

A Corpus-based Comparison of Near-Synonymous Adjectives in General English and in Academic Writing

Siaw-Fong Chung & Li-yin Chen, Taiwan

National Chengchi University

Abstract

In academic writing, certain near-synonymous adjectives are commonly found, among which are *critical* and *important*. This study aims to analyze the unique and shared patterns of these two near-synonyms in academic discourse and general use. In the British Academic Written English (BAWE) Corpus, both adjectives modify nouns such as *point*, *method*, and *data*. However, each has its individual uses: only *critical* modifies *lure*, *path*, and *period*, and only *important* modifies *element*, *thing*, and *implication*. However, the collocation patterns are somewhat different in a reference corpus. In the British National Corpus (BNC), *critical* modifies *scrutiny* and *evaluation*, while *important* modifies *implication* and *consequence*. The two adjectives behave differently in the two registers. By comparing the unique uses of a word in a specialized corpus, one can explore its regular patterns. Such patterns could be contrasted more clearly by comparing to a reference corpus. We argue that the collocation analysis from a reference corpus may need to be accompanied by a specialized corpus for the understanding of the vocabulary of a specialized field.

Keywords: academic writing, near-synonym, corpora, adjective

Introduction

Important is listed as one of the words in the 1000 General Service List (GSL), while *critical* is in the second 1000 GSL list (West, 1953). Because of this, the words are not listed in the Academic Word List (AWL) (Coxhead, 2 000), since GSL words

are automatically excluded from the AWL; both, however, are included in Paquot's (2010) Academic Keyword List (AKL). The AKL contains 355 nouns, 233 verbs, 180 adjectives, 87 adverbs, and 75 other words. Although it is not our intention to compare the word lists, it is worth pointing out that both *important* and *critical* are considered part of a general word list by some (e.g. West, 1953), but are included as academic keywords by Paquot. Yet, none of these word lists tells us how the usage of these two adjectives differs in general and in academic discourse. It is the aim of this work to compare them in the two genres using two corpora – the British National Corpus (BNC) and the British Academic Written English Corpus (BAWE). Our research questions are as follows:

- (a) How similar or different are the linguistic behaviors of *important* and *critical* in a reference corpus and a specialized corpus?
- (b) How will a corpus-based knowledge of the similarities and differences of the near-synonyms help raise awareness of them in ESP teaching and learning?

These two research questions will be answered based on our observations from the corpora.

As *important* and *critical* have at least one overlapping sense, they are qualified to be called near-synonyms (cf. Chung & Ahrens, 2008; Chung, 2011). The studies of near-synonyms are many and they mostly aim to disambiguate the senses of two closely related words. Our work includes not only sense disambiguation, but also the comparison of both *important* and *critical* in two corpora, focusing on their differences in meaning, as well as their different linguistic behaviors when they appear in different genres. As this pair of near-synonyms could appear in both general and academic discourse, they are hypothesized to possess different linguistic environments. Although there has been much research on English word lists and on near-synonyms, the study of near-synonyms in academic versus general discourse is less often seen, especially when examining a pair of near-synonyms that do not seemingly differ in most circumstances. *Important* and *critical* are less likely to evoke enough interest, for they are too general to display any significant differences. It is our purpose to figure out the differences between the two in the two different genres.

Literature Review

In the following, we first review previous research on adjectives in academic discourse, and then a brief review of academic vocabulary follows.

Adjectives in Academic Discourse

Adjectives are generally classified into two types, attributive adjectives and predicative adjectives. Biber, Johansson, Leech, Conrad, and Finegan (1999) analyzed the syntactic roles of the two types, showing that the attributive adjectives normally precede the nominal expression, and function to pre-modify nominal expressions as shown in (1) with the adjectives in bold.

(1) *One of the most **important** ways of achieving this is by the **regular** and **thorough** implementation of planned disinfection programmes in all livestock units.* (taken from Biber et al., 1999, p. 510)

As illustrated in this example, the attributive adjective *important* pre-modifies the head noun *ways*, and similarly, the second head noun, *implementation*, is pre-modified by two attributive adjectives, *regular* and *thorough*. Biber et al. further subdivided the attributive adjectives based on their semantic domains by forming two major classes, namely, descriptive and classifying. Descriptive adjectives include those describing size or amount (e.g. *great*, *low*), time (e.g. *new*, *young*), color (e.g. *black*, *dark*), and evaluation (e.g. *important*, *special*), whereas classifying adjectives refer to those which describe relations (e.g. *same*, *different*, *general*, and *final*), topics (e.g. *political*, *public*), and affiliations (e.g. *American*, *British*). In contrast to the other three registers that they established, namely, conversation, fiction, and news, the authors also found that academic prose attests far more classifying attributive adjectives.

The other class of adjectives, predicative adjectives, differs from attributive adjectives in terms of syntactic position, as shown bolded in (2).

(2) *The tendencies are not **significant** and get **weaker** when data are corrected for guessing.* (taken from Biber et al., 1999, p. 515)

Biber et al. found that predicative adjectives are commonly used following copular verbs including *be*, *become*, *get*, *look*, *feel*, *seem* and *appear*. The copula

be is the most prevalent verb used with predicative adjectives (over 20 times more frequent than any other), and it is the most frequent in academic prose. Some predicative adjectives (e.g. *sure, important, difficult, likely, necessary, and possible*) are particularly common in the academic genre since they are often used to express the author's stance, or the author's position in relation to the topic under discussion.

