

Exploring Source Use by Different EFL Students: Text Quality, Lexical Profiles, and Citation Features

Hsien-Chin LIOU

Feng Chia University, Taiwan

The study examined to which extent EFL students' language proficiency is related to text quality, lexical profile, and source use (specifically, the amount and type of textual borrowing) in an integrated summary and response task. Student essays from 42 EFL second-year college English-majoring students using a source text were analyzed with identification of summary, paraphrase, reformation of original source text strings, and verbatim copying which were measured with the exact string length of each citation. These essays came from a more proficient group (N=14) and a less proficient group (N=28). Besides documenting citation features, the text quality was also analyzed with their lexical frequency profiles compiled using the VocabProfile tool. Findings indicate that the more proficient group wrote longer essays with better rated quality and more complex vocabulary. They also demonstrated a slightly better skill in incorporating the source text into their essays. It seems that good source using skills required in such a summary-response genre grow with increase of the general language proficiency which involves more sophisticated word use and better and more fluent writing performance. However, a variety of misattribution problems were found in the two groups. Pedagogical implications and future research are discussed against relevant previous studies.

Introduction

Common in academic genre, source use has become an essential writing task for graduate students, and a common writing skill for undergraduate foreign language departments to train their learners in several ESL (e.g., Wette, 2010; Zhang, 2013) or EFL contexts (e.g., Chen & Su, 2012; Hu

& Lei, 2012; McDonough, Crawford, & Vleeschauwer, 2014; Yeh & Liou, 2000). Most of the previous studies have provided information on ESL faculty's and the students' perceptions as well as textual patterns of imperfect student writing and misappropriation of sources, or triangulation of the two perspectives (Pennington, 2010; Polio & Shi, 2012; Shaw & Pecorari, 2013). Yet, the EFL student population has been under-represented in this line of inquiry. Reasonably, treatment studies would be emphasized (Chen & Su, 2012; Hu & Lei, 2012; McDonough, Crawford, & Vleeschauwer, 2014) as instruction is essential to provide necessary input and also guidance for EFL learners to imitate and master the genre in a source-using writing environment. As little do we know how these learners acquire proper attribution skills, descriptive studies particularly on learners with various L2 proficiency levels can provide understanding on how they differ and how much the EFL learners of lower proficiency should make up in order to reach those of higher proficiency who show proper source-using skills.

Source-using writing tasks require the ability to control and create the target L2 language by integrating source texts with adequate general reading and writing abilities (Plakans, 2009). Due to a lack of L2 proficiency and training in using sources, many EFL student-writers tend to demonstrate wide varieties of misattribution phenomena, and the fact leads to a great amount of difficulty and challenges in the class (Chen & Su, 2012; McDonough, Crawford, & Vleeschauwer, 2014; Yeh & Liou, 2000). The purpose of this study is to examine how different proficiency levels are related to EFL undergraduate students' source use, a feature in academic writing, and their overall text quality and lexical complexity, features of general writing. By highlighting their unique problems in a common reading-based academic writing task, their needs in training proper attribution can be adequately addressed in the EFL learning context where limited exposure to source using texts but demand on writing such a genre is evident. Data on different source using problems by students of cross-proficiency levels can also further our understanding of different EFL learners' glitches with academic writing in order to devise proper instructional intervention to achieve their writing goals.

Relevant Studies

Factors on the L2 student-writers themselves play a key role in shaping their final source-using papers. L1 language backgrounds and the students'

proficiency levels have been examined in the past ESL literature. The contrast of source use by non-native speaking (NNS) and native-speaking (NS) groups indicates that mostly NS students demonstrated better attribution features in writing than NNS ones (Campbell, 1990; Keck, 2006; Shi, 2004). From another perspective into learners' background as related to writing performance, score levels obtained from testing settings match with different linguistic features such as source use (Gebril & Plakans, 2013; Plakans & Gebril, 2013; Weigle & Parker, 2012). Source use was also found to be significantly co-related with academic attainment (Donohue & Erling, 2012) where course grades were involved.

