

Frequent Adjective + Noun Collocations for Intermediate English Language Learners

Joshua Brook Antle
Hokkaido University of
Education

Reference Data:

Antle, J. B. (2014). Frequent adjective + noun collocations for intermediate English language learners. In N. Sonda & A. Krause (Eds.), *JALT2013 Conference Proceedings*. Tokyo: JALT.

In this paper, two published collocation lists are examined, Coxhead's (2000) Academic Word List (AWL) and Shin and Nation's (2008) Collocation List for Spoken English. A new list of 150 frequent adjective + noun collocations is then proposed. These collocations were selected because the meaning of the noun within the collocation varies depending upon the adjective with which it is used. Due to the high frequency of these collocations, they are likely to be particularly useful for intermediate-level English language learners. Additionally, due to the variable nature of the noun component's meaning, students are unlikely to acquire these structures through exposure alone. Suggestions for EFL classroom application of this list are then made.

本論では、二つの発表された連語リスト (Coxhead [2000] の学術的ワードリスト [AWL] およびShin and Nation [2008] の口語英語の連語リスト) を考察し、そして新しい150の頻度の多い形容詞+名詞の連語リストを提案する。これらの連語は、連語内の名詞の意味と一緒に使われる形容詞によって変わるので、とりわけ選ばれたものである。これらの連語は頻度が高いので、特に中級の英語学習者にとって有益であろう。さらに、名詞部分の意味が変わるので、学生がそれらに触れることのみでそれらの構造を獲得することはありそうもない。そこで、このリストをEFLクラスで利用することが示唆される。

COLLOCATION LISTS, and word lists in general, are useful for explicit instruction in language teaching. By creating these lists, teachers can partially compensate for the learners' lack of exposure to the target language. Nesselhauf (2003) stated that "it seems indispensable that a number of collocations be taught and learnt explicitly" (p. 238). Building on this point, Durrant and Schmitt (2010) claimed that an "explicit focus on target collocations would dramatically improve their acquisition" (p. 181). However, collocation lists cannot be randomly compiled. These lists are most useful when they target language items that are suitable for the learners' needs, occur frequently in natural English, and are at an appropriate level.

The target language included in the collocation lists that have been compiled is determined by the researchers' view of what a collocation represents. The term *collocation* has been defined in different ways. A common definition is strictly frequency-based and is commonly used by corpus linguists. For example, Durant (2009) stated that "collocations are sets of two or more words which appear together more frequently than their individual frequencies would lead us to expect" (p. 158). Another set of definitions add other qualities to frequency. Antle (2012) defined collocations as "a set of two or more words that frequently occur together, that rep-



resent a single choice in a native speaker's mental lexicon, and whose meaning cannot be easily determined by the individual words themselves" (p. 26). Regardless of the definition used by researchers, collocation lists can be valuable tools for language instructors.

In this paper, I will briefly explain the problems collocations pose for L2 learners. This will be followed by a description and discussion of two collocation lists that were previously compiled. The final section will present my own list of adjective + noun collocations and some ideas for its classroom use.

Collocation Problems of L2 Learners

For the purposes of this study *collocation* and *formulaic sequence* will be used interchangeably following Wray (2009), who stated that features attributed to formulaic knowledge can be attributed to collocations as well. Pawley and Syder (1983) stated that native speakers use collocations to quickly produce utterances that both are grammatical and sound native. However, the literature shows that the nature of formulaic sequences causes learners problems when using a second language productively. Initial studies using small elicitation tests such as cloze and translation tasks found that collocation was highly problematic for L2 learners (Webb & Kagimoto, 2009) and that it accounted for a significantly high proportion of learner errors in L2 writing (Bahns, 1993). Wray (2002) pointed out that learners, who are in the process of learning a language, tend to be more analytical than native speakers. This results in learners focusing on isolated aspects of the language as opposed to holistic aspects. A reason for the large number of collocation errors is that learners often rely on intuition to determine which lexical items collocate with one another (Chi, Wong, & Wong, 1994). In addition, Fan (2009) believed a language learner's L1 adversely affects collocation production.