Investigation the spoken and written registers in the university, Biber (2006:13-14) had also discovered some meaningful observations on adjectives. In this work, Biber discovered that adjectives could vary in frequency in different registers. He pointed out that in the *Longman Grammar of Spoken and Written English* corpus, "there are about 300,000 nouns per million words in academic prose, compared to only around 150,000 per million words in conversation" (pg. 14, citing data from Biber et al. (1999: 235)). As for adjectives, they are found to be far more common in academic prose (around 80,000 per million words) than other registers (around 60,000 per million words in fiction and newspaper and around 20,000 per million words in conversation).

Previous work on uses of adjectives in academic discourse was also undertaken in studies of evaluative language. In this line of research, adjectives were found to express the author's opinion (e.g. Hunston & Thompson, 2000) or the authorial stance (e.g. Swales & Burke, 2003). For example, Swales and Burke (2003) analyzed evaluative adjectives across the spoken and written registers in academic discourse. The spoken register data is from the Michigan Corpus of Academic English (MICASE), and the written corpus consists of Hyland's (2000) corpus of 80 research articles collected from eight disciplines. The evaluative adjectives considered by Swales and Burke were gradable adjectives occupying various positions on a continuum, including polarized adjectives at the extreme ends (e.g. *marginal, crucial, essential, fundamental, trivial, irrelevant, key*) and centralized adjectives occupying the middle positions (e.g. *central, important, main, major, peripheral, relevant, serious, unimportant*). The distribution of the adjectives showed that the centralized adjectives are prevalent in both registers (around 85%). The result of a statistical analysis showed that a significant difference holds between the frequencies of polarized adjectives of the spoken (69 per 100,000 words) and written (111 per 100,000 words)

registers. A detailed analysis was also carried out by further dividing the adjectives into seven sub-categories as shown in Table 1.

Table 1 Swales and Burke's (2003) classification of evaluative adjectives

Categories	Examples
acuity	<i>smart, stupid</i>
aesthetic appeal	centralized: <i>pretty, unattractive</i> polarized: <i>beautiful, hideous</i>
assessment	centralized: <i>bad, boring, dull, exciting, fair, good, interesting, uninteresting</i> polarized: <i>amazing, awesome, awful, excellent, fascinating, horrible, incredible, terrible, unbelievable</i>
deviance	centralized: <i>funny, odd, standard, strange, typical, unusual</i> polarized: <i>absurd, crazy, weird</i>
relevance	centralized: <i>central, important, main, major, peripheral, relevant, serious, unimportant</i> polarized: <i>irrelevant, fundamental, marginal, crucial, essential, trivial, key</i>
size	centralized: <i>big, large, little, small</i> polarized: <i>enormous, huge, infinitesimal, minute, teeny, tiny, tremendous</i>
strength	<i>weak, strong</i>

A comparison of the rate of occurrence in the two academic registers, spoken and written, showed that the former is more frequent in five out of the seven categories, namely, 'acuity', 'aesthetic appeal', 'assessment', 'deviance', and 'size'. The only two categories where instances in the written register outnumber the spoken are 'relevance' and 'strength'. Under the category of 'relevance', centralized adjectives such as *important, major, and relevant* (6.5, 2.7, and 2.4 tokens per 100,000 words, respectively) are used much more frequently in the written register. The adjective *important* is also found to be the most prevalent in the spoken register (4.9 tokens per 100,000 words).

Another approach to evaluative adjectives is by means of patterns. In Hunston and Sinclair (2000), the authors proposed the consideration of “local grammars” or “a categorization and terminology that is developed specifically for each area” (Hunston & Thompson, 2000, p. 74). Unlike earlier work, Hunston and Sinclair identified patterns of evaluative adjectives as summarized in Table 2 and argued for the use of patterns to distinguish evaluative from non-evaluative adjectives.

Table 2 Hunston and Sinclair’s (2000) patterns of evaluative adjectives

Patterns	Examples
<i>IT</i> + LINK VERB + ADJECTIVE GROUP + CLAUSE	<i>It was certain that he was much to blame.</i>
<i>THERE</i> + LINK VERB + <i>SOMETHING/ ANYTHING/NOTHING</i> + ADJECTIVE GROUP + <i>ABOUT/IN</i> + NOUN GROUP/ <i>-ING</i> CLAUSE	<i>There's something rather appealing about being able to spend the evening in a town.</i>
LINK VERB + ADJECTIVE GROUP + <i>TO</i> -INFINITIVE CLAUSE	<i>Horses are pretty to look at.</i>
LINK VERB + ADJECTIVE GROUP + <i>THAT</i> CLAUSE	<i>He was very angry that she had spoken to people about their private affairs.</i>
PATTERNS WITH <u>GENERAL NOUNS</u>	<i>The important <u>point</u> is to involve them in the decision.</i>

With the exception of the last pattern (patterns with general nouns) in Table 2, all the other five follow a linking verb to occupy a predicative position. As mentioned by the authors, evaluation is “an extrinsic quality, being a matter of judgement” (p. 94), implying that evaluative adjectives are less likely to be attributive where they immediately precede a noun. Hunston and Sinclair’s work sheds light on how patterns may serve to distinguish the types of adjectives (i.e. evaluative from non-evaluative).

