

## 台灣的大學入學考試英文作文與英文教科書課文在詞彙 分布與轉承語使用的相關性研究

丁鴻儒<sup>1</sup>

### 摘要

本研究目的在探討英語教科書對台灣高中生大學入學考試英文寫作能力的影響。本研究採用語料庫分析法，為建立可靠的語料庫，本研究收集大學入學考試中心所收錄的優秀英文作文手寫稿（2009 年至 2016 年），共 210 篇，並將手寫稿轉檔為打字稿電子檔。另一方面，本研究收錄目前通行英文教科書的四種版本（第一冊至第六冊），共 270 篇文章。語料庫建立完成後，接著分析和比較兩個語料庫中詞彙分布和轉承語使用情形。本研究有以下的研究發現。首先，學生在 GAST（學科能力測驗）和 AST（指定考科測驗）兩種考試的寫作表現，在學術詞彙和寫作平均長度方面，存在顯著的差異性。其次，四種版本英文教科書在詞彙分布和轉承語使用方面存在著顯著的相關性。第三，在詞彙分布方面，英文教科書課文和學生英文作文在常用字彙 1000 等級，學術詞彙和詞彙密度方面都有顯著差異。對於轉承語的使用，包括遞增型，轉折型，因果型和時間型在內的四個主要類別，學生英文作文和英文教科書文章中都顯示出相似的分布結果。基於上述研究結果，本研究提出未來台灣高中英文教科書製作可參考的編寫原則，並且對於近年台灣教育當局所倡導的十二年國民教育改革英文課程綱要提出建議。

**關鍵詞：**詞彙分布、轉承語、英文教科書、英文作文

---

<sup>1</sup> 私立大同高中教師

通訊作者：丁鴻儒，E-mail: adamting@ms.tsh.ttu.edu.tw

## Introduction

Writing has been recognized as the essential and practical access to communication of an individual's ideas and concepts toward a given topic. Therefore, it is a rather difficult language skill compared with the other language competences. Writing involves complicated arrangement of linguistic items, such as words, phrases, sentence structures, and paragraphs in order to achieve the goal of communication. English writing for academic purpose, different from everyday speech for communicative purpose, requires more conventionally-acceptable language use to meet the expectation of discourse community (Spycher, 2007). As fundamental linguistic element, vocabulary use plays an important role in the effectiveness of academic writing. Studies point out that appropriate employment of words can effectively lead to better understanding of the writer's viewpoint on the part of readers (Nation, 2001). In addition to vocabulary use in English writing, researchers have studied the influence of the use of connective ties on the quality of English writing (Eherlish, 1988). Studies related to connective ties (e.g. linking adverbials, transitional words, conjunctive adverbials) indicated that such linguistic features are important devices which provide readers with specific guidance as to how the writer's main points are directed (Chen, 2006; Liu, 2008).

When it comes to learning to writing in English, EFL learners at beginning level tend to utilize more general words, or lexical teddy bears (Hasselgren, 1994), in their attempt to complete a composition. When learners gradually develop better linguistic competence, they are able to employ a wider range of vocabulary and more appropriate rhetorical structures that convey their ideas more successfully (Laufer, 1994). However, the difference between low-proficiency writers and high-proficiency ones does not solely lie in vocabulary size; vocabulary use, such as transitional words, also plays a vital role that determines the quality of writing (Sugita, 2012).

As EFL learners in the non-native English speaking environment, Taiwanese senior high school students usually study English following a regular process, starting with vocabulary, phrases, grammar, and then text. For Taiwanese English learners, they usually receive formal English writing instruction or do writing exercises most frequently at the third year in senior high school, because they are faced with the competitive college entrance examination which include formal English writing test (Chang, 2006). English Textbooks have always been the model texts for Taiwanese EFL learners when they are studying English in school setting. Accordingly, English teachers usually need to spend a lot of time preparing teaching materials, part of which may be supplied by the textbook publishers, before the new lessons begin. While they are performing formal instructions, teachers may particularly instruct

some focused vocabulary or sentence patterns mentioned in the textbooks and ask students to do some practices. This is to enable students to be familiar with the accurate use of the vocabulary or sentence patterns in the subsequent writing assignment. Thus, it is likely that Taiwanese EFL learners tend to employ English vocabulary or sentence patterns similar to those used in their English textbooks to certain extent in their English writing.

There have been studies aiming at investigating how Taiwanese English learners use vocabulary in academic writing, most of which focused on vocabulary size (e.g. Chen, 2008). On the other hand, although some researchers have tried to study how Taiwanese college-leveled English learners utilize transitional words in their writing (Chen, 2006; White, 2015), there is not any study trying to analyze the transitional words use in Taiwanese senior high school students' writing, not to mention the study combining both vocabulary size (or vocabulary distribution) and transitional words use. To bridge the gap, it is necessary to conduct a study which incorporates the analysis of the above-mentioned two factors in students' writing, so that the result of which can shed light on the effect of current English textbooks on Taiwanese senior high school students English learning.

The most straightforward approach to understanding Taiwanese senior high students' writing in respect of vocabulary distribution and transitional words use is through analyzing how they write in a formal writing test.

Writing tests can provide teachers with important information regarding their students' current writing competence. Formal writing assessments for Taiwanese senior high school graduates include timed short essay writing tests in GSAT (General Subject Aptitude Test) and AST (Appointed Subject Test), two nationwide tests for college entrance admission in Taiwan. As stated in guidelines of college entrance examination center (CEEC, 2010), one of the objectives for GAST and AST is "to evaluate ability to understand English words, phrases (including content words, function words, fixed expressions, and transitional words, etc.), sentences and paragraphs using context clues." The analysis of Taiwanese high-proficiency writers' writing in terms of vocabulary distribution and transitional words use can help the researcher to unveil the overall phenomenon of how those advanced-level students make use of what they learned from textbooks into their written works. The results can benefit both senior high school students and Taiwanese English education practitioners (teachers and textbook editors) for their future English learning and teaching design.

Based on the rationale as mentioned above, this study intended to examine the relationship between advanced Taiwanese senior high school students' writing and

four commercialized English textbooks with respect to vocabulary distribution and the use of transitional words and tried to answer the research questions as follows,

1. What are the vocabulary distribution and transitional word use by the proficient Taiwanese EFL high school students in college entrance examination?
2. What are the vocabulary distribution and transitional word use in Taiwanese English textbooks for senior high school students?
3. Is there any significant difference between advanced students writing and the lessons in the four commercialized English textbook in terms of vocabulary distribution and transitional word use?

