to most reversible. Table 3 shows the mean and standard deviation for Controls and ELLs for each band, as well as for the Aggregate group of targeted binomials.

Table 3. Mean and St. Dev. For Control and ELL responses on Preference Task.

	Control Mean (n = 9)	Control St. Dev.	ELL Pretest Mean (n = 3)	ELL Pretest St. Dev.	ELL Post-test Mean (n = 3)	ELL Post-test St Dev.
Frozen (100) (n = 11)	2.495	.591	3.394	.378	2.818	1.000
97 – 99 (n = 12)	2.306	.402	2.538	.833	1.860	.774
94 – 96 (n = 14)	2.691	.661	3.738	.352	2.214	1.067
91 – 93 (n = 11)	3.020	.467	3.939	.845	2.727	.328
86 – 90 (n = 12)	3.009	.490	3.861	.488	2.139	.536
81 – 84 (n = 10)	2.933	.812	3.667	.153	2.600	1.418
72 – 76 (n = 7)	3.556	.471	4.333	.082	3.517	.429
64 – 69 (n = 7)	3.634	.447	4.761	1.593	2.761	.705
51 – 62 (n = 12)	4.037	.477	4.473	.614	3.528	.481
Targeted binomials	2.871	.537	3.667	.309	1.968	.695

(1 = Canonical order best; 4 = neutral; 7 = Reversed order best)

The table shows a clear progression for the native speaker controls, from highly fixed binomials (94-100) that consistently fell in the mid-range between somewhat and strongly preferred in their canonical order, to a score of 3 (somewhat preferred) for binomials with a fixedness of 80-93, to a preference that ranged between somewhat preferred and neutral (3.5-3.6) for binomials with lower fixedness ratings. Those binomials that are reversible in nature were recognized as such by the controls, and had the highest mean: 4.04 (neutral). Despite a reluctance to choose "best" (1) for the highly fixed binomials similar to that reported by Mollin (2014), native speakers do mirror the corpus data, suggesting both a familiarity with the binomials and an overall sensitivity to their degree of fixedness. Confident that L1 responses on the preference task reflected COCA fixedness, overall, and that this vocabulary thus met learnability standards within a college-level curriculum, we turned our attention to ELLs.

No similar progression is noted for the ELLs on the pretest preference task, who, with one notable exception, hover around the neutral score. Their score for entirely frozen binomials is a bit lower (3.3), thus between somewhat preferred and neutral, than that for binomials fixed at 97-99 at 2.538. While it is difficult to know if several of the binomials in this second band were already known to them or if this marginally lower score is the byproduct of some lucky guesses, the overall means suggest that this group of learners was largely unfamiliar with the binomials being tested, and hovered around the neutral response. On the post-test, however, participants in the ELL group have shifted their responses to more closely mirror those in the COCA corpus. Means for binomials in the most highly fixed bands (90-100) range from 1.8 – 2.8 (highly preferred to somewhat preferred),

while those in the less highly fixed bands (64-84) range from 2.6 – 3.5 (somewhat preferred to neutral). Those that are generally reversible also fall midway between the somewhat preferred and neutral ranges (3.52). Binomials that were specifically focused on in class were obviously retained by learners, as their canonical order was highly preferred (mean score 1.968).

Inferential statistics conducted after the intervention confirmed that the ELLs' knowledge of binomials shifted to match corpus data, particularly for those binomials that were a specific focus of attention in class. A two-tailed independent t-test was run on the aggregate scores within each band of binomials to determine if the differences observed between the pre- and post-test results were statistically significant. The t-tests showed that differences were significant for two bands of binomials: targeted ($t(4)=3.881$, $p=.018$) and those fixed at 72-76 ($t(4)=3.024$, $p=.039$, while those in the 91-93 band neared significance ($t(4)=2.317$, $p=.081$). While no significant differences were noted amongst binomials in other bands, the direction of scores on the preference task was always downward, toward a more canonical order. This slope was greater for the targeted binomials than for others, and not as great (non-significant) in bands where ELLs demonstrated some previous knowledge of the expressions (the 97-99 band, for example). Large standard deviations within some bands also likely resulted in non-significant results, despite what appear to be important improvements from pre- to post-test. These results suggest that a focus on binomials and their constraints over the course of a 10-week term can lead to greater awareness of these expressions, and improved intuitions across multiple levels of fixedness.