The last line of research is synonymy studies where near-synonymous adjectives are compared for their linguistic behavior. In an earlier study, Gries (2001) examined

the differences between *-ic* and *-ical* adjectives in English (e.g. *politic(al)*, *economic(al)*, *classic(al)*, *numeric(al)*, and *problematic(al)*). Gries conducted a corpus-based analysis to identify the R1 collocates (the first word to the right of the adjective) and common collocates shared by each pair aided by statistical tools (e.g. log-likelihood, chi-square). The results showed variations among each pair, although these pairs of adjectives had been traditionally considered to be synonymous. Some pairs (e.g. *logistic(al)* and *symmetric(al)*) exhibited divergent patterns based on the identification of non-overlapping or discriminating collocates. Gries' study provides us with a new method to measure the degree of similarity between synonymous words and dispels our misconception in conceiving all *-ic/ical* pairs as synonymous. In a more recent study on adjectives, Gries and Otani (2010) focused on the adjectives of size and investigated a set of synonyms (*big*, *great*, and *large*) and their antonyms (*little*, *small*, and *tiny*) with a more sophisticated method involving statistical analyses including hierarchical agglomerative cluster analysis, dendrograms, and snakeplots to measure the degrees of similarity among the words. Their results support their assertion that antonym pairs such as *big/little* and *large/small* are canonical antonyms. They also provided evidence to show that *tiny* behaves like *smallest* based on categories of the characteristics of collocating neighbors. The authors investigated these size adjectives with a wide range of categories including morphological features (tense, voice, and transitivity of the co-occurring finite verb), syntactic features including the syntactic position (attributive/predicative) and clause level (dependent/main clause), as well as semantic features such as whether the noun is modified (the original term by the authors was 'modifiee or whether the noun is countable, whether count/non-count, concrete/abstract, human/organization, quantity/ongoing processes/punctual events, whether the modifier is literal/metaphorical or quantitative/evaluative, and others. In other words, the annotation was multidimensional, considering various aspects of the co-occurring conditions of the target adjective set.

In a similar fashion, Liu (2010) examined five near-synonyms of attributive adjectives (*chief*, *main*, *major*, *primary*, and *principal*) and found that the internal semantic structure of this synonym set is quite intricate based on a behavioral

profile approach “for describing the distributional patterns of lexical items” (p. 59). Specifically, the nouns pre-modifying each of the adjectives were examined and classified into six categories including ‘abstract’ (e.g. *concern, reason*), ‘concrete’ (e.g. *road, dish*), ‘dual’ (e.g. *source, component*), ‘institution’ (e.g. *school, corporation*), ‘position-title’ (e.g. *executive, counsel*), and ‘non-position title’ (e.g. *sponsor, author*). Liu applied Gries’ (2004) hierarchical configural frequency analysis program (HCFA 3.2 for R) to his categories of pre-modified nouns. The results showed that all five adjectives pre-modify ‘abstract’ and ‘dual’ nouns, but *main* is specific for modifying ‘concrete’ nouns, *principal* is for ‘non-position title’ nouns, *chief* is for ‘position title’ nouns, and *major* and *primary* are for ‘institution’ nouns. Furthermore, a cross-register comparison among the spoken, fiction, newspaper, magazine and academic writing registers indicated that the five adjectives also differ in degree of formality when they pre-modify abstract/dual nouns. Liu concluded that *main* is the least formal as it is more common in the spoken register and *primary* is the most formal and commonly found in academic writing. *Major, chief, and principal* are considered to be of neutral formality. In a recent study of near-synonyms of six adjectives (*nice, kind, lovely, friendly, gorgeous, and pleasant*), Hoffmann (2014) investigated their nominal collocates across five registers (spoken, fiction, magazines, newspaper, and academic writing) and suggested that style variation appears to influence the behavior of the adjectives. For example, these six adjectives are rarely found in academic writing but vary in distribution in the other four less formal registers. Both Liu and Hoffmann’s work considered register variations, and their results demonstrated lexical specificity for various registers. In addition to this line of studies, research in academic vocabulary has also recognized the importance of disciplinary specificity, leading to the compilation of specialized corpora, to be discussed in the next section.

Academic Vocabulary

Previous work on near-synonyms appears to make comparison between different registers but few studies have contrasted a specialized corpus with a general one. However, researchers in academic vocabulary have long recognized the linguistic

variation of a specialized corpus from a reference corpus (e.g., Chung & Nation, 2003; Hyland and Tse, 2007; Sutarsyah, Nation & Kennedy, 1994; Vongpumivitch, Huang & Chang, 2009). More recent studies such as Chen and Ge (2007) and Khani and Tazik (2013) have demonstrated a very different composite of vocabulary content, derived from a specialized corpus, as compared to the AWL which is derived from a general academic corpus. Ample research can be found to underscore the necessity of compiling discipline-specific word lists derived from specialized corpora for English for specific purposes (ESP) learners (e.g., Martinez, Beck & Panza, 2009; Wang, Liang & Ge, 2008). For example, Martinez, Beck and Panza (2009) constructed a specialized corpus of research articles in the agricultural sciences and compared their specialized word list with the AWL. Their results showed that only 72 word families coincide with those in the AWL, or approximately 10% of the specialized corpus. The authors also pointed out that some frequent academic words such as *culture* and *strategy* have technical meanings and collocations that are discipline-specific. For example, in agriculture the word *strategy* collocates with *control*, *management*, and *adaptation* (compare, e.g., *learning strategy* in applied linguistics and *marketing strategy* in business).