Ying and O'Neill (2009) described three problems learners have due to a lack of collocational competence:

1. Use of longer phrases and utterances because of an inability to express themselves concisely;
2. Odd word combinations—often a result of L1 influence;
3. Overuse of a few general items, leading to an oversimplified, flat, uninteresting style.

Shih (2000) investigated this overuse phenomenon and found that collocations of high frequency in learner English tend to be used to express vague ideas when more specific meanings should be conveyed. Shih also found that learners are apt to apply those collocations to cases in which more concise expressions are preferred. She presented two examples taken from the Taiwanese Learner Corpus of English to illustrate these phenomena:

1. "... but I have a big, big, big problem, that is, that I don't have a camera However, a camera is really very expensive."
2. "... it will be a big trouble to move all my things to another place." (Shih, 2000, p. 286)

Shih explained that in (1), "big problem" is used to convey the meaning of "financial problem," and in (2), "a big trouble" can be replaced by "troublesome."

Fan (2009) argued that collocation problems are prevalent regardless of the proficiency of the learner. Native speakers acquire collocation knowledge subconsciously and gradually through exposure, but L2 learners do not have this opportunity. Although collocations are particularly important for learners who are trying to achieve a high level of proficiency, they are also important for learners with more modest goals, as they lead

to more fluent and accurate English. However, it is unclear how and which collocations should be taught, especially considering the large number of collocations (Nesselhauf, 2003).

Previous Collocation Lists

Durrant (2009) and Shin and Nation (2008) used these types of definitions as the basis for creating collocation lists. Durrant created a list for English for Academic Purposes (EAP) and Shin and Nation focused on the highest frequency collocations in spoken English.

Durrant's EAP collocation list

Durrant's study (2009) produced a list of 1,000 two-word collocations. Durrant's aim was to produce a list of highly frequent EAP collocations that could be used as a pedagogical tool. It is intended to be a pedagogically manageable body of learning to which learners should pay special attention.

The corpus used to compile this list was created by collecting 5 million words from five different faculties (life sciences, science and engineering, social-psychological, social-administrative, and arts and humanities). Durrant used a four-word span, which means that there had to be a co-occurrence of the individual lexical items within four words to qualify. For example, both *there were no significant differences* and the *differences were not significant* would have been identified. He characterized academic collocations as those pairs that appear significantly more frequently in academic than in nonacademic texts. This disparity was calculated by comparing the total frequency of collocations in the academic corpus with their frequency in an 85 million-word subsection of the British National Corpus (BNC), comprising only nonacademic texts. Each collocation had to appear at least once per million words in each of the five parts of the corpus. Collocations were removed if (a) they

included an acronym or abbreviation, a proper name, an article, or a number or ordinal other than *one* or *first*; (b) the collocation corresponded to a single Latin item (e.g., *ad hoc*, *percent*); or (c) the majority of their occurrences appeared to be in writing outside the main text of the articles, for example, in bibliographies, copyright information, or acknowledgements.

The resulting list had 763 collocations that were *grammatical*, meaning that one of the words was nonlexical, such as a preposition, a determiner, or a modal verb (e.g., *related to*). Durrant (2009) justified including these collocations by stating, "One benefit to learners of a listing of high-frequency grammatical collocations is that the most typical versions of the patterns they need, and the most typical patterns of the words they need, can be brought to their attention" (p. 163). However, Woolard (2000) believed it was better to restrict the use of the term collocation to relationships between nouns, verbs, adjectives, and adverbs.

One interesting finding was a lack of overlap between this collocation list and the items on Coxhead's (2000) Academic Word List (AWL). Of the 1,000 collocations on Durrant's collocation list, only 425 include an item from the AWL. Durrant argued that this lack of overlap indicates a shortcoming of traditional approaches to identifying academic vocabulary, rather than a weakness of his list. The AWL excludes items that are on West's (1953) General Service List (GSL) on the grounds that students of EAP are likely already to have mastered these items. Durrant believed the strategy of eliminating all high-frequency words from academic word lists seems suspect: Many items that are excluded by this strategy may be of considerable importance for learners of EAP.