Repetition Task

The repetition task was transcribed and analyzed according to criteria for oral speech established by Cutler (2017). In his study examining spontaneous speech, Cutler determined that four criteria were helpful in identifying the use of formulaic sequences by Japanese speakers of English: filled pauses, unfilled pauses >.25 seconds, syllable lengthening > .4 seconds, and repair or retracing. Following these criteria, we transcribed all sentences produced by participants during the repetition task. We shortened the threshold for unfilled pauses to ≥ .2 seconds to account for the nature of the repetition task, and also noted filled pauses, word omissions, word substitutions, and reorderings. Cutler's conventions of noting unfilled pauses between parentheses (.35) are followed here. Both authors initially transcribed the data from three participants chosen at random from the entire pool of native speakers as well as from the three ELLs, with 98% and 95% inter-rater reliabilities respectively. Deeming this inter-rater reliability acceptable, we chose our nine controls. Only the first author of the study rated these participants. Forty sentences in total contained binomial expressions, counter-balanced to present half in their canonical and half in their reversed order. Transcriptions were carefully analyzed to determine whether participants *maintained word order* of the binomial as they heard it. If not, the following possibilities were coded: binomial *reordered*; binomial *reworded* (one or both terms within the binomial replaced with another, to produce a new binomial, including hapax legomena); binomial *abandoned* or skipped entirely; or *omission* of "and" and one term from within the binomial. Additionally, disfluencies such as pauses and lengthenings were noted. Because of space limitations, we have summarized general results for this task below.

Data from controls are again included to provide insight into the extent to which these binomials are known to a peer group. While clearly more fluent in English than their ELL counterparts, it can be argued that they have no more content knowledge in academic subject areas. Their ability to produce these MWEs, as opposed to simply comprehending them, must therefore not be assumed. As illustrated in Table 4, controls maintained the order of the binomial that they heard in the vast majority of the sentences that they repeated, with little difference between reversed and canonical contexts. Spontaneous reordering of items within the binomial occurred on eight reversed

binomials, by a total of 14 participants, compared to only three canonical binomials. While the number of rewordings shown in Table 5 appears similar at first glance, the transcriptions reveal differences. A total of nine reversed binomials accounted for 13 rewordings, while only four canonical binomials posed difficulty for speakers, with one, "(the) review and evaluation" accounting for five of the ten rewordings. Here, it appears that several of the native speakers first analyzed "review" as a verb, thus leading to later processing difficulties with the second noun in the binomial. Only one speaker abandoned a sentence entirely, and omissions were observed as well. These appear to have affected less frequent and less highly fixed binomials to a greater extent than others. This is a point that will be examined in the discussion section, below.

Table 4. Repetition task results – Control participants.

	Word Order Maintained	Reordered	Reworded	Abandoned	Omissions
Canonical Binomials (max = 180)	160 (88.9%)	3 (1.6%)	10	1	6
Reversed Binomials (max = 180)	153 (85%)	14 (7.8%)	13	0	3

As expected, ELL speech on the repetition task was much less fluent overall than native-speaker speech. This was a highly challenging task, during which participants heard the sentence only once, and had only 250 ms to process what they heard before being prompted to repeat. At the same time, ELLs were clearly already familiar with half of the binomials, which had been specifically taught and practiced in class. We therefore distinguish between targeted and new binomials, a distinction that was not made for the native speakers.

Table 5. Repetition task results – ELLs.

	Word Order Maintained	Reordered	Reworded	Abandoned	Omissions
Targeted binomials, canonical (max = 30)	19 (63.3%)	1	6	1	3
Targeted binomials, reversed (max = 30)	16 (53.3%)	3	6	3	2
New binomials, canonical (max = 30)	12 (40%)	1	5	6	3
New binomials, reversed (max = 30)	6 (20%)	3	2	5	9

As Table 5 illustrates, ELLs were more successful in retaining binomials that had been specifically taught in class, and repeated these binomials with greater success when heard in their canonical order than when reversed. When not repeated as heard, one item of these targeted binomials was often retained, as the column showing rewordings indicates. ELLs reworded a higher number of binomials than did their control counterparts. Within these rewordings, the canonical constraints were frequently observed, even if the creation was novel. New (untaught) binomials were also repeated more successfully in their canonical order, suggesting that ELLs had generalized the constraints that they learned and could apply them with some degree of success to new binomials. It is not surprising that a higher percentage of abandonments occurred in non-targeted binomials, as the sentences were difficult enough that vocabulary that had not been specifically studied was not retained after a single listen. What is noticeable, however, is that omissions occurred to a much greater extent in new reversed binomials than in any others. In these instances, learners retained one item of the binomial, but not the other. The fact that this affected the reversed but not the canonical new binomials is again suggestive of generalization of the constraints. In the canonical

order, the entire new binomial was retained whereas in the reversed order, only one item was held in memory. This is a point to which we will return in the discussion section.

Disfluencies were common throughout the task as learners struggled to remember and produce not only the binomials but also academic vocabulary in general (*bias, structures, architect*). Of the pauses that occurred within the binomials themselves, 13 out of 18 affected reversed binomials, whether targeted (and thus studied throughout the semester) or not. This is again an indication that ELLs were sensitive to the order of the items within the binomials. While less fluent overall in their repetition, reversed binomials provoked greater disfluency than those in their canonical order.

Examples 1 and 2, below, shows speaker performance on two sentences from the repetition task. These sentences contrasted *conclusions and results,* a reversed binomial, with *findings and results*, which participants heard in its canonical order. Native speakers produced both sentences quite effortlessly, although many speakers reduced the reversed binomial to a singular "conclusion." ELLs, on the other hand, struggled to repeat the reversed binomial. The second posed fewer difficulties for all but two speakers, one ELL who deleted the binomial altogether, and two controls, one who corrected a false start at the beginning of the sentence, and one who displayed disfluency during the production of the binomial. ELL difficulties often appear to stem from other vocabulary (the verb *combine*, for example) which places cognitive demands on them. With one exception, when the binomial appeared in its canonical order (sentence 2), they were able to deal with it relatively effortlessly; when it did not (sentence 1), the result was disfluency both following the verb and within the binomial.