In a cross-disciplinary study of sciences, social sciences, and engineering, Hyland and Tse (2007) also illustrated that some academic words vary in terms of their meanings and distribution and collocational patterns due to their distinct usage in each discipline. In addition, the study indicated intra-discipline variation by comparing the vocabulary concentration for each sub-discipline (e.g., biology, physics, and computer science for the sciences discipline). For example, about 52% of all unique words (283 items) in engineering cover more than 65% of all items occurring in the engineering sub-corpus. In the sciences, 244 families (or 43%) cover more than 65% of all items in the sciences sub-corpus, and 128 families (22.5%) in the social sciences sub-corpus cover more than 65% of all items. Hyland and Tse urged ESP practitioners to note that the patterns they found “suggest a more complex picture of language use in the disciplines than notions of a general academic vocabulary allow, pointing to more specialized language uses” (p. 243). In other words, the difference between a

specialized corpus and a reference corpus from which academic vocabulary is derived has been recognized.

Overall, early studies demonstrated the effectiveness of a corpus-based approach to analyzing synonymous sets of adjectives. Although the comparison of a specialized corpus against a general one reflects the different composition of a corpus, such a contrast may not be easily recognized by second/foreign learners of English. Therefore, the purpose of this study is to compare the linguistic behavior of a near-synonym pair in two corpora.

Methodology

Two methods were employed to carry out the analysis of *important* and *critical*. First a sense disambiguation analysis was applied involving the consultation of a learner's dictionary. Second, a corpus-based approach was taken to examine the collocates of the two adjectives. The following subsections explain the research procedures in detail.

Sense Disambiguation Analysis

We first searched for the meanings of both adjectives in the online *Merriam-Webster Learner's Dictionary* (<http://www.learnersdictionary.com/>). A learner's dictionary was our target of investigation, because the results of the analysis can be directly applied to teaching. The sense entries of a learner's dictionary are restricted to a limited word list and provide the most common linguistic uses for language learners (Kernerman, 1996). This particular dictionary was selected for two main reasons. First, it is publicly accessible and, second, it provides at least three examples for each sense. Because our sense disambiguation analysis relied heavily on examples to provide specific linguistic contexts for making comparisons between two words, this dictionary suits our purposes well. We expect similar results would be produced if other dictionaries were used. The sense disambiguation analysis involved three major steps. First, the online *Merriam-Webster Learner's Dictionary* was used to search for *important* and *critical*, respectively. The search results were transferred to an Excel

file for further analysis. Next, the senses from both words were compared along with the examples. Lastly, a replaceability test, which substituted one target word for another, was then applied to the dictionary examples. Then the authors judged whether the replacement would invoke a different meaning from the original in context. If the replacement did not involve any changes in meaning, the sense would be considered as overlapping for both words.

Table 3 *Meaning, Examples, and the Replaceability Test of Important and Critical*

Senses	Examples from the Dictionary (square brackets taken directly from the dictionary)	Replaceability Test (square brackets added by the author)
(1) Important		
(a) (i) having serious meaning or worth	1. She's an important [=significant] part of the team 2. Diet and exercise are important for health.	1. ?She's a critical part of the team 2. ?Diet and exercise are critical for health.
(ii) deserving or requiring serious attention	1. an important problem 2. In his editorial, he made several important points. 3. It's important that you remember to send these forms on time.	1. a critical problem 2. In his editorial, he made several critical points. 3. It's critical that you remember to send these forms on time
(b) having power, authority, or influence	1. He's one of the most important scholars in his field. 2. an important artist	1. ?He's one of the most critical scholars in his field 2. a critical artist =[a dying artist or one that has a critical judgment?]
(2) Critical		
(a) expressing criticism or disapproval	1. You're always so critical . 2. They are often critical of the mayor's policies.	1. ?You're always so important . 2. *They are often important of the mayor's policies.
(b) of or relating to the judgments of critics about books, movies, art, etc.	1. The book received much critical acclaim. [=many critics said good things about the book] 2. critical writings/theory 3. The movie was a critical success [=critics liked the movie], but it didn't make much money.	1. The book received much important acclaim. 2. important writings/theory [important ≠ critical] 3. ?The movie was an important success, but it didn't make much money.

(c) using or involving careful judgment about the good and bad parts of something	<ol style="list-style-type: none"> 1. <i>The program presents a critical analysis of the government's strategies.</i> 2. <i>She has a talent for critical thinking.</i> 3. <i>We need to look at these proposed changes with a critical eye before we accept them.</i> 	<ol style="list-style-type: none"> 1. The program presents an important analysis of the government's strategies. [important ≠ critical] 2. *She has a talent for important thinking. 3. *We need to look at these proposed changes with an important eye before we accept them.
(d) extremely important	<ol style="list-style-type: none"> 1. <i>We have reached a critical phase of the experiment.</i> 2. <i>It is absolutely critical [=vital, essential] for us to remain together.</i> 3. <i>This is a matter of critical importance to the future of our country.</i> 	<ol style="list-style-type: none"> 1. <i>We have reached an important phase of the experiment.</i> 2. <i>It is absolutely important for us to remain together.</i> 3. *<i>This is a matter of important importance to the future of our country.</i>
(e) medical: relating to or involving a great danger of death	<ol style="list-style-type: none"> 1. <i>He suffered critical injuries in the accident.</i> 2. <i>The patient is in critical condition. = The patient is very sick or injured and may die]</i> 3. <i>a nurse who specializes in critical care [=the care of patients who are in critical condition]</i> 4. <i>patients who are on the critical list; sometimes used figuratively (a government program that is on the critical list) [=that is in danger of failing or being eliminated]</i> 	<ol style="list-style-type: none"> 1. *<i>He suffered important injuries in the accident.</i> 2. ?<i>The patient is in important condition. ?The patient is important.</i> 3. *<i>a nurse who specializes in important care</i> 4. *<i>patients who are on the important list; *a government program that is on the important list</i>