Durrant conceded two weaknesses of his collocation list. First, he acknowledged that by limiting his search to two-word collocations, he was likely missing many valuable collocations of three or more words. Secondly, this analysis looked only at the forms, as opposed to the functions of the collocations. Therefore,

although the collocations appeared at least five times in each part of the corpus, it is not clear that all disciplines use them in the same way. However, analysis of the use of collocations would need to be undertaken manually and would be a labour-intensive task.

Shin and Nation's Collocation List for Spoken English

The Shin and Nation (2008) study determined the most frequent collocations in spoken English based on the spoken section of the BNC. In Shin and Nation's study, collocation referred to a group of two or more words occurring together with each collocation having two parts: a pivot word and its collocate(s). The pivot word is the focal point of the collocation. For example, in the collocations *high school*, *high court*, and *too high*, *high* is the pivot word and *school*, *court*, and *too* are the collocates. Shin and Nation investigated the 1000 most frequent word types from the spoken section of the BNC as pivot words. They used six criteria to find collocations in the corpus:

- Each pivot word was a word type rather than a word family. Therefore, *books* and *book* were treated separately.
- Only nouns, verbs, adjectives and adverbs were considered as pivot words.
- Each pivot word had to be in the first 1000 words in the spoken word frequency list by Leech, Rayson, and Wilson (2001).
- Each collocation had to occur at least 30 times within 10 million running words in the BNC spoken corpus.
- Each collocation should not cross an immediate constituent boundary (e.g., *I saw you at that place* has five immediate collocational constituents: *I saw you at that place*, *saw you at that place*, *saw you*, *at that place*, and *that place*. *You at that place* does not meet this criterion).
- Different senses of collocations with the same words and

word forms were considered different (e.g., *looking up* can mean *to improve* or *to search*; these were counted separately) (Shin & Nation, 2008, pp. 342-343).

The final list contained 4,698 collocations indicating there are large numbers of grammatically well-formed high-frequency collocations. The list also showed that pivot words that are more frequent have a greater number of collocates. The first 100 pivot words have an average of 20.5 collocations, but the second 100 words have 8.4. Additionally, two-word collocations account for 77% of the total list.

This collocation list was originally designed for elementary learners of English. Considering the target learners, Shin and Nation (2008) conceded several weaknesses of their list from a pedagogical perspective. First, many collocations are strongly colloquial and may not be suitable for explicit instruction. Secondly, frequency is just one of several criteria that should be considered when deciding what to focus on, such as learner need, range of use (spoken and written), difficulty, teachability, and suitability for the age and background of the learners. Finally, greetings such as *good morning* and *how are you* do not appear within the 100 most frequent collocations, which indicates frequency alone should not dictate what language to target.

Although each of these two collocation lists represents one aspect of English language (spoken English and Academic English) and is suitable material for a certain level of learner (beginner students and EAP students), they cannot be appropriate for use in all ELT contexts. I therefore considered it necessary to create a new list to assist my specific groups of learners in their understanding and acquisition of collocations. Like the above lists, however, the one I compiled also focuses on a specific type of collocation (frequent adjective + noun collocations) and is suitable for a specific level of English language learner (intermediate).

A Proposed Adjective + Noun Collocation List

For this collocation list, I used Hill, Lewis, and Lewis's (2000) list of "problematic but really useful words: 47 nouns whose meanings depend on the adjectives used with them" as the pivot words (see Figure 1). This list was chosen because the nouns are frequently used in English, but their meanings are opaque. For example, the meaning of the noun *view* is different in the collocation *clear view* and in the collocation *general view*. Because of these two qualities, I feel the collocations are appropriate for intermediate English language learners. Furthermore, the collocation list was trialled with high beginner students using the activities presented at the end of this paper. Most of the students had success with the material, but I felt it was slightly above their current level.