1. This researcher combined her conclusions and results in one section. (R)

C1, C2, C8: Exact repetition, no disfluencies.
C3, C4, C5, C6, C7: ***conclusion*** *and results*
C9: This researcher combined her conclusion and results (.5) into one section.
EL1: This researcher combine (.4) his (.9) own (.65) results in one sentence.
EL2: This researcher combined her (.25) conclusion and result (.5) in (.6) one (.7) solution.
EL3: This researcher combined her (1.9) conclu (.4) sion and re (.45) sult …

2. This researcher combined her findings and conclusions in one section. (C)

C1, C2, C3, C4, C6, C8: Exact repetition, no disfluencies.
C5: This researcher- (.35) researcher combined her findings and conclusions in one section.
C9: This researcher combined her findings (.6) and (.3) conclusions in one section.
EL1: This researcher combined his (.85) including and (1.7) in one section.
EL2: This researcher combined her finding and conclusion in one section.
EL3: This recherche- (.35) research combined <uh> (.85) finding and conclusion in (.35) her research.

Sentences 3 and 4 display similar results, again with targeted binomials for the ELLs. In this instance, they heard the canonical order first during the task (*design and build*), and held the entire sentence in memory quite successfully, with one learner substituting the word *actress* for *architect*, and a second omitting the second half of the binomial. When the matched binomial *design and develop* appeared in reverse order, however, learners displayed much greater difficulty with the sentence as a whole as well as with the binomial itself. All learners produced the word *design*, suggesting that they were aware that this term should come first, with one actually placing *developed* after, but with no conjunction. Other vocabulary in the sentence was recalled less

correctly as well, and much longer disfluencies were attested, again suggesting that the learners were distracted by the order of the binomial that they had heard, compared to the order that they were expecting. Here, the ELLs patterned somewhat differently from their peers, who maintained the word order for both binomials as they heard them, despite displaying somewhat longer hesitations for the reversed binomial. As will be noted in the discussion session below, overall binomial frequency may have played a role for the native speakers, as *design and develop* is a low frequency expression. For the ELLs, on the other hand, because both binomials were taught in class, overall binomial frequency likely played less of a role than setting frequency, which in this case was equal.

3. This architect always designs and builds unusual structures. (C)

C1, C2, C3, C4, C5, C8: Exact repetition, no disfluencies.
C6: This architect always bui- (.2) designs and builds unusual structures.
C7: This architect always designs and builds (.2) unusual structures.
C9: This architect(.3)t (.2) always designs and builds (.6) beautiful structures.
EL1: This actress always design and build u- (.55) unusual structure.
EL2: This architect are designed and build (.35) with (.3) unusual …
EL3: This architect designed (1.6) very unusual structures.

4. This architect always develops and designs unusual structures. (R)

C1, C3, C4, C5, C8, C9: Exact repetition, no disfluencies.
C2: This architect always (.4) develops and designs unusual structures.
C6: This architect always develops and designs (.45) unusual unique structures in his (.75) designs.
C7: This architect always (.45) develops and <ah> (.8) design (.4) unusual structures.
EL1: This archi-check always design and (1.2) in- architect.
EL2: This architect always design an unusual (2.1) creature.
EL3: This architecture designed (.8) developed and (1.7) <huh> (.75) innovative (1.75) structure.

A final example is drawn from binomials that were not specifically taught in class: processes and outcomes versus goals and outcomes. In the first sentence, participants heard the binomial in its reversed order, in the second, in its canonical order. In this instance, only one ELL managed to retain any portion of the reversed binomial (the word outcomes), and none were able to repeat the entire sentence in any form. ELLs were considerably more successful retaining the binomial in its canonical order despite the fact that this was a new expression that had not been explicitly practiced. Their ability to repeat (or not) these two sentences reflects that of their peers, who also demonstrated difficulties with the reversed binomial (several rewordings, one omission, long disfluencies) compared to the canonical order.

5. The report outlined the outcomes and processes of the experiment. (R)

C1, C2, C4: Exact repetition, no disfluencies.
C3: The report (.4) outlined the outcomes and (.6) evidence (.25) of the experiment.
C5: The report outlined the outcome and processes of the experiment.
C6: The outline gav- (.95) did the outline and procedures of the processes.
C7: The report outlined the outcomes and processes of each experiment.
C8: The report outlined the (.9) processes of the experiment.
C9: The report outlined the outcomes and processes (1.15) of the experiment.
EL1: The report outline the <uh> ….