Collocates

In order to understand the two adjectives better, we then compared their collocates in two different genres – the British National Corpus (BNC) and the British Academic Written English (BAWE) corpus. The frequencies of *important* and *critical* in the BNC are respectively 38,716 and 5,627, while those in BAWE are 5,485 and 1,469. From the comparisons, we can say that *important* has more general use because it occurs much more often in the reference corpus – i.e., 6.88 times more than *critical*. Yet, in BAWE, its frequency is 3.73 times higher than *critical*, indicating the more prominent role of *critical* in academic than in general discourse.

Most earlier studies seem to have neglected the possibility that the usage of a pair of near-synonyms might differ when they appear in different genres. From Table 4, we can see clearly how *important* and *critical* differ in a reference corpus (BNC) and in academic writing (BAWE). The collocates were taken from the Sketch Engine and arranged in descending saliency.

From Table 4, *important* does not vary greatly in the two genres, with some shared collocates (bolded) but they appear in different orders. What differs greatly is the adjective *critical*, where in the reference corpus, *critical* modifies *acclaim* and *appraisal* more (sense 2b), but in the academic one, it modifies *lure*, *essays*, *path*, etc. which are totally different from the collocates in the reference corpus. Examples (3) below show the use of *critical* in the academic discourse.

(3) (a) *In the case of our experiment, the **critical lures**, by their nature, come very close to the familiarity threshold level of the old words that are actually on the lists.*

(b) *In technical terms, the **critical path** of a project or a system can be defined as the path in which the "float" of all the processes is equal to "Zero".*

From the next analysis of shared and non-shared collocates, the differences will become more apparent.

Table 4 Noun Collocates Modified by Important and Critical

BNC (General)		BAWE (Academic)	
<i>Important</i>	<i>Critical</i>	<i>Important</i>	<i>Critical</i>
<i>factor</i>	<i>acclaim</i>	<i>role</i>	<i>lure</i>
<i>role</i>	<i>appraisal</i>	<i>factor</i>	<i>essays</i>
<i>aspect</i>	<i>mass</i>	<i>aspect</i>	<i>path</i>
<i>part</i>	<i>evaluation</i>	<i>part</i>	<i>realism</i>
<i>feature</i>	<i>scrutiny</i>	<i>issue</i>	<i>dialogue</i>
<i>issue</i>	<i>judgement</i>	<i>element</i>	<i>thinking</i>
<i>thing</i>	<i>factor</i>	<i>thing</i>	<i>period</i>
<i>element</i>	<i>comment</i>	<i>implication</i>	<i>wage</i>
<i>point</i>	<i>thinking</i>	<i>feature</i>	<i>introduction</i>

<i>implication</i>	<i>faculty</i>	<i>point</i>	<i>evaluation</i>
<i>contribution</i>	<i>pulsatance</i>	<i>determinant</i>	<i>discussion</i>
<i>consideration</i>	<i>examination</i>	<i>source</i>	<i>load</i>
<i>difference</i>	<i>importance</i>	<i>tool</i>	<i>angle</i>
<i>source</i>	<i>theory</i>	<i>distinction</i>	<i>theorist</i>
<i>question</i>	<i>temperature</i>	<i>consideration</i>	<i>interrogation</i>
<i>step</i>	<i>juncture</i>	<i>contribution</i>	<i>survey</i>
<i>consequence</i>	<i>frequency</i>	<i>theme</i>	<i>pg</i>
<i>respects</i>	<i>analysis</i>	<i>reason</i>	<i>factor</i>
<i>component</i>	<i>load</i>	<i>question</i>	<i>hypothesis</i>
<i>influence</i>	<i>review</i>	<i>step</i>	<i>analysis</i>
<i>matter</i>	<i>discourse</i>	<i>component</i>	<i>passions</i>
<i>distinction</i>	<i>writings</i>	<i>influence</i>	<i>value</i>
<i>function</i>	<i>edition</i>	<i>parameter</i>	<i>realist</i>
<i>lesson</i>	<i>stance</i>	<i>decision</i>	<i>guide</i>
<i>determinant</i>	<i>reflection</i>	<i>information</i>	<i>anthology</i>

As for the modifiers of *important* and *critical*, the results in Table 5 were found. From here, one could see that the collocates of *important* are modified by quite similar modifiers (*most*, *more*, *particularly*, *very*, etc.). Comparatively, *critical* has fewer modifiers in academic than in general discourse. *Openly*, *fiercely*, *sharply*, *harshly*, *strongly*, *bitterly*, etc., most of which contain an evaluative feature, are hardly found in academic writing.