With a focus on linguistic analysis of L2 undergraduate students' source-using texts, the following seven studies (ESL=5, and EFL=2) as summarized in Table 1 may provide an initial overview of discourse features in student writing. Campbell (1990) examined how ten native-speaking and twenty non-native speaking English student-writers used background knowledge to write a source-based article. Although the NS students were found to write better than NNS students, all of the students did not make proper attribution of the original writing. The NNS students were also found to use quotations for attributions most often. Aiming to reveal task effects, Shi (2004) compared two groups (an ESL and an NS) of student essays concerning textual borrowing on two writing tasks: summary or opinion. Data analysis involved identifying exact or near verbatim retention of strings of words from sources with or without references (a string can be a word, several words, a sentence or a sentence plus words). The summary task was found to elicit more borrowed words than the opinion task, and the ESL group documented the source less properly than the NS group. In a series of studies, Keck (2006, 2010, 2014) analyzed why and how undergraduate students of both L1 and L2 backgrounds paraphrased a source for summary writing. Based upon the precise ranges of amount of source texts borrowed from the source information and detailed linguistic characteristics, Keck (2006) found L2 students made more near copies while their NS students used more moderate or substantial revisions. To further examine their grammatical strategies adopted to avoid plagiarism, Keck (2010) found that both groups were found to make changes on components of the subject, main verb, or the direct object category of the original excerpts. More recently, Keck (2014) found that although the L2 group copied more, the finding came from only a small portion of the students.

Three studies examined the relationships between ESL or EFL students’ proficiency levels and their source use in testing settings. Weigle and Parker (2012) collected student texts from an institutional testing context, transcribed them verbatim, and analyzed them for source borrowing (adapting Shi [2004]’s scheme). Only 9% of the students borrowed extensively from the source texts. Minor differences in patterns of borrowing across topics, student groups (undergraduate and graduate), and various proficiency levels were located. Gebril and Plakans (2013) found moderate relations between source use and proficiency levels where their ESL lower scorers did not use as many indirect formats (summary and paraphrase) as higher scorers. Adding an additional source with a different modality, Plakans and Gebril (2013) examined how TOEFL test-takers of both ESL and EFL students incorporated a reading text and a lecture into their writing. Higher scorers relied on the listening text and incorporated less copying of words and phrases than lower scorers. Source use explained over 50% of the score variance. Using texts on the Web as sources, Yeh and Liou (2000) investigated how the English-major EFL college students composed research papers. They found that when citing articles to write, better readers also wrote better papers. Analyses of the students’ citation strategies showed that half of the students had difficulty using the correct citation format, and direct quotation was the most frequently used citation format.

Table 1. Summary of Relevant Studies on Textual Description of Student Writing

Relevant Studies/ Features	Task Difference/ Setting	Source Use	Length (Words)	Text Quality	Lexical Density
Campbell, 1990 (2 NNS/ESL groups vs. 1 NS group, <i>n</i> =30, 1o in each group)	1 source as background, 1-hour timed writing/ in class	Averagely 5.7 incidents of text borrowing per essay; NNS cited more	NA	NS better (essay NA scores in 3 groups: *87 vs. 76.6 vs. 74.1/100)	NA
Shi, 2004 (NNS/ ESL vs. NS, <i>n</i> =87)	2 sources, summary or opinion task/ Outside class	(1) NNS borrowed more in summaries than opinions; (2) NS performed better with shorter strings of borrowed texts with references.	NS (137–348 words), NNS (272–295)	NA	NA

(Cont'd Table 1)