### Literature review

In this section, previous research regarding vocabulary learning, evaluation of vocabulary size, and vocabulary distribution in academic text are reviewed. Furthermore, the use of transitional words is discussed, particularly how they are used in Taiwanese advanced EFL learners' writing is reviewed. Finally, the importance of combining vocabulary distribution and transitional words into a study aiming at analysis of high-proficiency English learners' writing is discussed.

Compared with other language skills, writing is usually considered a more difficult language skill to learn for EFL learners. How to produce a piece of good writing frequently puzzles either students or writing teachers. Writing definitely involves the use of vocabulary, which paves the way for writing at different language level.

Vocabulary learning is not instantaneously acquired, but rather, "they are gradually learned over a period of time from numerous exposures." (Schmitt, 2000)." For EFL learners, knowing how to use vocabulary appropriately to meet the expectations of academic writing in English takes time. Academic writing is different from writing for everyday communication, which may not stress some special linguistic features, such as specialized use of vocabulary or connective ties within sentences. Academic writing, narrowly defined, refers to writing in school settings that prepares senior high school students to continue their academic careers into college. To be equipped with sufficient knowledge concerning academic writing appears essential for senior high school students, for without it students are unable to cope with upcoming discipline-oriented writing task in the university, like thesis writing, or term paper.

Knowledge concerning academic writing for intermediate EFL learners, at lexical domain, involves vocabulary size and vocabulary use, which are deemed necessary to determine the quality of writing (Laufer and Nation, 1995). In their study,

Laufer and Nation proposed four popular measures for vocabulary use in writing: lexical originality, lexical density, lexical sophistication, and lexical variation, among which lexical density and lexical variation are suitable for analysis of writing of EFL learner at intermediate level. The other two measures, lexical originality and lexical sophistication, tend to focus on more advanced vocabulary use and therefore not appropriate for analyzing young EFL learners' writing.

Lexical density calculates the percentage of content words to function words. The former are practical words that include nouns, verbs, adjectives, and adverbs, while the latter refers to grammatical words, such as pronouns, determiners, prepositions, modal verbs, auxiliary verbs, and conjunctions. Theoretically, the higher lexical density a text displays, the better versatility a text can indicate. On the other hand, lexical variation, or lexical diversity, is often indicated by way of presentation of type-token ration. Word types refer to different words (i.e. repeated words are counted one word), while word tokens are running words (total number of words) in the text. The calculation of type-token percentage is more effective in distinguishing basic-level writing and advanced-level writing; basic-level writers usually display less vocabulary size while advanced-level writers usually possess stronger word power. Compared with lexical diversity, the calculation of lexical density for analysis of an English composition appears more appropriate for advanced-level writers.

With the advance of computer technology, current studies of vocabulary size, or vocabulary distribution, usually rely on corpus analysis. Corpus studies in vocabulary learning and teaching in recent years have yielded important information in terms of vocabulary distribution in academic writing. Early corpus study on vocabulary distribution was Kučera and Francis' research (1967), who displayed vocabulary size and text coverage in the Brown University corpus, a corpus focusing on American English. Another study made by Coxhead (1998) listed the coverage by the different kinds of vocabulary in an academic corpus. Table 1 displayed the vocabulary type and text coverage in the two mentioned corpora of academic writing.

Table 1. *Vocabulary Distribution in two Corpus Studies*

Type of vocabulary	Text coverage (Coxhead, 1998)	Text coverage (Kucera & Francis, 1967)
1 <sup>st</sup> 1000 words	71.4%	72.0%
2 <sup>nd</sup> 1000 words	04.7%	07.7%
Academic words list (570 w)	10.0%	N.A
Others	13.9%	N.A

*Notes.* Others refers to low-frequency words

As indicated in Table 1, the first and second 1000 words cover 75 percent to 80 percent of text in academic writing. Academic words include 10 percent of the text, while others, meaning low frequency words in different content, are around 14 percent, depending on the specific genre type or field-oriented writing.

English learners' vocabulary size in formal writing test has also been studied by researchers. Aiming to investigate overall Canadian undergraduate novices' level of writing competence, which is deemed necessary for students to cope with university-level reading and writing task, Douglas (2013) collected writing samples from undergraduate students who passed the Effective Writing Test (EWT) and built a corpus of 120 papers, 62,309 running words. He found that the General Service List (GSL) and the Academic Word List (AWL) cover an average of 94% of a typical paper, with 3,000 word families and 5,000 word families covering 95% and 98% respectively of each paper. The result indicated that even native English-speaking students tend to utilize more high frequency words than low frequency words in their writing, compared with formal academic English writing.

Although researchers, like Laufer and Nation (1995), repeatedly mentioned the importance of vocabulary size as indicators to determine the quality of writing (Nation, 2001), a study that aims at evaluating the quality of writing with exclusive focus on vocabulary distribution is never complete enough. Writers who possess limited vocabulary size are able to produce good writing, on condition that those words are properly chosen and skillfully used. As Steve Peha (2016) mentioned in his article, *what is good writing*, one of the six traits of good writing is Word Choice that is specific and memorable. Good writing uses just the right words to say just the right things. He stressed that "Good word choice doesn't mean using big, fancy, unusual words. It means using the right words to say the right thing in just the right way." For EFL learners, a piece of good writing is not a text which involves only large vocabulary size; rather, it is concerned with how to present the words logically, or coherently, which is even more important.

In addition to vocabulary size, another important factor leading to well-organized writing is the use of linking devices. Similar terms used in related studies on topic of linking devices for L1 Chinese students include, among other things, "transitional markers" (White, 2015), 'conjunctive adverbials' (Chen, 2006), "linking adverbials" (Leedham and Cai, 2013). In this study, the linking devices is termed as transitional words, in order to conform to CEEC's guidelines (2010).

The importance of transitional words in academic writing lies in their direction-guiding function. Transitional words bridge ideas within sentences, or to a larger extent, paragraphs. A writer's ideas are logically connected by way of

transitional words. Therefore, skillful use of transitional words are likely to make up insufficient vocabulary size in term of communication purpose.