Table 5 *Modifiers of Important and Critical*

BNC (General)		BAWE (Academic)	
<i>Important</i>	<i>Critical</i>	<i>Important</i>	<i>Critical</i>
<i>most</i>	<i>highly</i>	<i>most</i>	<i>absolutely</i>
<i>more</i>	<i>openly</i>	<i>very</i>	<i>highly</i>
<i>particularly</i>	<i>fiercely</i>	<i>particularly</i>	<i>most</i>
<i>very</i>	<i>sharply</i>	<i>extremely</i>	<i>more</i>
<i>equally</i>	<i>harshly</i>	<i>more</i>	<i>very</i>
<i>as</i>	<i>strongly</i>	<i>especially</i>	<i>also</i>

<i>extremely</i>	<i>bitterly</i>	<i>as</i>	
<i>less</i>	<i>especially</i>	<i>also</i>	
<i>vitaly</i>	<i>particularly</i>	<i>increasingly</i>	
<i>increasingly</i>	<i>mildly</i>	<i>equally</i>	
<i>especially</i>	<i>severely</i>	<i>so</i>	
<i>so</i>	<i>absolutely</i>	<i>less</i>	
<i>also</i>	<i>extremely</i>	<i>highly</i>	
<i>crucially</i>	<i>equally</i>	<i>therefore</i>	
<i>all</i>	<i>unusually</i>	<i>vitaly</i>	
<i>therefore</i>	<i>increasingly</i>	<i>all</i>	
<i>really</i>	<i>deeply</i>	<i>crucially</i>	
<i>terribly</i>	<i>most</i>	<i>hugely</i>	
<i>obviously</i>	<i>less</i>	<i>critically</i>	
<i>strategically</i>	<i>more</i>	<i>greatly</i>	
<i>quite</i>	<i>as</i>	<i>quite</i>	
<i>too</i>	<i>very</i>	<i>obviously</i>	
<i>potentially</i>	<i>both</i>	<i>still</i>	
<i>least</i>	<i>so</i>	<i>economically</i>	
<i>that</i>	<i>too</i>	<i>now</i>	

When we compared these two adjectives in terms of their shared and non-shared collocates using the Sketch Differences feature, the following results were found. In Table 6, we only considered those that are not particularly more frequent with any of the adjectives, or in other words, the collocates that are shared equally by both adjectives. From Table 6, one can see that the shared collocates of *important* and *critical* are different in the BNC and BAWE under the same grammatical relation. For instance, the modifiers of *important* and *critical* are *clearly/especially/often* etc. (i.e., *clearly/especially/often important/critical*) but only *highly important/critical* is equally shared by *important* and *critical* in BAWE. This table tells us that the two adjectives are behaving differently in the two registers.

Table 6 Shared Collocates of Important and Critical (in descending saliency values)

Grammatical Relations	BNC	BAWE
Modifier (e.g., <i>clearly important/critical</i>)	<i>clearly, especially, often, now, both</i>	<i>highly</i>
Modified (e.g., <i>important/critical variables</i>)	<i>variables</i>	<i>point, method, data, component, difference, parameter, activity, stage, condition, character, level, system</i>
and/or (e.g., <i>important/critical and/or useful</i>)	<i>useful, recent, theoretical, certain, serious, great, current, active, old, historical, contemporary, substantial, major, independent, modern, traditional, popular</i>	<i>social, difference, new, legal</i>

After we had compared the shared collocates, we then looked at the non-shared collocates. Table 7 provides this information. From Table 7, one can see how the same adjective could differ in use between two genres. For instance, *important* is used differently in the BNC and BAWE except for *vitally*, *crucially*, and *all* (bolded items are found in both general and academic use, though still exclusively used with *important* only). This means that these collocates co-occur exclusively with *important* (and not *critical*, and vice versa) but there are still differences in the two corpora because the BAWE collocates are examples of academic use. The fewer bold entries there are in the table, the rarer are the collocates for a particular genre.

Table 7 Non-shared Collocates of Important and Critical

	important		critical	
	BNC	BAWE	BNC	BAWE
modifier	<i>vitally</i> <i>crucially</i> <i>all</i> <i>therefore</i> <i>terribly</i> <i>obviously</i> <i>strategically</i> <i>potentially</i> <i>least</i> <i>that</i> <i>critically</i> <i>sufficiently</i>	<i>particularly</i> <i>extremely</i> <i>especially</i> <i>as</i> <i>increasingly</i> <i>equally</i> <i>so</i> <i>less</i> <i>therefore</i> <i>vitally</i> <i>all</i> <i>crucially</i>	<i>openly</i> <i>fiercely</i> <i>sharply</i> <i>harshly</i> <i>strongly</i> <i>bitterly</i> <i>mildly</i> <i>severely</i> <i>absolutely</i> <i>unusually</i> <i>deeply</i> <i>little</i>	<i>absolutely</i>
modified	<i>implication</i> <i>contribution</i> <i>consideration</i> <i>source</i> <i>consequence</i> <i>respects</i> <i>influence</i> <i>matter</i> <i>lesson</i> <i>determinant</i> <i>exception</i> <i>ingredient</i>	<i>element</i> <i>thing</i> <i>implication</i> <i>determinant</i> <i>tool</i> <i>distinction</i> <i>consideration</i> <i>contribution</i> <i>theme</i> <i>reason</i> <i>question</i> <i>step</i>	<i>acclaim</i> <i>appraisal</i> <i>mass</i> <i>evaluation</i> <i>scrutiny</i> <i>judgement</i> <i>thinking</i> <i>faculty</i> <i>pulsatance</i> <i>importance</i> <i>temperature</i> <i>juncture</i>	<i>lure</i> <i>essays</i> <i>path</i> <i>realism</i> <i>dialogue</i> <i>thinking</i> <i>period</i> <i>wage</i> <i>introduction</i> <i>evaluation</i> <i>discussion</i> <i>load</i>
and/or	<i>urgent</i> <i>interesting</i> <i>influential</i> <i>archaeological</i> <i>valuable</i> <i>only</i> <i>additional</i> <i>strategic</i> <i>contributory</i> <i>prognostic</i> <i>worthwhile</i> <i>structural</i>	<i>single</i> <i>desirable</i> <i>ethical</i> <i>conservation</i> <i>influential</i> <i>clinical</i> <i>interesting</i> <i>other</i> <i>strategic</i> <i>popular</i> <i>useful</i> <i>relevant</i>	<i>constructive</i> <i>informed</i> <i>creative</i> <i>analytical</i> <i>reflective</i> <i>biographical</i> <i>scholarly</i> <i>hostile</i> <i>sceptical</i> <i>cynical</i> <i>bourgeois</i> <i>verbal</i>	<i>realist</i> <i>contemporary</i> <i>ongoing</i> <i>creative</i>