EL2: The (.25) research import the outline … <hhh>
EL3: The report (1.1) outcomes (.6) <uh> (2.05) <uh>

6. The report outlined the goals and outcomes of the experiment. (C)

C1, C3, C4, C5, C6, C7, C8: Exact repetition, no disfluencies.
C2: The report (.4) outlined the goals and outcomes of the experiment.
C9: The report outlined the goals and outcomes (.4) of the experiment.
EL1: The report outlined the goals (.95) and outline (.3) of the (.45) accident.
EL2: The report (1.6) outlined (.55) goals and outcome (1.95) for an audience.

EL3: The report (.8) shows (.7) the outline of the (2.8) goals and outcomes (.95) <mhhm> …

Discussion

In response to our first research question, we note that results from the preference task corroborated previous work concerning learner vs. native speaker performance on lexical bundles and binomial word order (Mollin, 2014; Morgan & Levy, 2016, Siyanova-Chanturia, 2017). Pretest preference task results for the ELLs confirmed that at the start of their term learners were not familiar with these binomials, and appeared to simply guess at a response, with a mean for most bands averaging somewhere near the neutral response (4). By contrast, the native speaker controls mirrored patterns found in the COCA corpus, but did so conservatively. Rather than choosing "best" (or Mollin's "only possible"), even for highly fixed/frozen binomials, they leaned toward "highly" or "somewhat preferred." Previous research has suggested that ELLs may need a specific intervention to help them notice lexical bundles in the input, understand them *as bundles*, and grasp the constraints that will help them generalize them (AlHassan & Wood, 2015; Arnon & Christiansen, 2017; Christiansen & Arnon, 2017; El Dakhs et al, 2017; Hoang & Boers, 2016; Wood, 2009). The intervention that we provided, focusing on ten keywords and 20 binomials, was intended to do exactly that, and results indicate that while ELL intuitions did not mirror COCA corpus data on the pretest, they moved considerably closer on the post-test. Significant differences are observed between ELLs pre- and post-test preference task scores for the targeted binomials, specifically, as well as for binomial in the 74-76 fixedness band, with near significant differences observed in the 91-93 band. All other bands display movement toward the canonical order, but to different degrees, suggesting that learners have also recognized that some binomials are more highly fixed than others. Taken together, this indicates that information presented in class was not overgeneralized; learners distinguished binomials with variable word order from those with fixed orders in a way that was not observed on the pretest. For an intervention to be effective, learners must be both willing and able to apply information learned beyond the specific context of use. They must not, however, overextend the rules to cases where they do not apply. Results from the preference task suggest that by highlighting three specific constraints and focusing on them repeatedly and in context, using examples with specific binomials that were highlighted in class and further reinforced with rhetorical information and practice, learners were able to develop a system that served them well. They were able to apply specific item information to exemplars that they had learned in class, but also extend rules concerning constraints to new binomials as they encountered them. The fact that these constraints were not applied haphazardly signals a developing interlanguage system that allows for nuances in a complex system. In response to research question two, our results therefore suggest that the learner intuitions have become more native-like following the intervention, and not only for targeted binomials, but also for new binomials.

This evidence of generalization was reinforced in the repetition task, where learners also had greater success repeating binomials in their canonical order, not only for those binomials that were

specifically taught in class, but also for new binomials. Together, the two tasks suggest that the ELLs in the study were adding both individual binomials and the constraints that govern them to their interlanguage. In response to research question three, we can therefore argue that the eight-week intervention allowed learners to add both individual item knowledge and rule-based knowledge to their interlanguage system, allowing them to deal quite efficiently with both known and unknown binomials. Our study corroborates previous studies (Morgan and Levy, 2016) that suggest that speakers work from a combination of bundles entrenched in memory and generalized constraints. For those bundles that were entrenched in memory, both ELLs and native speakers produced the canonical form most fluently, whether this was the form that they heard, or whether they spontaneously reordered binomials that they heard in reverse order. For bundles that may have been less well known, fluency of repetition was again affected by order of presentation. Reorderings, rewordings and omissions were far more frequent with reversed binomials than with binomials presented in their canonical order. Morgan and Levy (2016) posit a "trade-off between reliance on abstract knowledge for infrequent items and reliance on direct experience for frequent items" (p. 395). Native-speaking participants experienced more difficulty repeating binomials that were somewhat less fixed (80%) and also those that were lower in frequency. ELLs, in contrast, heard a combination of known and new binomials throughout the task. Setting frequency (Mollin, 2014), rather than overall frequency, may therefore have played a role in ELLs' ability to repeat the binomials that they heard. For both groups, one can argue that a combination of direct experience and abstract knowledge of constraints was in evidence, as speakers relied on both during the task. For the ELLs, targeted binomials in canonical order were most easily repeated, followed by targeted reversed binomials. New reversed binomials provided many indications of memory decay, and to a much greater extent than new canonical binomials. It is this last category of binomials that allows us to argue that generalization of constraints did indeed occur, as their production, while less fluent than that of targeted binomials, was more fluent than that of new reversed binomials. Had the intervention included only specific binomials, with no explanation of the constraints that govern them, we are not as confident that these same results would have been obtained, as previous studies (Laufer & Girsai, 2008; Tajeddin & Daraee, 2013) have likewise shown the importance of form-focused instruction for MWEs.