Researchers centering on transitional words have tried to make a classification of those words in their studies, most of which followed Halliday and Hasan's classification (1976). Celce-Murcia and Larsen-Freeman's proposed a simplified version (1999) of Halliday and Hasan's four-way classification which divides conjunctions into the following four types: additive, adversative, causal and temporal. This study adopted Celce-Murcia and Larsen-Freeman's simplified version (see appendix A) in the analysis of the two corpora.

Chen (2006) investigated the use of conjunctive adverbials (CAs) in the academic papers of advanced Taiwanese EFL learners. In her study, she built two corpora, one learner corpus from Taiwanese EFL learners, and the other reference corpus in which articles from published Journals were collected, and then compared the use of linking devices in writing from two academic discourse communities. She found that in learner corpus, additive category occupied the highest percentage of use of CAs, 33.8%, causal had the second highest percentage, 24.6%, temporal and adversative ranked the third and the fourth respectively, 21.6% and 20.6%. As for the reference corpus, the highest ranking of percentage use of CAs was adversative category, 37%, followed by additive (30.9%), causal (18.9%), and temporal (13.2%) as the lowest ranking. The top 10 most frequently used CAs in learner corpus displayed slightly different ranking compared to those in reference corpus. Table 2 indicated the top 10 most frequently used CAs in the two corpora. ,

Table 2. *The top 10 most frequently used CAs*

Learner corpus			Reference corpus		
Rank	CA	Percentage*	Rank	CA	Percentage*
1	however	12.5%	1	however	26.5%
2	therefore	12.3%	2	for instance/ example	7.2%
3	for instance/ example	8.9%	3	thus	4.9%
4	thus	8.3%	4	therefore	4.9%
5	moreover	4.1%	5	that is	4.6%
6	besides	4.1%	6	furthermore	4.1%
7	also	4.1%	7	then (causal)	3.6%
8	first	4.1%	8	in fact (additive)	3.3%
9	then (temporal)	3.9%	9	in other words	2.8%
10	in addition	3.2%	10	on the other hand	2.8%
65%			67%		

Notes: percentage\*; % of the overall CAs used

White (2015) compared Taiwanese graduate students' transition markers use with SSCI journal articles. By building up two corpora, one from 30 MA theses completed by Taiwanese graduate students and the other from 30 published articles from two SSCI journals, he found that both the Taiwanese graduate students and the published journal writers used similar transition markers. The top ten transition markers included *and*, *also*, *but*, *however*, *because*, *therefore*, *thus*, *although*, *so*, and *while*. He mentioned that the main differences between the two corpora were the frequency of use.

Both Chen (2006) and White (2015) made use of corpus approach to investigate the differences between Taiwanese advanced EFL learners' writing and published journal articles regarding linking devices. Their studies did not distinguish the use of linking devices in different rhetorical task type. In addition, the two studies aimed their target at Taiwanese EFL learners at graduate school level, whose writing were essentially required to meet the expectation of academic writing. Therefore, the results obtained cannot account for the general phenomenon of Taiwanese senior high school students' writing performance, particularly not applicable to formal timed writing test. On the other hand, Halliday and Hasan (1976) pointed out that linking devices are vital since they make texts coherently and coherently; however, Eherlich (1988) argued, "the frequency with which cohesive devices occur does not determine the overall effectiveness of these devices in establishing cohesive and coherent text.... It is appropriate fit of these devices to the context and their distribution throughout a text which determines their effectiveness (p. 113)."

Given the fact that studies focusing on either vocabulary size or the frequency counting of transitional words alone are unable to predict the effectiveness of good writing, a study combining the analysis of both vocabulary distribution and transitional words use is likely to grasp an overall picture of how Taiwanese EFL learners' writing are made up of. Based on the above-mentioned notion, the researcher made an attempt to incorporate analysis of vocabulary distribution as well as transitional words into the study of how Taiwanese high-proficiency high school students' writing is organized and to what extent their English textbook can affect their writing. The researcher believed that observation of students' writing which have been evaluated excellent may provide an opportunity for learners and researchers to analyze how vocabulary and transitional words are used in English writing, so that they can serve as models for future English writing learning and teaching, and provide a guidance for Taiwanese English textbook editors in their compiling works of teaching materials.



## Methodology

This study intended to compare Taiwanese advanced senior high school students' writing and their textbook's lessons in terms of vocabulary distribution and the use of transitional words. To achieve the goal, this study collected excellent writing samples published by CEEC (2009-2016, <http://www.ceec.edu.tw/>), and then transcribed those handwriting compositions into electric files, building up a student writing corpus. In this way, the vocabulary distribution and the use of transitional words can be calculated and compared. Table 3 indicated the examinees' writing samples in recent years.

Table 3. *A corpus of student writing samples from the year of 2009 to 2016*

Number of students writing samples by years										
Test	word	2009	2010	2011	2012	2013	2014	2015	2016	TTL
GAST	34049	13	10	10	15	14	15	15	15	107
AST	28842	10	10	15	13	12	14	15	14	103

Note. GAST (General Scholastic Ability Test); AST (Advanced Subjects Test)

As indicated in Table 3, examinees' writing samples include 107 compositions from GSAT and 103 from AST. All these compositions made up a corpus of 62,891 words. The corpus can provide a general picture for how Taiwanese advanced senior high school students present their writing competence in the formal writing test settings.

On the other hand, four versions of English textbooks currently available for Taiwanese senior high students were collected and then built up an English textbook corpus for further analysis. The English textbook publishers, four in total, are Far East, San Ming, Lungteng, and Nani, all of which have been reviewed by NAER (National Academy Education Research) as approved English textbooks and are adopted by senior high schools in Taiwan. (<http://review.naer.edu.tw/Bulletin/PA02.php>). The detailed words number of different Book volume in the four versions of textbooks were shown in Table 4.