Relevant Studies/ Features	Task Difference/ Setting	Source Use	Length (Words)	Text Quality	Lexical Density
Keck, 2006 (NNE/ ESL vs. NS, <i>n</i> =153)	1 source, 45-min 1-paragraph summary	Borrowed 5 paraphrases per summary (45%); L2 writers used significantly more Near Copies than NS writers; NS writers, more Moderate and Substantial Revisions.	NS (176 words/ summary) yielded longer texts than NNS (151 words)	NA	NA
Weigle & Parker, 2012 (score levels of ESL, <i>n</i> =63)	1 source, Writing section of an in-house proficiency test/testing; argumentative type	(1) Students borrowed 9% from the source; (2) no clear patterns were observed across various comparisons, given their background factors	NA (instruction given on word limit, 200–300 words)	4 score bands	NA
Gebriel & Plakans, 2013 (score levels of ESL, <i>n</i> =136)	2 sources as support, a 1-hour timed argumentative writing test/testing	(1) Moderate relation between proficiency and source use amount; (2) Lower scorers used less indirect source use (paraphrase and summary) but no difference was found on quotation use across students of three levels.	134–285 words/ essay, length increased with proficiency level	6-level (TOEFL-iBT holistic integrated writing scale); 3 proficiency levels	Average word length as lexical sophistication; no difference found among score levels
Plakans & Gebriel, 2013 (score levels of both ESL/EFL, <i>n</i> =480)	TOEFL-iBT integrated writing section with 2 sources (lecture and reading)/ Testing	(1) Students' source use was very closely correlated to inclusion of ideas in lectures, and choosing important ideas, but negatively correlated with ideas in the reading and verbatim copying. (2) Students who were copying did not include important ideas.	Text length increased across levels of writing proficiency	9 levels (1.0–5.0)	NA
Yeh and Liou, 2000 (EFL, 1 group, <i>n</i> =15)	Multiple-source research paper/ Outside class	(1) 11.2 incidents averagely (2) 90% of the citations were relevant	4–6 pages (48.6 T-units averagely)	*74.7/ 100	NA

Note. NA: not applicable or data not provided; * both rated using ESL Composition Profile (Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey, 1981); N=total students included in the study.

From the review of the seven studies, types of citation including summary, paraphrase, and quotation, and their amounts seem to fluctuate with different task demand and proficiency levels. Where there were patterns to be found, quotation use was associated with students of lower proficiency or L2 background (Campbell, 1990; Keck, 2006; Plakans & Gebriel, 2013; Shi, 2004). L1 students wrote better and more proficient L2 students also composed better source-using papers with more proper attribution. Lexical aspects were much less documented across the seven studies. Laufer and Nation (1995) show that a lexical frequency profile with ratios of general service words and academic words can discriminate ESL learners' writing among texts of different proficiency levels. Evidence was inconclusive as Astika (1993) and Cumming et al. (2005) demonstrated the significant component of lexical measures in student writing, but Baba (2009) did not support this position. Recent computational tools can help with the analysis of specific linguistic features of student texts (e.g., Adel, 2014; Cobb, 2007).

Most students in EFL contexts have less access to source-using texts or writing demand and thus fewer opportunities for practice, compared with those in ESL settings. However, with the spread of English as a lingua franca for global communication, English departments in EFL countries start to emphasize training of source-using tasks. Compared with EFL graduate students facing the pressure of scholarly publications, source use by the undergraduate students have been much less investigated. The literature review indicates a gap in documenting the source use by EFL undergraduate students with different proficiency levels. If we believe the acquisition of source use in academic English writing needs to start from EFL undergraduate education and L2 proficiency is an important mediating factor for the text quality, then a study documenting these issues—both source use for academic writing as well as other features for general writing—is needed. The present descriptive study attempts to respond to the issues by addressing to what extent textual borrowing is related to the EFL students' proficiency levels, lexical complexity and text quality.

Methods

Participants

Forty-two second-year students were recruited from the English departments of two universities in an Asian EFL country where the

learners' first language was Chinese: seven were male, thirty-five, female. On average, they were 21 years old, and had studied English as a school subject for eleven years in this country, when they provided the writing samples. Those from university A (N=14, named as Group A here) happened to show a higher English proficiency level with an average TOIEC score of 870, whereas the remaining students were from university B (N=28, named as Group B) belonged to a less proficient level (average TOIEC score=637). They all signed a consent form of giving permission to use their essays for research purpose before the study started. Both writing classes in the two universities focused on expository writing, where their curricula both built upon the students' first-year writing programs, and were to be followed by their third-year programs, all of the three being required credits in the two departments (12 credits in their 4-year college education). Exposure to source using writing may be available in other content courses such as British literature, novel reading, etc.; however, formal writing units on source use were not arranged nor given in their first-year writing programs.