For my collocation list (see Table 1), the BNC was used as the corpus with frequency being the criterion for inclusion. The goal

was to produce a pedagogical tool for intermediate learners of English. Selection of collocations from the BNC was made based on five criteria:

1. Nouns from the list of 47 *Really Useful Words* in Figure 1, that were not within the first 1000 most frequent words were not investigated (33 nouns met this cutoff).
2. All adjectives are within the first 1000 most frequent words except if the adjective + noun collocation occurred at least 200 times within the BNC.
3. All collocations must occur at least 50 times in the BNC.
4. Only the five most common adjectives for each noun were included.
5. However, all collocations that have 200 or more occurrences in the BNC are included.

47 problematic nouns (frequency rank on the General Service List)

account (607)	decision (615)	method (NoL)	question (230)	style (NoL)
action (297)	difference (461)	move (134)	reason (296)	system (183)
answer (412)	discussion (897v)	performance (694)	relationship (654)	theme (NoL)
approach (NoL)	effect (326)	plan (187)	result (229)	theory (NoL)
argument (688v)	feature (NoL)	policy (NoL)	scheme (NoL)	use (65)
behaviour (942)	idea (328)	position (383)	situation (458)	view (409)
change (198)	information (319v)	problem (180)	solution (NoL)	vision (NoL)
circumstance (NoL)	interest (161)	programme (179)	state (55)	way (84)
condition (331)	issue (NoL)	project (NoL)	story (523)	work (71)
consequences (NoL)	manner (757)			

Figure 1. Really useful words. Adapted from Hill, Lewis, and Lewis (2000, p. 101). NoL = not within the 1000 most frequent words on the GSL; ##v = in verb form on the GSL.

The first two criteria were chosen because intermediate students (TOEIC score 400-600) would likely be familiar with words in this range, which should aid in the acquisition of the collocations. Woolard (2000) stated that “learning more vocabulary is not just learning new words, it is often learning familiar words in new combinations” (p. 31). The final three criteria were chosen to ensure all 33 of the nouns were represented on the list without omitting any highly frequent collocations.

Table I. Frequent Adjective Plus Noun Collocations for Intermediate Learners

Adjectives	GSL	BNC	Useful nouns
current	720	370	account
detailed	606	108	
full	349	92	
legal	noL	463	action
further	235	172	
political	426	163	
social	247	158	
only	56	105	answer
right	133	94	
short	436	92	
simple	301	91	
strong	385	55	argument
main	698	51	
human	304	242	behaviour
social	247	190	
good	102	157	
bad	588	80	

Adjectives	GSL	BNC	Useful nouns
social	247	487	change
major	noL	309	
significant	noL	185	
		189	changes
		172	
political	426	146	change
sudden	571	118	
good	102	382	condition
economic	noL	222	
poor	807	92	
necessary	387	89	
human	304	85	
final	noL	317	
right	133	78	
political	426	76	
difficult	674	68	
conscious	815	65	
significant	noL	328	difference
		296	
only	56	202	
big	262	172	
important	284	102	
main	698	100	
further	494	179	discussion
public	166	76	
immediate	541	158	effect
good	102	128	
direct	280	124	
significant	noL	120	
great	89	82	

Adjectives	GSL	BNC	Useful nouns
good	102	1851	idea
whole	329	177	
clear	269	154	
new	57	132	
better	261	123	
further	494	1113	information
new	57	303	
other	45	201	
general	136	116	
private	540	100	
public	166	881	interest
particular	338	377	
high	108	339	
special	336	324	
great	89	322	
similar	noL	81	manner
high	108	269	performance
economic	noL	214	
poor	807	131	
good	102	117	
best	315	77	
local	357	278	plan
(5)-year	58	88	
new	57	83	
financial	noL	231	position
present	169	133	
new	57	131	
better	261	129	
strong	385	123	

Adjectives	GSL	BNC	Useful nouns
main	698	400	problem
major	noL	400	
		227	problems
real	146	287	
economic	noL	250	
only	56	248	
serious	644	216	
particular	338	213	
financial	noL	211	
new	57	100	programme
full	349	56	
important	284	165	question
whole	329	124	
difficult	674	118	
good	102	93	
open	165	83	
good	102	735	
main	698	379	
only	56	309	
other	45	275	
real	146	207	
close	203	236	relationship
special	336	215	
good	102	111	
new	57	80	
direct	280	66	
direct	280	228	result
good	102	64	
present	169	146	situation
political	426	141	
current	720	132	
particular	338	113	