Table 4. *Lesson numbers from four textbook publishers*

	Textbook A	Textbook B	Textbook C	Textbook D
Book I	12 (4,461)	12 (3,916)	12 (4,754)	12 (5,279)
Book II	12 (5,081)	12 (5,158)	12 (5,287)	12 (5,603)
Book III	12 (5,459)	12 (5,530)	10 (5,133)	12 (6,672)
Book IV	12 (7,196)	12 (6,954)	10 (5,853)	12 (7,477)
Book V	10 (6,010)	10 (5,942)	10 (5,951)	12 (8,358)
Book VI	10 (7,026)	10 (6,604)	10 (6,649)	10 (7,056)
TTL words	68 (35,223)	58 (34,104)	64 (33,627)	70 (40,445)

After the two corpora were established, the data were analyzed by way of the online software [www.lextutor.ca](http://www.lextutor.ca), which provides necessary results in respect of word vocabulary distribution and percentage of a variety of word types, such as content words, functional words, and lexical density. All of the results obtained from the two corpora were compared and analyzed, using one-way ANOVA, to indicate the correlation and difference.

In addition to analysis of vocabulary distribution using online software, the frequency of transitional words in the two corpus were obtained by using Microsoft Office Word. With its powerful sorting function, the results of frequency of each transitional word and locations were clearly indicated. Most importantly, the researcher was able to judge among those listed items if a given transitional word could be properly counted as accurate frequency. For example, *then* can be a temporal adverb meaning “at that time”; it can also be a transitional word indicating sequence, like “..., and *then*...”. In this case, the researcher needed to rule out inappropriate occurrence of a transitional word with multiple meaning, so that the calculation of frequency of each transitional word can comply with the definition of what is being taken a transitional words.

## Result and discussion

College entrance examination is usually the main access for Taiwanese senior high students to apply for college entrance admission. Every year, after the two college entrance examinations, GSAT (General Scholastic Ability Test) and AST (Advanced Subjects Test), held in January and in July respectively, CEEC will make public ten to fifteen excellent composition samples written by examinees on the website (<http://www.ceec.edu.tw/>), so that learners who are going to take the same tests next year can have a clear picture regarding how advanced English short writing is organized and evaluated.

For Taiwanese university-bound senior high graduates, the two nationwide college entrance writing tests in Taiwan, GSAT and AST, put emphasis on different test purposes, as indicated in their names, and therefore both adopt different rhetorical task types. The former appears to focus on narrative easy writing, while the latter tend to stress the evaluation of examinees' expository and descriptive writing competence (Tsai & Cheng, 2009).

To obtain different results of different writing genre, the building of examinees' writing corpus was built separately and went through some particular process. Student writing examples in this study, though rated as excellent writing, contained a number of grammar errors or wrongly spelt words. To accurately calculate vocabulary

distribution in each written work, it is necessary to correct all these errors ahead of digitization of the handwritten compositions.

In order to further investigate if there exists any significant difference of vocabulary distribution between examinees' writing performance in GSAT and AST, in terms the first 2000 words (General Service List) and Academic Word List, one way ANOVA was used to do the comparison. The results were shown in Table 5.

Table 5. *Comparisons of percentage at different vocabulary word level of students writing in GSAT and AST*

Items	GSAT	AST	<i>p</i> value
1000 word level	82.72	82.95	.69
2000 word level	6.46	5.89	.07
Academic word	2.84	3.78	.00**
Off word list	7.95	7.38	.19
Lexical density	.49	.49	.93
Average words	321	291	.00**

Note. \*\* $p < .01$

As indicated in Table 5, the examinees' performance in either GAST or AST did not reveal significant differences in the first 2000 word level (the most frequently used words), both of which reached a coverage of 88% of vocabulary distribution in each typical English writing. This result is identical with Douglas' study (2013) on novice Canadian undergraduate writing competence. In another two categories, off word list, which refers to those words beyond the level of the first 2000 words, and lexical density, which is the ratio of content words to function words indicating the text versatility, all examinees did not manifest remarkable dissimilarity.

However, there are two categories, academic word (words that may occur in any academic discipline) and average words (total words in each composition), showing remarkable dissimilarity between students' writing in GAST and AST. The latter employed more academic words, 3.78%, than the former, 2.84. ( $p < .001$ ). The average words in GSAT (321 words) are more than those in AST (291 words, ( $p < .003$ )). These differences are caused due to the distinctive requirement of the two tests by nature. As its name revealed, GSAT intends to evaluate test-takers' general competence in English writing, and thus the test allows longer time for students to complete their compositions; the overall test lasts for 100 minutes. Unlike GSAT, AST sets its target to assess students' academic writing competence with higher standard: the testing time is 80 minutes, which is obviously shorter than GSAT. Another significant difference between the two tests lies in the vocabulary competence requirement. GSAT is held during the winter break in each school year, normally in January; therefore, students

need to equip themselves with a vocabulary size of 4,500 word level. Those who intend to attend AST must enrich their vocabulary competence to 7,000 word level, which is almost 1.5 times larger of vocabulary size than that needed for GSAT. Without a doubt, the examinees' writing performance in GSAT will indicate larger academic words than those in AST.

On the other hand, rhetorical task type may lead to different writing performance. Studies of the effect of rhetoric task type on EFL learners' writing performance have produced mixed results, Tsai and Cheng (2009) attributed these mixed findings to variation of writing research focuses and designs. For example, Faigley et al (1981) found in their study that the effect of writer's apprehension on writing performance was significant in descriptive and narrative essay, not in argumentative essay. Koda's study (1993) indicated that narrative discourse involves more demanding linguistic processing, at varying levels, than descriptive discourse. Nevertheless, a number of studies which compared different rhetorical task type revealed that narrative writing is deemed more preferable than expository and argumentative writing on the part of EFL learners. For instance, Cheng's study (2006) found that Taiwanese college students produced longer narrative essay than argumentative essay, although the former included less syntactic complexity while the latter contained higher syntactic complexity.

In addition to analysis of vocabulary distribution in both examinees' corpus, those used in textbook lessons is discussed. English textbooks provide the major sources for Taiwanese senior high school students to learn English. The lessons in each English textbook follow the curriculum guideline issued by MOE (Ministry of Education), which regulates detailed criteria, such word levels, grammatical or sentence patterns, and content focuses, for publishers to edit their textbooks. Therefore, it is clear that the difficulty level in textbook lessons from Book I to Book VI may show different vocabulary distributions and grammar focus. Table 4 indicates the word numbers in each English textbook from the four publishers in Taiwan.