<p>adj_comp_of (e.g., <i>Such factors are considered important when choosing books and audiovisual materials...</i>)</p>	<p><i>consider</i> <i>seem</i> <i>deem</i> <i>regard</i> <i>view</i> <i>sound</i> <i>feel</i> appear <i>play</i> <i>look</i> <i>think</i> <i>see</i></p>	<p><i>deem</i> <i>prove</i> consider seem appear <i>show</i></p>		go
<p>adj_subject (e.g., <i>The distinction is important because...</i>)</p>	<p>distinction <i>friendship</i> <i>sex</i> <i>accuracy</i> <i>issue</i> consideration <i>continuity</i> <i>outcome</i> <i>phenomenon</i> <i>aspect</i> <i>religion</i> <i>availability</i></p>	<p><i>preface</i> <i>witchcraft</i> <i>can</i> distinction <i>communication</i> consideration <i>autonomy</i> <i>variable</i> <i>customer</i> <i>factor</i> <i>criterion</i> <i>association</i></p>	<p><i>mission</i> <i>article</i> <i>design</i> <i>committee</i></p>	
<p>np_adj_comp_of (e.g., <i>You forget everything important,...</i>)</p>	<p><i>forget</i> <i>consider</i> <i>miss</i> <i>think</i> <i>regard</i> <i>remember</i> feel <i>believe</i> <i>lose</i> <i>find</i> <i>see</i></p>	<p><i>feel</i> <i>make</i> <i>find</i></p>	<p><i>publish</i></p>	
<p>pp_in-p (e.g., <i>...that the toxin is so important in the pathogenesis of disease...</i>)</p>	<p><i>pathogenesis</i> <i>aetiology</i> <i>context</i> <i>respect</i> <i>identification</i> <i>evaluation</i> <i>run</i> <i>climate</i> <i>interpretation</i> <i>formation</i> <i>right</i> <i>assessment</i></p>	<p><i>metabolism</i> <i>estimation</i> <i>decade</i> <i>respect</i> <i>context</i> <i>society</i> <i>movement</i> <i>industry</i> <i>development</i> <i>world</i> <i>approach</i> <i>year</i></p>		

<p>pp_to-p (e.g., ...<i>that does not mean that it is less important to our future economic <u>well-being</u>.</i>)</p>	<p><i>well-being Christian survival economy client individual animal child Britain woman community people</i></p>	<p><i>debate understanding business individual development company people</i></p>		<p><i>placement success project</i></p>
<p>pp_for-p (e.g., <i>not as important for its own <u>sake</u></i>)</p>	<p><i>sake survival purpose reason researcher pilot future understanding interpretation parent teacher reader</i></p>	<p><i>lodge archaeologist shareholder plc professional development organisation success growth business reason company</i></p>		
<p>pp_at-p (e.g., ...<i>but it is important at the <u>outset</u> to see what it amounts to...</i>)</p>	<p><i>outset stage level time point</i></p>			
<p>pp_as-p (e.g., <i>Far more important as a <u>source</u>...</i>)</p>	<p><i>source</i></p>	<p><i>source</i></p>		
<p>pp_of-p (e.g., ...<i>while critical of some <u>aspects</u> of CNAAs procedure...</i>)</p>	<p><i>factor</i></p>		<p><i>aspect failure policy lack proposal way plan practice quality approach government decision</i></p>	

As shown in Table 7, *important* is more register-specific in grammatical relations where its nominal collocates (underlined) are present in a post-modifying prepositional phrase (e.g. *important in the pathogenesis of disease; less important to our future economic well-being*).

Discussion

In the present study, we compared the linguistic behaviors of a near-synonym pair, *important* and *critical*, in both a reference corpus and a specialized corpus. We found that the two words behave divergently in the two genres, and this would very likely affect how we would go about teaching these particular adjectives. As illustrated in Tables 4 and 5, the noun collocates and modifiers of *important* and *critical* for both corpora differ to a certain extent. Furthermore, as shown in Table 6, where shared collocates of *important* and *critical* are presented, we could not find any overlap between the two corpora. This finding indicates that the two corpora contain different linguistic contexts for our target words and, thereby, give rise to different results when a collocation analysis is applied. When we paid attention to the near-synonym pair, the corpus results showed that the two near-synonyms displayed both similarities and differences. The most obvious difference can be seen in the results of non-shared collocates of the two adjectives (see Table 7); the two adjectives vary in the way they are used in the two corpora. More shared collocates between the two corpora can be identified for *important* than *critical*.