Adjectives	GSL	BNC	Useful nouns
present	169	201	state
new	57	168	
current	720	164	
solid	951	97	
free	369	79	
whole	329	226	story
short	436	201	
different	327	155	
true	381	152	
long	92	120	
political	426	738	system
new	57	649	
legal	noL	333	
economic	noL	220	
educational	noL	217	
present	169	202	
good	102	261	use
effective	554	160	
full	349	147	
best	315	145	
better	261	112	
different	327	159	view
clear	269	117	
fair	666	114	
general	136	94	
good	102	86	

Adjectives	GSL	BNC	Useful nouns
long	92	1749	way
only	56	1748	
other	45	1452	
best	315	1166	
different	327	502	
wrong	769	392	
right	133	347	
good	102	343	
effective	554	317	
similar	noL	293	
new	57	260	
better	261	256	
easy	351	255	
particular	338	202	
social	247	1235	work
hard	477	1224	
good	102	306	
recent	375	240	
other	45	216	
new	57	209	

Note. GSL = frequency rank on the General Service List. BNC = number of occurrences for that collocation in the BNC. NoL = not within the 1000 most frequent words on the GSL.

Three collocations were removed from the list: *industrial action* (283-354), *political behaviour* (426-71), and *general discussion* (136-68). Despite meeting all five criteria, they were removed to bring the final number of collocations to 150. I judged these three collocations to be difficult to teach and less needed by intermediate students, relative to the other collocations on the list. The noun *move* did not have any adjective collocates that met the criteria.

As in the two published collocation lists described above, there are several weaknesses in this proposed list. First, frequen-

cy alone was used to determine inclusion. Other factors such as learner need, teachability, and level should also be considered when deciding how and when to use this list. Second, the original list of 47 nouns appears to have been compiled through intuition alone. Hill, Lewis, and Lewis (2000) did not specify how they determined that these nouns were problematic. Finally, these collocations were extracted from a corpus, which is a compilation of language and does not represent the language used by any individual person.

Discussion

Target vocabulary selection requires the instructor to identify categories of language that are suitable for explicit instruction. These categories of language likely consist of structures that are difficult to acquire through exposure alone. I believe frequent adjective + noun collocations consisting of nouns from the list of *Really Useful Words* have this quality. The nouns from the frequent adjective + noun collocations have meanings that are largely dependent on the adjectives with which they are used. For example, the meaning of the noun *way* has a different meaning in the collocation *best way* than in the collocation *long way*. It is logical to assume that a noun like *way* would be more difficult to acquire through exposure alone than nouns with a more concrete meaning. However, by focusing on the complete collocation, these structures become less problematic. I believe that exposing students to large quantities of comprehensible input is an effective approach to English language instruction, but this approach should be supplemented with explicit teaching of structures such as the collocations previously listed.

Using the Collocation List in Class

Although the focus of this paper is how the collocation list was compiled, it is helpful to consider how it can be used in EFL

classrooms. Possible procedures for explicitly teaching, and then practicing, these collocations are briefly described below. These activities were trialed with high beginner students. The students were engaged and had success, but the activities themselves were difficult to monitor.

Example Procedures for Using the Collocation List in Class

1. Create a translation and cloze exercise using 15 of the collocations from the list. Assign one list for homework each class, for eight consecutive classes (120 total collocations). I have created a website that can be used to initially expose and provide translations for 120 of the collocations at <https://sites.google.com/a/joshuaantle.com/adjektive-nouncollocations/>
2. The next class, spend the final 10 to 15 minutes doing one of the following speaking exercises targeting the previous lesson's collocations:
 - » Half a crossword: Create two crossword grids. On grid A, write half of the previous lesson's collocations as the across answers. On grid B, write the other half of the previous lesson's collocations as the down answers. Do not provide any clues on either crossword. The students work in pairs: One student has crossword grid A and the other student has B. One student asks for a hint by saying "What is number 2?" Their partner has to provide a hint or hints. A hint could be a Japanese translation, a cloze sentence, or a description of the targeted collocation. The students continue asking for and giving hints until both crossword grids are complete.
 - » Create a conversation: Each student works with a partner. Give each pair of students one of the previous lesson's collocations. The students work together and write a short two-person four-line conversation that includes

the targeted collocation. As the students are working, walk around the class and correct their English. The most common problem is that the students create an unnatural use for the targeted collocation. After most of the pairs of students have finished, they post their completed short conversations on the wall. The students then walk around the class with their partner reading out loud each other's conversations, providing further exposure to the target language.