As shown in Table 4, the word numbers in each textbook of the four versions display the similar distribution, and it is clearly indicated that word numbers increase greatly from Book IV to Book V in any version of English textbook. This means that the lessons for the students during their second and third year of senior high school have raised difficulty level either in depth and or in width. Based on the 2010 version of the Senior High School Curriculum Guidelines, the edition of English textbook for first-year senior high school students, Book I and Book II, should integrate the four language skills averagely, which means English writing will not be exclusively taught during normal instruction for tenth graders, and the focuses are aimed at connecting

English lessons for Grade 1-9 curriculum guideline. Therefore, it is until the second year that senior high school students start to receive formal English writing. The vocabulary distribution of the four commercialized English textbooks based on General Service List (West, 1953) and Academic Word List (Coxhead, 2000) are shown in Table 6.

Table 6. Comparisons of percentage at different vocabulary word level of commercialized English textbooks

Items	B1	B2	B3	B4	B5	B6	<i>p</i> value
1000 word level	83.24	83.20	81.53	79.95	80.00	78.36	.00**
2000 word level	6.69	6.64	6.69	6.61	6.56	5.95	.67
Academic word	1.81	2.52	3.13	3.91	3.51	4.58	.00**
Off word list	7.97	7.69	8.65	9.29	9.94	11.09	.00**
Lexical density	0.52	0.51	0.52	0.53	0.52	0.54	.02*
Average words	383.65	442.33	498.07	600.65	605.48	682.4	.00**

Note. \*\* $p < .01$ , \*  $p < .05$

Table 6 apparently indicated that, except the vocabulary at the second 1000-word level, the lessons in each volume of the commercialized English textbooks present significantly different vocabulary distribution at the first 1000-word, academic word, and off-word list. In addition, the lexical density in each book volume does have different percentage. The total words through Book I to Book VI also display a significant tendency of increased word number. Book I and Book II, for the first-year senior high school learners, introduce the highest frequency word, the first and second 1000 words, because they are necessary for beginning EFL learners to learn. Those words cover at least 80 percent of text in academic articles (Coxhead, 1998). Book III through Book VI gradually reduce the amount of high frequency words, and increase academic words and low frequency words, which are important for university-bound high school students to start to cultivate their academic writing competence and skills (CEEC, 2010). Book VI can see the highest vocabulary distribution percentage in respect of academic words, 4.58%, as well as low frequency words, 11.09%. In this volume, average words per lesson can be as many as 682 words. Either vocabulary distribution or longer articles in Book VI aimed to prepare the twelfth graders for their academic learning in college level.

According to 2010 version of the Senior High School Curriculum Guidelines, the edition of vocabulary in English textbooks should follow gradual development principle, that is, high frequency word should be taught prior to low frequency word. On the other hand, the total amount of vocabulary in Book I and II for basic level learners, tenth graders, is suggested to be less than 600 words. As for intermediate level

learners and above (Book III, IV for eleventh graders, Book V and VI for twelfth graders) the overall new words should be around 700 words in each volume. Based on the above mentioned difference between Book I, II and those above, this study decided to dispense with Book I and Book II in the comparison of vocabulary distribution between student written works and articles in English textbooks. By doing so, vocabulary use in both corpora could be compared in similar linguistic condition. Table 7 showed the comparison between student writing corpus and textbook corpus.

Table 7. *Comparisons of percentage at different vocabulary word level between student writing and textbook lessons*

Items	Students Writing	Textbooks	<i>p</i> value
1000 word level	82.83	80.04	.00**
2000 word level	6.17	6.47	.19
Academic word	3.30	3.77	.03**
Off word list	7.67	9.69	.00**
Lexical density	.49	.53	.00**

Note. \*\* $p < .01$ , \*  $p < .05$

A further comparison of the two corpora regarding vocabulary distribution indicated that student writing significantly differed from textbook lessons in four categories, except the 2000 word level. In the category of the first 1000 word level, student writing (82.83%) employ more high frequency words than textbook lessons (80.04%), while the former use less academic word (6.17%) than the latter (6.47%). As for the other categories, off word list and lexical density, student writing does not use as many academic words as textbook articles, 3.3% for the former and 3.77% for the latter. In respect of lexical density, which is meant to indicate vocabulary richness in a text, it is apparent that textbook authors are more competent than student writers in the use of content words in the essay writing.

English textbooks for Taiwanese EFL senior high school students are graded readers, which vary in compliance with learners' current linguistic competence. For third-year high school students, they are provided nearly-authentic English reading materials for the purpose of enabling them to develop higher level of competence, which is fundamental requirement for them to engage in the academic demands of university studies. The increasing of academic words and low frequency words are meant to add their exposure to authentic articles written for native English users.

Evidently, those college-oriented examinees are suggested to enlarge vocabulary size, particularly academic words, for those words are necessary requirement for any discipline at college level and above.

In addition to the comparison between examinees' writing samples and lessons in the English textbooks regarding vocabulary distribution at different word levels, this study aimed to investigate the use of transitional words in both corpora. When student writing corpus was compared with textbook corpus in respect of transitional words, a surprisingly correlative outcome can be found, as shown in Table 8: either the students writing in GSAT or in SAT indicated significant correlation (both had a high correlation coefficient relation,  $r = .867$ ,  $p < .05$ ) to the four commercialized English textbooks. Even among the English textbooks, they also showed high correspondence regarding transitional word use (see Table 8). These results may imply that senior high school students learned how to make use of transitional words in the writing mainly from the instruction in the textbooks. This result echoed Leedham and Cai's study (2013), which demonstrated the influences of teaching materials on the students' learning of linking adverbials.

Table 8. *Correlation of transition word frequency among students writing and their textbooks*

	GSAT	AST	Text. A	Text. B	Text. C	Text. D	TTL words
GSAT	1	.86**	.46**	.70**	.60**	.57**	34,006
AST		1	.46**	.54**	.52**	.54**	29,846
Text. A			1	.59**	.87**	.91**	35,223
Text. B				1	.85**	.76**	34,104
Text. C					1	.93**	33,627
Text. D						1	40,445

Note. \*\* $p < .01$

The classification of linking adverbials has been discussed by researchers, among them, Celce-Murica and Larsen Freeman's simplified version of Halliday and Hasan's original four-way classification system was widely used (Chen, 2006, Liu, 2008). Chen (2006) added more conjunctive adverbials to the abovementioned system in order to "provide a more reliable benchmark which is less arbitrary in nature (p. 335)." Given the fact that Chen's study focused on examining Taiwanese EFL learners' use of conjunctive adverbials in academic writing, which provided valuable findings in respect of how Taiwanese college-level students make use of linking adverbials in academic writing, this study adopted Chen's modified version to compare the transitional words use in senior high school student writing and those lessons in textbooks.