Pedagogical implications

Binomial expressions are a significant component of both spoken and written texts of a variety of registers and genres. They also represent an area of vocabulary study previously underappreciated by the ESL community. This study suggests that attention to the form and underlying constraints governing their structure can enhance learners' knowledge of these expressions.

As the cognitive load for vocabulary study can be heavy at any level, in order for learners to successfully take up these expressions or find the study of such expressions beneficial, the presentation of binomials should highlight their utility, not require extra vocabulary practice, and should be linked to existing target vocabulary items. Learners are motivated when they know that the effort they put into learning will have multiple applications. In line with curricular recommendations from AlHassan & Wood (2015), the results of the present study suggest reading and writing activities that encourage noticing and elevate the saliency of target expressions. Further, our study suggests adding explicit teaching of rules to help students more easily process novel binomial expressions. Similarly, explicit testing could also encourage recognition. Participants in our study were not quizzed on the target binomials. Yet they still demonstrated an increased knowledge of binomial expressions after instruction. It is therefore conceivable that formative assessment of the target expressions, particularly *repeated* formative assessment could further promote learning of binomials.

Essential components of instruction on these types of expressions seem to be the relevance of the target binomials to other course content and presentation of rules that learners can base extended learning on. Repetition is another key feature of instruction that may promote automatization. Binomials should be highlighted when they appear in course materials. Instructors must be cognizant of teachable moments and opportunities to introduce binomials and present them as extensions of other vocabulary being encountered and taught. In addition to classroom implications, results of this study suggest that designers of instructional materials, textbooks and supplements alike, extend their treatment of multi-word expressions to include more binomial expressions. These presentations should include information on meaning, structure, use, and constraints that determine order.

Incorporation of binomial expressions and the constraints that govern them in our English language classes can provide language learners at various levels creative additions to their range of expression, thus contributing to their goals of fluency and natural-sounding language. Inclusion of these expressions should be encouraged at all levels of instruction.

Conclusions

Results of this study suggest that focused interventions similar to the one conducted based on instruction of academic vocabulary and related multiword expressions — in this case, binomials — provide great potential for expanding ELLs' knowledge of composite units of speech. While this study is limited in scope by the small number of learners that were involved, the inclusion of both known and new binomials, tested across a combination of intuition and production data allow us to show learners' ability to generalize constraints that were taught, beyond a mere knowledge of individual exemplars. This is a promising avenue for future classroom-based research. Despite these gains, however, we cannot argue that learners have begun to incorporate these academic binomials in their writing after eight weeks of instruction. Samples of learner writing collected throughout the term and examined for evidence of binomial use did not show spontaneous use of binomials by the learners in the study. As receptive abilities often precede production, it was to be expected that uncued production of binomials would not emerge in writing. In a similar pre-test/post-test study, Peters and Pauwels (2015) found that an increase in production of formulaic sequences was observable two terms after the original instruction. Therefore, further valuable extensions to our study would include an examination of writing samples later in the learners' path of study.

About the Authors

Theresa A. Antes is an Associate Professor of French and Linguistics at the University of Florida. She holds a PhD in French Applied Linguistics from Cornell University. Her research interests include second language acquisition and pedagogy, particularly as it relates to vocabulary acquisition. She has recently published articles in *Foreign Language Annals*, *Language and Sociocultural Theory, System* and *TESL-EJ*.

Errin Beck is a Senior Instructor II in the Intensive English Language Program at Portland State University. She holds an MA in TESOL from Portland State University and specializes in curriculum design and classroom assessment in language teaching. Her professional and research interests also include advanced level reading and writing instruction.

Acknowledgments

The authors wish to thank Lori Barkley, the classroom instructor without whose participation this study could not have taken place, and Mark Antes for assistance with data entry.

References

AlHassan, L. & Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes, 17,* 51-62. http://dx.doi.org/10.1016/j.jeap.2015.02.001

Arnon, I. & Christiansen, M. (2017). The role of multiword building blocks in explaining L1-L2 differences. *Topics in Cognitive Science*. 9, 621-636. https://doi.org/10.1111/tops.12271

Benor, S., & Levy, R. (2006). The Chicken or the egg? A probabilistic analysis of English binomials. *Language, 82,* 233-278.

Bestgen, Y. & Granger, S. (2014). Quantifying the development of phraseological competence in L2 English writing: An automated approach. *Journal of Second Language Writing, 26*, 28-41. http://dx.doi.org/10.1016/j.jslw.2014.09.004

Biber, D., Conrad, S., & Cortes, V. (2004). *If you look at …*: Lexical Bundles in University Teaching and Textbooks, *Applied Linguistics*, *25*, 371–405. https://doi.org/10.1093/applin/25.3.371

Boers, F. & Lindstromberg, S. (2008). Structural elaboration by the sound (and feel) of it. In F. Boers & S. Lindstromberg, (Eds.), *Cognitive Linguistic Approaches to Teaching Vocabulary*, (pp. 329-353), Mouton de Gruyter. http://dx.doi.org/10.1515/9783110199161.3.329

Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, *5*, 31-64.