- » Conversation cloze: This activity is similar to *Create a conversation*; however, in this activity the teacher writes all of the short two-person four-line conversations beforehand. The conversation is a cloze activity with the targeted collocation removed and written on the back of the card, so the students can check their answers. Student A reads the first line of the conversation. Student B reads the second line. This continues until the conversation is completed. When a student encounters the blank space, he or she has to fill it in with the correct collocation for the given situation. All of the collocations are from the previous lesson's list. The students walk around the classroom, going from one cloze task to the next, until all 15 conversations have been completed.

Conclusion

A collocation list can be a valuable tool for language instructors. It can help compensate for a lack of exposure to the target language. However, the criteria used to create such a list largely determine its value for a particular group of students. The three lists described represent pedagogical tools suitable for students at three different proficiency levels (Durrant's, 2009, list for EAP students; Shin and Nation's, 2008, list for elementary English language learners; and my list for intermediate English language learners). The activities described can be used to

review and repeatedly expose language learners to the targeted language. Given the problematic nature of collocations for second language learners, explicit instruction using lists such as these will help students overcome the productive challenges of collocations. However, more investigation into the pedagogical challenges of using collocation lists and explicitly teaching collocations is needed.

Bio Data

Joshua Brook Antle is currently a lecturer at Hokkaido University of Education. His research interests include collocations, fluency, and listening.

References

- Antle, J. B. (2012). Two words are better than one. In A. Stewart & N. Sonda (Eds.), *JALT2011 Conference Proceedings*. Tokyo: JALT.
- Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal*, 47, 56-63.
- Chi, M. L., Wong, P. Y., & Wong, C. P. (1994). Collocational problems amongst ESL learners: A corpus-based study. In L. Flowerdew & A. K. Tong (Eds.), *Entering text* (pp. 157-165). Hong Kong: University of Science and Technology.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-238.
- Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, 28, 157-169.
- Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research*, 26(2), 163-188.
- Fan, M. (2009). An exploratory study of collocation use by ESL students: A task-based approach. *System*, 37, 110-123.

- Hill, J., Lewis, M., & Lewis, M. (2000). Classroom strategies, activities and exercises. In M. Lewis (Ed.), *Teaching collocation: Further developments in the Lexical Approach* (pp. 88-117). Hove, UK: Language Teaching Publications.
- Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English: Based on the British National Corpus*. London: Longman
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24, 223-242.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). New York: Longman.
- Shih, R. H. H. (2000). Collocation deficiency in a learner corpus of English: From an overuse perspective. In A. Ikeya (Ed.), *The Proceedings of The 14th Pacific Asia Conference on Language, Information and Computation* (pp. 281-288).
- Shin, D., & Nation, I. S. P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal*, 62, 339-348.
- Webb, S., & Kagimoto, E. (2009). The effects of vocabulary learning on collocation and meaning. *TESOL Quarterly*, 43, 55-77.
- West, M. (1953). *A general service list of English words*. London: Longman.
- Woolard, G. (2000). Collocation: Encouraging learner independence. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 28-46). Hove, UK: Language Teaching Publications.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. (2009). Navigating L2 collocation research. In A. Barfield & H. Gyllstad (Eds.) *Researching collocations in another language* (pp. 232-244). New York: Palgrave Macmillan.
- Ying, Y., & O'Neill, M. (2009). Collocation learning through an "AWARE" approach: Learner perspectives and learning process. In A. Barfield & H. Gyllstad (Eds.) *Researching collocations in another language* (pp. 181-193). New York: Palgrave Macmillan.