The further comparison between the two categories in respect of transitional words was indicated in Table 9.

Table 9. *Total frequency comparison by type across corpora*

Transition	GSAT	AST	Text A	Text B	Text C	Text D	<i>p</i>
Additive	42.05 (1)	55.62 (1)	41.72 (1)	51.65 (1)	29.80 (1)	55.65 (1)	.96
Adversative	30.88 (2)	32.50 (2)	21.57 (3)	26.86 (2)	25.63 (2)	25.96 (3)	.99
Causal	18.53 (4)	23.79 (4)	22.99 (2)	20.96 (3)	23.24 (3)	37.59 (2)	.95
Temporal	20.88 (3)	30.82 (3)	17.32 (4)	17.32 (4)	21.99 (4)	20.26 (4)	.88
Total	112.34	142.73	103.6	116.79	100.66	139.46	

Notes. 1. Frequency per 10,000 words, 2. Number 1 to 4 in parentheses indicating the ranking

The four-way classification of transitional words includes additive, adversative, causal and temporal. The first type, additive transitional words, are meant to indicate semantically emphatic, appositional, and comparative relation within a sentence or between sentences. The second, adversative transitional words, are functioned as providing contrastive, correction, and dismissal relationship within a sentence or between sentences. The third type, Causal transitional words, serve as subsequent direction of an anterior cause, including general causal relation and causal conditional relation within a sentence or between sentences. As for the fourth type, temporal category of transitional words, refer to those linking words that point out sequential or summarizing relations within a sentence or between sentences.

The above results, as shown in Table 9, indicated that Taiwanese advanced senior high school examinees used similar transitional words in their writing to those used in English textbook ( $p > .05$ ), which echoes White's finding (2015) that Taiwanese graduate students and the published journal writers used similar transitional words. Generally speaking, additive category occupies the largest occurrence among the four types of transitional words: except Textbook C, the average frequency for the other five text groups rank form 41.72 per 10,000 words (Textbook A) to 55.65 per 10,000 words (Textbook D). As for the less frequently used type of transitional words, temporal, the result indicates that except AST, the other five text groups only use this type of transitional words at the frequency of 17.32 (Textbook A) per 10,000 words to 21.99 (Textbook C) per 10,000 words.

In view of the high correlation for both examinee writing samples (GAST and AST) and lessons of the four commercialized English textbooks, this study put two sets of examinee's writing as one category and the four English textbooks as another category, so that transitional words in both corpora can be specifically displayed.



Table 10. *Top ten additive transitional words in both students and textbook corpora*

Rank	Students' W.	raw <sup>a</sup>	ptw <sup>b</sup>	Textbook lessons	raw <sup>a</sup>	ptw <sup>b</sup>
1	also	151	23.65	also	301	20.99
2	In addition	23	3.60	for instance/ for example	113	7.88
3	Besides	23	3.60	in fact	72	5.02
4	For instance/example	22	3.45	in addition	44	3.07
5	Moreover	17	2.66	moreover	18	1.26
6	Furthermore	15	2.35	What's more	17	1.19
7	What's more	15	2.35	Besides	13	0.91
8	Most importantly (Most important of all)	11	1.72	at the same time	13	0.91
9	In fact	9	1.41	Additionally	10	0.70
10	What's worse	7	1.10	In other words	10	0.70

<sup>a</sup> raw frequency<sup>b</sup> Frequency per 10,000 words

As shown in Table 10, in student writing, *also* was used most frequently (23.65), the next four are *in addition* (3.60), *besides* (3.60), *for instance/example* (3.45), and *moreover* (2.66). Similarly, textbook lessons had *also* (20.99) as the top one transitional word in additive category; the following were for *instance/example*, *in fact*, *in addition*, and *moreover*. The ranking of additive transitional words in both corpora display close resemblance.

Table 11. *Top ten adversative transitional words in both students and textbook corpora*

Rank	Students' W.	raw <sup>a</sup>	ptw <sup>b</sup>	Textbook lessons	raw <sup>a</sup>	ptw <sup>b</sup>
1	however	114	17.85	However	198	13.81
2	nevertheless	23	3.60	Instead	43	3.00
3	on the other hand	12	1.88	at least	33	2.30
4	despite this	11	1.72	on the other hand	28	1.95
5	instead	11	1.72	despite this	19	1.32
6	in (by contrast)	7	1.10	Nevertheless	15	1.05
7	meanwhile	7	1.10	in (by contrast)	11	0.77
8	at least	7	1.10	Otherwise	7	0.49
9	Nonetheless	4	0.63	Meanwhile	3	0.21
10	On the contrary	3	0.47	in the meantime	2	0.14

<sup>a</sup> raw frequency<sup>b</sup> Frequency per 10,000 words

Table 11 showed that the highest rank of adversative transitional words in student writing is *however* (17.85), the following are *nevertheless* (3.60), *on the other hand* (1.88), *despite this* (1.72), and *instead* (1.72). Interestingly, textbook lessons also presented similar ranking, *however* as the first, then followed by *instead*, *at least*, *on the other hand* and *despite this*.

Table 12. Top ten Causal transitional words in the two corpora

Rank	Students' W.	raw <sup>a</sup>	ptw <sup>b</sup>	Textbook lessons	raw <sup>a</sup>	ptw <sup>b</sup>
1	Thus	42	6.58	then	235	16.39
2	therefore	35	5.48	Thus	57	3.97
3	Then	20	3.13	Therefore	40	2.79
4	As a result	18	2.82	As a result	34	2.37
5	Consequently	6	0.94	Consequently	6	0.42
6	Hence	6	0.94	Hence	5	0.35
7	In turn	4	0.63	In turn	2	0.14
8	Accordingly	2	0.31	For that reason	1	0.07
9	In consequence	1	0.16	accordingly	1	0.07

<sup>a</sup> raw frequency<sup>b</sup> Frequency per 10,000 words

The top five causal transitional words in the two corpora are identical, as shown in Table 12, with only slightly different order between them; the top five items in students writing are *thus* (6.58), *therefore* (5.48), *then* (3.13), *as a result* (2.82), and *consequently* (0.94), while the top five in textbook lessons are *then* (16.39), *thus* (3.97), *therefore* (2.79), *as a result* (2.37) and *consequently* (0.42).