Christiansen, M. & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science, 9, 542-551.* https://doi.org/10.1111/tops.12274

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213-238. http://dx.doi.org/10.2307/3587951

Coxhead, A & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing, 16*, 129-147. http://dx.doi.org/10.1016/j.jslw.2007.07.002

Cutler, S. (2017). The use of psycholinguistic formulaic language in the speech of higher-level Japanese speakers of English. *Vocabulary Learning and Instruction, 6*, 48-60. http://dx.doi.org/10.7820/vli.v06.1.2187-2759

Davies, Mark. (2008). *The Corpus of Contemporary American English: 450 million words, 1990-present*. Retrieved from http://corpus.byu.edu/coca/

El-Dakhs, D., Tasneem Prue, T. & Ijaz, A. (2017). The effect of the explicit instruction of formulaic sequences in pre-writing vocabulary activities on foreign language writing. *International Journal of Applied Linguistics and English Literature, 6,* 21-31. http://dx.doi.org/10.7575/aiac.ijalel.v.6n.4p.21

Ellis, N., & Simpson-Vlach, R. (2009). Formulaic language in native speakers: Triangulating psycholinguistics, corpus linguistics, and education. *Corpus Linguistics and Linguistic Theory, 5*, 61-78. http://dx.doi:10.1515/CLLT.2009.003

Ellis, Nick C., Simpson-Vlach, Rita, & Maynard, Carson. (2008). Formulaic Language in Native and Second Language Speakers: Psycholinguistics, Corpus Linguistics, and TESOL. *TESOL Quarterly, 42*, 375-396. http://dx.doi.org/10.1002/j.1545-7249.2008.tb00137.x

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning, 51*, 1-46. https://doi.org/10.1111/j.1467-1770.2001.tb00013.x

Hoang, H. & Boers, F. (2016). Re-telling a story in a second language: How well do adult learners mine an input text for multiword expressions? *Studies in Second Language Learning and Teaching, 6,* 513-535. http://dx.doi.org/10.14746/ssllt.2016.6.3.7

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge.

Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making*. Niemeyer. http://dx.doi.org/10.1515/9783110937923

Hyland, K. (2012). Bundles in academic discourse. *Annual Review of Applied Linguistics, 32,* 150-169.

Jones, M., & Haywood, S. (2004). Facilitating the Acquisition of Formulaic Sequences: An Exploratory Study in an EAP Context. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 269-300). John Benjamins. http://dx.doi.org/10.1075/lllt.9.14jon

Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, *29*, 694-716. https://doi.org/10.1093/applin/amn018

Li, J., & Schmitt, N. (2009). The Acquisition of Lexical Phrases in Academic Writing: A Longitudinal Case Study. *Journal of Second Language Writing, 18*, 85-102. http://dx.doi.org/10.1016/j.jslw.2009.02.001

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32, 111-129.* http://dx.doi.org/10.1017/S0267190512000128

Mollin, S. (2014). *The (ir)reversibility of English binomials*. John Benjamins. http://dx.doi.org/10.1075/scl.64

Morgan, E., & Levy, R. (2016). Abstract knowledge versus direct experience in processing of binomial expressions. *Cognition, 157,* 384-02. http://dx.doi.org/10.1016/j.cognition.2016.09.011

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524759

Peters, E. & Pauwels, P. (2015). Learning academic formulaic sequences. *Journal of English for Academic Purposes, 20,* 28-39. http://dx.doi.org/10.1016/j.jeap.2015.04.002

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158. https://doi.org/10.1093/applin/11.2.129

Schmitt, N., Jiang, X. & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95,* 26-43. http://dx.doi.org/10.1111/j.1540-4781.2011.01146.x

Singleton, D. (1999). *Exploring the second language mental lexicon*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139524636

Siyanova-Chanturia, A. (2017). Researching the teaching and learning of multi-word expressions. *Language Teaching Research, 21*, 289-297. https://doi.org/10.1177/1362168817706842

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. (2011). Seeing a phrase 'time and again' matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37,* 776-84.

Tajeddin, Z. & Daraee, D. (2013) Vocabulary acquisition through written input: Effects of form-focused, message-oriented, and comprehension tasks. *TESL-EJ, 16,* 1-19.

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 /Revue Canadienne De Linguistique Appliquée, 12*, 39-57.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511519772

Appendix A – PowerPoint Presentation

Title slide: PAIRED EXPRESSIONS: A way to expand your vocabulary

Slide 1: Importance of vocabulary development
Increasing your vocabulary can help you improve in writing, reading, listening, and speaking. Vocabulary is a fundamental basis of all of the skills of language. How do you know what to learn? You can use lists like the Academic Word List (AWL). You can also learn collocations…words that often come together.

Slide 2: What are they?
Words that are often grouped together or naturally occur with one another.
Examples
To MAKE a decision (not 'to have a decision')
Fast food (not 'quick food')
Why?
Learn and use these expressions to make your speech and writing sound more natural. Read and listen more fluently.