Previous studies on near-synonyms tend to draw research results directly from a single corpus source, neglecting the different results which may be obtained from different corpora, and in different registers. Here in our study we directly addressed this issue by extending the research findings to practice. Our data revealed degrees of variation when collocation analysis is applied to different corpora. Even for frequent adjectives such as *important* and *critical*, the corpus-derived difference is apparent. Consequently, when it comes to teaching and learning these adjectives in ESP/EAP classes, we need to take into account corpus type. A reference corpus such as BNC is useful due to its generality as well as size; however, when general or disciplinary

variations are of concern, a specialized corpus would be needed to accompany the general one. While the reference corpus provides us with a wide range of possible collocates, the specialized corpus provides us with needed specificity.

Conclusions

In this study, we analyzed a near-synonym pair, *important* and *critical*, in a reference corpus (the BNC) and a specialized corpus (BAWE). As demonstrated by the corpus results of the near-synonym pair, a reference corpus does not produce a set of collocates that is very similar to a specialized corpus. We urge ESP/EAP practitioners to be aware of the differential results derived from different corpora and on genre variations when applying corpus analysis. Future work can be carried out to extend our research findings to the classroom and to investigate the pedagogical effect of using a reference corpus and a specialized corpus for ESP/EAP learners in writing.

Acknowledgements

The authors would like to thank the Taiwan Ministry of Science and Technology for supporting the research herein under the research grant 104-2420-H-004 -003-MY2.

The BAWE data in this study comes from the British Academic Written English (BAWE) corpus, which was developed at the Universities of Warwick, Reading and Oxford Brookes under the directorship of Hilary Nesi and Sheena Gardner (formerly of the Centre for Applied Linguistics [previously called CELTE], Warwick), Paul Thompson (formerly of the Department of Applied Linguistics, Reading) and Paul Wickens (Westminster Institute of Education, Oxford Brookes), with funding from the ESRC (RES-000-23-0800). The BNC data cited herein have been extracted from the British National Corpus, distributed by Oxford University Computing Services on behalf of the BNC Consortium. All rights in the data cited are reserved. (http://www.reading.ac.uk/internal/appling/bawe/sketch_engine_bawe.htm)

References

- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam: Benjamins.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Chen, Q., & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL Word families in medical research articles (RAs). *English for Specific Purposes*, 26, 502–514.
- Chung, S.-F. (2011). A corpus-based analysis of “create” and “produce.” *Chang Gung Journal of Humanities and Social Sciences*, 4(2), 399–425.
- Chung, S.-F. and Ahrens, K. (2008). MARVS Revisited: Incorporating Sense Distribution and Mutual Information into Near-Synonym Analyses. *Language and Linguistics: Lexicon, Grammar and Natural Language Processing*, 9(2), 415–434.
- Chung, T., & Nation, P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, 15, 103–116.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34, 213-238.
- Gries, S. Th. (2001). A corpus linguistic analysis of English -ic vs -ical adjectives. *ICAME Journal*, 25, 65-108.
- Gries, S. Th. (2004). HCFA 3.2. A program for R. Available upon request from S. Th. Gries at: <http://www.linguistics.ucsb.edu/faculty/stgries/>.
- Gries, S. Th., & Otani, N. (2010). Behavioral profiles: A corpus-based perspective on synonymy and antonymy. *ICAME Journal*, 34, 121-150.
- Hoffmann, K. (2014). A corpus-based analysis of the near-synonyms “nice”, “kind”, “lovely”, “friendly”, “gorgeous” and “pleasant”. Retrieved January 3, 2015 from http://www.academia.edu/7541436/A_corpus-based_analysis_of_the_nearsynonyms_nice_kind_lovely_friendly_gorgeous_and_pleasant_
- Hunston, S., & Sinclair, J. (2000). A local grammar of evaluation. In S. Hunston & G. Thompson (Eds.)

- Hunston, S., & Thompson, G. (Eds.) (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- Hyland, K. (2000). *Disciplinary discourse*. Harlow: Pearson Education.
- Hyland, K., & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Quarterly*, 41, 235-253.
- Kernerman, L. (1996). English learners’ dictionaries: How much do we know about their use? In M. Gellerstam et al. (Eds.), *Euralex '96 Proceedings*, Vol. II, pp. 405-414.
- Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC Journal*, 44, 209-232.
- Liu, D. (2010). Is it a *chief, main, major, primary, or principal* concern? A corpus-based behavioral profile study of the near-synonyms. *International Journal of Corpus Linguistics*, 15, 56-87.
- Martinez, I. A., Beck, S. C., & Panza, C. B. (2009). Academic vocabulary in agricultural research articles: A corpus-based study. *English for Specific Purposes*, 28, 183-98.
- Paquot, M. (2010). *Academic vocabulary in learner writing: From extraction to analysis*. London & New York: Continuum.
- Sutarsyah, C., Nation, P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based study. *RELC Journal*, 25, 34-50.
- Swales, J. M., & A. Burke (2003). “It’s really fascinating work”: Differences in evaluative adjectives across academic registers. In P. Leistyna & C. F. Meyer (Eds.) *Corpus analysis: Language structure and language use* (pp. 1–18). New York: Rodopi.
- Vongpumivitch, V., Huang, J., & Chang, Y. (2009). Frequency analysis of the words in the academic word list (AWL) and non-AWL content words in applied linguistics research articles. *English for Specific Purposes*, 28, 33-41.
- Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical academic word list. *English for Specific Purposes*, 27, 442-558.
- West, M. (1953). *A General Service List of English Words*. London: Longman.