Table 13. Top ten Temporal transitional words in the corpora

Rank	Students' W.	raw <sup>a</sup>	ptw <sup>b</sup>	Textbook	raw <sup>a</sup>	ptw <sup>b</sup>
1	Finally	43	6.73	Next	90	6.28
2	First/firstly	20	3.13	Later	88	6.14
3	Second/secondly	20	3.13	Finally	54	3.77
4	Eventually	18	2.82	Eventually	27	1.88
5	First of all	14	2.19	First/firstly	26	1.81
6	Later	11	1.72	Second/secondly	16	1.12
7	Next	6	0.94	At last	10	0.70
8	All in all	6	0.94	Third/thirdly	8	0.56
9	Initially	5	0.78	First of all	4	0.28
10	To sum up	5	0.78	In short	2	0.14

<sup>a</sup> raw frequency<sup>b</sup> Frequency per 10,000 words

Like the results of ranking in other categories, the temporal transitional words in student writing display highly resembling order as those in textbook lessons (see Table 13); the top five of the former are *finally* (6.73), *first/firstly* (3.13), *second/secondly* (3.13), *eventually* (2.82), and *first of all* (2.19), while the first to the fifth ranking of the latter are *next* (6.28), *later* (6.14), *finally* (3.77), *eventually* (1.88), and *first/firstly* (1.81).

As discussed in above-mentioned results, the frequency of transitional word (or words) per 10,000 words in each category present highly similar distribution in the two corpora (see Table 10). These results suggested that Taiwanese third-year senior high school students tend to utilize transitional words that they learn in the textbooks in their English writing. This result echoed Leedham and Cai's (2013) and Liu's finding (2008) in which they claimed Chinese EFL students' use of linking adverbial is deeply influenced by the textbooks they learned during senior school days. Some transitional words are particularly overused. Apparently, Taiwanese EFL high school students show similar tendency in their English writing to those English learner of L1 Chinese.

In view of the significant influence that English textbook can have on Taiwanese advanced learners in respect of transitional words, the result obtained in this study can be useful for less-proficient senior high school students who intend to attend college entrance examination. Generally speaking, ordinary Taiwanese senior high school students, unlike high-proficiency learners, do not possess large vocabulary size that allows them to produce better writing (Chang, 2006). It takes much effort for average-level students to accumulate sufficient word power to do their writing. Nevertheless, as Peha (2016) described, good writing uses just the right words to say just the right things. As long as used skillfully and logically, the first 2000 words, as indicated in Table 7, are enough for average-level English learners to produce conventionally acceptable writing (Ting, 2011).

Since transitional words make important part in English writing, among the four types of transitional words, additive transitional words usually enable a writer to lengthen the sentence, and therefore are likely to make the writer's idea understood easily. Learning how to make use of additive transitional words to present an intended idea does not require complicated and unusual words. It is worth the effort for ordinary senior high school students to spend time learning to use additive transitional words in their writing as early as possible.

## Conclusion

Based on the data analysis and discussion of the study result, three major findings with the three research questions are summarized as follows.

First, in respect of vocabulary distribution, students writing performance in the college entrance examination is that the first 2,000 words occupied 87 per cent of each standard composition, and the academic words have a percentage of 3 per cent. However, there are significant differences between students writing in GAST and AST in respect of academic-word level and average length of writing (see Table 5).

Second, there is significant correlation among the four commercialized textbooks in terms of vocabulary distribution, transitional word use. All the four commercialized English textbooks edited their English textbooks in a gradual level-raising principle in accordance to the 2010 version of the Senior High School Curriculum Guidelines (see Table 4). Averagely, the first 2,000 high frequency words have occupied 83 per cent of a standard lesson in Book V and Book VI, which are closest to Taiwanese senior high school students' current English proficiency level. Academic words only have a lower percentage in each lesson, 4 percent (see Table 6).

Third, in terms of vocabulary distribution, the students writing samples and the articles in the four textbooks showed significant differences in regard to first 1000-word level, academic-word level, and lexical density (see Table 7). As for transitional words use, the four major categories, including additive, adversative, causal, and temporal, displayed similar results either in the student writing samples or in the textbooks. Additive transitional words have the highest percentage of use in both student writing and textbook lessons, compared to the other four categories (see Table 9).

The finding of this study can provide textbook publishers and English teachers with some pedagogical applications. In view of the close connection between senior high school students' writing performance and their English textbooks, it is suggested that genre-specific reading materials, such as scientific research papers for natural science-oriented students, should be edited and offered to the Taiwanese senior high school students, particularly twelfth graders, so that students can learn how to make use of proper academic words in field-oriented writing. On the other hand, the use of transitional words, additive ones in particular, should be instructed formally in general English textbooks for senior high school students as early as possible. Those most frequently used transitional words can be taught to less-proficient Taiwanese EFL learners, since they are not difficult to learn, and they can contribute to better writing quality, even though for writers without large vocabulary size.

These years, Taiwanese education authorities concerned have called for providing more elective courses for high school students in accordance with the current educational policy of 12-year compulsory education, which emphasize student learning not focused on united curriculum syllabus, but on individualized curriculum design. To stand up for the goal of current English education policy, it is appropriate to design elective English writing courses which incorporate reading of academic research papers for advance-leveled senior high school students. For those who intend to develop better English writing competence in their college-level learning, they do need to read more genre-specific academic articles with a view to familiarizing conventions of English writing and facilitating necessary English writing skills. It will never too early for advanced Taiwanese EFL learners to read academic articles in senior high school. After all, good writing skills must come from observing good articles, no matter what field a college student majors in.