Slide 3: One particular type of collocation…
A short phrase that has two words of the same part of speech.
– Adjective and adjective
– Verb and verb
– Noun and noun
Examples
– Black and white (to describe photos or movies)
– Come and go (to describe a repetitive movement)
– Eyes and ears (to describe a person who substitutes for another)

Slide 4: What is special about these expressions?
They are common in native speaker speech and writing (at all levels and situations, academic, technical, general)
Some of them ONLY appear in one order
Some of them are flexible in order

Slide 5: How is the order decided?
There are some rules that decide the order of these paired expressions
– Sound
Short words come before long ones
Rules and regulations
– Meaning
Time – some activities and events naturally come before others
Yesterday and today
Relationship to speaker – things that are close to the speaker come before those that are farther away
Here and there
Good ideas come before bad ones
Positive and negative

[Back to article]

Appendix B – Sample Worksheet

Vocabulary Partners for Academics

In academics, we often find pairs of words that collocate together. These words appear in the same order every time they occur. In our class in the next few weeks, we'll look at the meaning of some of these pairs and the ideas they are related to from our book.

Here is the list of words for this week:

fair and accurate	a balanced account of ideas presented without bias and based on facts
identify and describe	to be able to locate a piece of information and tell about it or its connection to other ideas
recognize and appreciate	understanding facts, ideas or actions, and find value in those facts, ideas, or actions
large and complex	a big range of ideas with intricate connections
findings and conclusions	information that comes from a study and a statement of what the data reveals

Where we might see these phrases in our reading

- *News reporting, laws, and employment practices are all expected to be 'fair and accurate.'*
- *Students are often asked to 'identify and describe' ideas on a test, in a study or from their own experience.*
- *People often 'recognize and appreciate' things like art, literature, achievements, ideas…*
- *When describing structural relationships, one might describe something as 'large and complex.' For example, a collection of data or set of rules may be described as large and complex.*
- *We often read about 'findings and conclusions' at the end section of a study, analysis, or report. It is one of the key sections of reading you want to examine to learn what happened.*

In the exercises below, select words from the list to complete the sentences. Do not look back at your vocabulary list.

Accurate	*describe*	*large*
appreciate	*fair*	*recognize*
complex	*findings*	
conclusions	*identify*	

1. At the lower levels of Bloom's taxonomy, we're often asked to _____ and _____ information in order to show that we understand the main ideas of a reading.
2. Other Wes' brother Tony can _____ and _____ his influence over his younger brother. He tries to use this power to get Wes to listen to him and stay away from a street-life. Unfortunately, the power of money, drugs and notoriety Wes can gain from dealing drugs is stronger than Tony's influence.
3. Despite the fact that drugs are a street business, the drug trade in Baltimore was just as _____ and _____ of an organization as any major company in business.
4. Author Wes interviewed Other Wes' family, neighbors, and friends to learn more about why his life had turned out so differently. His _____ and _____ about the differences between the two men were shockingly simple. There was very little that separated the two men from their very different destinies.

5. In TOWM, the author says his purpose is not to excuse other Wes' crime. Rather, to show a _____ and _____ story where simple choices can make all the difference.

TIPS for remembering the order of Vocabulary Partners

If you can remember the 'constraints' of the pair ordering, you can get it right every time!

* A constraint is a rule the pair follows, a way the data is organized. Below are some common constraints for the order of words in pairs.

A. Metrical constraints – *order is determined by the length of the word. Always the short word will come before the long one*
Ex. Hot and cold
Ex. Fat and thin
Ex. Tall and short

B. Chronological constraints – *order is based on which idea or action would come first in a sequence*
Ex. Four or five ideas
Ex. Show and tell
Ex. Read and revise

C. Cultural centrality – *order is based on telling the smallest group or idea and expanding outward to include a bigger group*
Ex. The use of devices and technology
Ex. Players and teams
Ex. Ideas and information

Can you identify which rule is being used with our vocabulary?

- ___ fair and accurate
- ___ identify and describe
- ___ recognize and appreciate
- ___ large and complex
- ___ findings and conclusions

[Back to article]

Appendix C – Binomials preference task

Create and maintain
Open and available
Applications and concepts
Charts and graphs
Documents and interviews
Guidelines and standards
Vast and varied
Commitment and love
Evaluate and identify
Complex and large
Build and design
Principles and practices
Produced and distributed
Broad and diverse
Current and emerging
Evolution and origin
Fans and media
Use and create
Dental and medical
Form and function
Civil and military
Emerging and new
Processes and outcomes
Producers and consumers
Long and complex
Consequences and costs
Laws and procedures
Class and status
Medical and mental
Conclusions and results
Friendship and cooperation
Instruction and assessment
Distributed and printed
Evaluation and review
Liberal and conservative
Policies and programs
Status and wealth
Summary and conclusion
Tradition and history
Consumers and businesses
Federal and local
Liberal and democratic
Design and develop
Outcomes and goals
Summary and discussion
Accurate and fair
Location and name

Civil and religious
Injuries and illnesses
Protests and demonstrations
Provincial and federal
Fees and interest
Recognize and appreciate
Accurate and clear
Teachers and community
Photographs and documents
Date and location
Strategies and policies
Schools and community
Recovery and addiction
Rules and procedures
Protect and promote
Free and available
Findings and conclusions
Large and diverse
Development and evaluation
Rich and varied
Innovation and tradition
Causes and consequences
Collective and individual
Nature and function
Communal and individual
Sustain and create
Treatment and recovery
Maps and charts
Taxes and fees
Identify and describe
Formation and evolution
Develop and sustain
Strikes and demonstrations
Work and commitment
Instructors and students
Books and documents
Appreciate and understand
Federal and private
Promote and support
Words and concepts
Government and media
Principles and techniques
Facts and circumstances
Rules and guidelines
Trust and cooperation
Time and circumstances
[Back to article]

Appendix D: Repetition Task sentences, in order heard by participants.

Practice sentences:

1. This is my first practice sentence; I'm going to repeat exactly.
2. If I don't know a word, I'll repeat what I think I hear.
3. I'm going to speak loudly and clearly so the researchers can hear me.
4. I don't have any questions and I'm ready to start the task.

Task sentences:

1 The student asked to take a trip with her parents during the holidays.
2 This researcher combined her findings and conclusions in one section.
3 Teenagers sometimes get into trouble because they make bad decisions.
4 This author writes about national and global politics in her books.
5 If you're too hot, unbutton your jacket and take your hat off.
6 Lawyers must learn all procedures and rules before going to trial.
7 This city has a large and complex network of public transportation.
8 I asked him to speak louder because I was having a hard time hearing him.
9 The newspaper reported the circumstances and time of the accident.
10 Lawmakers in Washington often argue about which laws to pass.
11 These scientists use new and emerging technologies in the lab.
12 My employer provides mental and medical health insurance for me.
13 My presentation would have been better, but I ran into problems.
14 Profession often affects status and class in modern cultures.
15 My neighbor promised to keep an eye on my dog while I was out of town.
16 The speaker stated his conclusions and summary very clearly.
17 People sometimes decide on easy answers instead of hard choices.
18 The report outlined the goals and outcomes of the experiment.
19 Children can learn to appreciate and recognize differences at school.
20 Our boss is going to do us all a favor and reorganize the office.
21 Cities invest in community and teachers for the sake of their children.
22 The mother was tired, so she decided to sleep in and not fix breakfast.
23 Civil engineers develop and sustain urban renewal projects.
24 It's hard to come back to the same place when you've had a problem.
25 Employers always try to attract and retain the best employees.
26 Experts assure the evaluation and development of all products.
27 The homework was hard, so it helped to know that we were all on the same boat.
28 Businesses have to maintain and create appealing websites non-stop.
29 The professor was running late so his assistant started class.
30 News reports must be accurate and fair to avoid bias.
31 Teachers should identify and evaluate problems for their students.
32 In good marriages, each spouse makes up for the other one's faults.
33 This architect always develops and designs unusual structures.
34 This comedian went all out to entertain his audience.
35 Many lawyers fight for communal and individual rights for all.
36 My neighbor promised to keep up with my dog and me while we went for a jog.
37 Children can learn to understand and appreciate differences at school.
38 The homework was hard, but it helped to know that we were all in the same boat.
39 This researcher combined her conclusions and results in one section.
40 I asked if I could stay longer because I was having a good time with them.

41 The newspaper reported the facts and circumstances of the accident.
42 Teenagers sometimes get themselves out of jams by making good decisions.
43 This city has a complex and vast network of public transportation.
44 People sometimes settle on easy answers instead of hard choices.
45 Profession often affects wealth and status in modern cultures.
46 The mother was tired, so she decided to sleep late and not fix breakfast.
47 Lawyers must learn all laws and procedures before going to trial.
48 Civil engineers sustain and create urban renewal projects.
49 The student asked to take the test early so she could leave for the holidays.
50 Experts assure the review and evaluation of all products.
51 This comedian went out with a member of his audience.
52 This author writes about global and domestic politics in her books.
53 Businesses have to create and use appealing websites non-stop.
54 It's hard to come up with a solution for every problem.
55 My employer provides medical and dental health insurance free of charge.
56 Our boss is going to do away with these books and reorganize the office.
57 Cities invest in schools and community for the sake of their children.
58 These scientists use emerging and current technologies in the lab.
59 I could barely hear her when she spoke; she had a very soft voice.
60 Employers always try to retain and hire the best employees.
61 My presentation would have been better, but I ran out of time.
62 News reports must be clear and accurate to avoid bias.
63 In good marriages, each spouse makes an effort to accept the other's faults.
64 This architect always designs and builds unusual structures.
65 If you're too cold, button up your jacket and put your hat on.
66 The report outlined the outcomes and processes of the experiment.
67 The professor was feeling ill so his assistant started class.
68 Many lawyers fight for individual and collective rights for all.
69 Lawmakers in Washington often argue over which laws to pass.
70 Teachers should describe and identify problems for their students.
71 The speaker stated his summary and discussion very clearly.
72 I could hardly hear her when she spoke; she had a very soft voice.

[Back to article]