## References

- CEEC, (2010). *Senior High School Curriculum Guidelines*. Retrieved August 1, 2016, from [http://www.ceec.edu.tw/CeecEnglishWeb/E07Process\\_GSAT\\_02.aspx](http://www.ceec.edu.tw/CeecEnglishWeb/E07Process_GSAT_02.aspx)
- Celce-Murcia, M. & Larsen-Freeman, D. (1999). *The Grammar Book: An ESL/EFL Teacher's Course* (2nd ed.). Boston: Heinle & Heinle Publishers.
- Chang, W. (2006). English Language Education in Taiwan: A Comprehensive Survey. *Educational Resources and Research*, 69, 129-144.
- Chen, C. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1), 113-130.
- Chen, M. (2008). *From learning to production: a case study of vocabulary use of senior high school students in Taiwan*. Unpublished master thesis, National Chengchi University.
- Cheng, Y. (2006). *The effects of writing anxiety on writing performance and self-efficacy*. Paper presented at Asia TEFL 2006 international conference, Fukuoka, Japan.
- Coxhead, A. (1998). *An Academic word list, Occasion Publication Number 18, LALS*, Victoria University of Wellington, New Zealand.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Douglas, S. (2013). The breadth of undergraduate novice level writing competency. *The Canadian Journal of Applied Linguistics*, 16(1), 152-170.
- Eherlich, S. (1988). Cohesive devices and discourse competence. *World English*, 7(2), 111-118.
- Faigley, L., Daly, J. A., Witte, S. P. (1981). The role of writing apprehension in writing performance and competence. *Journal of Educational Research*, 75, 16-21.
- Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. Harlow: Longman Group UK Limited.
- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4(2), 237-258.
- Koda, K. (1993). Task-induced variability in FL composition: language-specific perspectives. *Foreign Language Annals*, 26(3), 332-346.
- Kučera, H., Francis, W. (1967). *Computational analysis of presentday American English*. Providence, RI: Brown University Press.
- Laufer, B. (1994). The lexical profile of second language writing: does it change over time?. *RELC Journal*, 25(1), 21-33.

- Laufer, B.; Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Leedham, M.; Cai, G. (2013). Besides . . . on the other hand: Using a corpus approach to explore the influence of teaching materials on Chinese students' use of linking adverbials. *Journal of Second Language Writing*, 22(4), 374-389.
- Liu, D. (2008). Linking adverbials: An across-register corpus study and its implication. *International Journal of Corpus Linguistics*, 13(4), 491-518.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Peha, S. (2016). *What is good writing?* Retrieved Aug 10, 2016, from [https://www.ttms.org/writing\\_quality/writing\\_quality.htm](https://www.ttms.org/writing_quality/writing_quality.htm).
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge Press.
- Spycher, P. (2007). Academic writing of adolescent English learners: Learning to use "although". *Journal of Second Language Writing*, 16, 238-254.
- Sugita, Yoshihito. (2012). Enhancing students' fluency in writing: Learning to use transition words. *Open Journal of Modern Linguistics*, 1(1), 18-25.
- Tsai, P., Cheng, Y. (2009). The effects of rhetorical task type, English proficiency, and writing anxiety on senior high school students' writing performance. *English Teaching and Learning*, 33(3), 95-131.
- Ting, H. J. (2011). A study on Taiwanese high school students' English reading awareness, vocabulary knowledge and their writing performance. *Proceedings of Canada International conference on Education*. (pp. 631-636). Toronto: Infonomics Society.
- West, M. (1953). *A general service list of English Words*. London: Longman.
- White, GB (2015, November). *Taiwanese students' transition markers use with comparisons to SSCI Articles*. Paper presented at The 24th International Symposium and Book Fair on English Teaching, Taipei City, Taiwan.

投稿日期：2017/08/10 接受日期：2017/12/08

## Appendix A

### 1. Additive

*Emphatic*: in addition, additionally, moreover, furthermore, besides, also, in fact, at the same time

*Apposition*: that is, that is to say, in other words, for instance/example

*Comparative*: likewise, similarly

### 2. Adversative

*Proper adversative*: however, nevertheless, nonetheless, despite this, in (by contrast)

*Contrastive*: conversely, in fact, on the other hand, at the same time, in the meantime, meanwhile, otherwise

*Correction*: instead, rather, on the contrary, at least

*Dismissal*: in any case/event, anyhow, at any rate

### 3. Causal

*General causal*: therefore, consequently, for that reason, thus, as a result (consequence), hence, thereby, accordingly, in consequence

*Causal conditional*: then, in that case, in turn

### 4. Temporal

*Sequential*: then, in turn, next, first, second, third, fourth (fifth...), first of all, firstly (secondly...), last, finally, later, initially

*Summarizing*: in short, in summary, in brief, in sum, in conclusion, to sum up, to conclude, to summarize, overall, all in all



# **A Comparison of Excellent Taiwanese EFL Learners’ English Writing in College Entrance Examinations and Their English Textbook Lessons in Respect of Vocabulary Distribution and Transitional Word Use**

Hung-Ju Ting<sup>1</sup>

## **Abstract**

This study aimed to investigate the influences that English textbooks may have on Taiwanese EFL senior high students’ writing performance in the college entrance examination. This study adopted corpus analysis methodology. To build up reliable corpora, 210 pieces of students written compositions in GAST and AST, issued by the CEEC (2009 to 2016), were collected and transcribed into electronic files. On the other hand, 270 articles in the current four commercialized English textbooks, Volume I through Volume VI, were collected and edited as a corpus. In this way, vocabulary distribution and transitional words used in the two corpora could be analyzed and compared. The major findings are as follows. First, there are significant differences between students writing in GAST and AST in respect of academic-word level and average length of writing. Second, there is significant correlation among the four commercialized textbooks in terms of vocabulary distribution, transitional words. Third, with respect to vocabulary distribution, the students’ writing samples and the articles in the four textbooks showed significant differences in regard to first 1000-word level, academic-word level, and lexical density. As for the usage of transitional words, the four major categories, including additive, adversative, causal, and temporal, displayed similar results either in the students’ writing samples or in the textbooks. Based on the above findings, this study proposed some principles for further textbook compilation. Some suggestions are also made to the Taiwanese English education authorities concerned as regards their design of curricular guideline for 12-year compulsory education.

**Keywords: vocabulary distribution, transitional word, English writing, English textbook**

---

<sup>1</sup> English Teacher, Private Tatung Senior High School

Corresponding Author: Hung-Ju Ting, E-mail: adamting@ms.tsh.ttu.edu.tw