Task, Material, and Procedure

One narrative source text "Darkness at Noon" (text by Harold Krents, cited in Wyrick, 2007, pp. 529-530) from a textbook chosen by the course instructor of both groups was required to read, before all the students came to class to work on a writing task. The text has 824 words with a Flesch-Kincaid ease score of 46 (out of 100), and grade level, 11.3. This represents a text written by a native speaker. Using one text allowed juxtaposing the students' texts with the source in order to identify incidents of textual borrowing. Sixty minutes were given for the students to do a summary-and-response writing task, and students were allowed to read the source while composing their essays. None of our data was an incomplete text so the length of time given was perhaps enough to show the students' "understanding and practices of using source texts" (cf., Shi, 2004, p.177). Such a summary-and-response writing task was common in both English programs, with its various purposes of task assignments. We also confirmed with writing class instructors of the two groups about the suitability for the task to their students.

Data Analysis and Related Instruments

We collected forty-two student essays which came to a total of 11,445 running words as our corpus. Data analysis procedures involved categorizing textual borrowing, and assessing text quality and lexical complexity (using lexical frequency profiles). The categorization provides linguistic perspectives on how different learners borrow text segments from the source, and the assessment gives numerical measures for essay quality and its related vocabulary use.

To categorize kinds of textual borrowing in student essays, by following Shi's (2004) coding scheme as a starting point, we took the sentence as the unit of analysis to locate whether borrowed texts were referenced or not. In Shi's scheme (pp. 196–197), her seven codes are classified into verbatim copying from the source text, modification (substitution of synonyms or deletion of words), or reformulation (changes onto the phrase or sentence structure). Different from Shi's scheme, incidents of paraphrase and summary coded in Campbell (1990) were added into our scheme where paraphrase was defined as a re-written sentence from the original source text with no direct borrowing of two or three consecutive words. Summary represented the gist of information from the source text which condensed borrowed text segments of more than one sentence. The seven codes represent a progressive degree of heavy reliance on the source linguistically to having none; however, as patchwriting (explained later) is involved in various cases of the seven, and thus different degrees of reliance are not completely clear-cut.

As only one article was referenced in the task, we found some students used "the article," "the author," or "he" as referents, instead of using "Harold Krents" (the author's name) or "Darkness at Noon" for proper referencing. Strictly speaking, those incidents cannot be considered as proper attribution; however, it may be reasonable to name them as an intermediate, or semi-referenced category for these developing writers. They were shown in the cases of verbatim copying, modification, paraphrase, and summary (again, moving from complete reliance on the source to less).

After several rounds of discussions and revisions with our research assistants (one post-doctoral fellow holding a Ph.D.-TEFL degree and another, a TEFL-MA graduate student), we developed our coding scheme by merging the two used by Campbell (1990) and Shi (2004)

with adaptation, because such a combined scheme fits our data. Altogether, the scheme included thirteen codes under three major categories: referenced, semi-referenced, or non-referenced types of source use (see Appendix A for the codes and examples). Under each of the three major categories, how and how much text amount of references were further categorized: those with reference to the author or the source text (5 codes: RC, RM, RR, RP, and RS), those with only common nouns as referents (4 codes: xNC, xNM, xNP, and xNS), those without references (3 codes: NC, NP, and NS), and one with quotation (1 code: QC). To account for the amount of textual borrowing, a word in student essays was taken as the unit of analysis. For the incidents of borrowing exact words from the source, the number of every word was included in sub-categories of verbatim copying (NC, xNC, RC, or QC), modification of words (xNM or RM), and reformulation (RR). For the incidents of summary (NS, xNS, or RS) and paraphrase (NP, xNP, or RP), total words in the entire sentence in the student essay were counted, as they demonstrated borrowing ideas from the source.

Patchwriting was defined as a textual borrowed sentence which showed a summary, paraphrase, modification, or reformulation together with verbatim copying. It was coded and counted with the main function plus some verbatim copying, e.g. RS+(RC) or xNP+xNC (see Appendix A for illustration). Two research assistants and the author worked together to identify incidents of all thirteen codes and patchwriting. Because all 42 essays had different lengths, the percentages of words of the thirteen types of textual borrowing in each essay were tallied by means of dividing them by the total number of words in each essay (in the unit of 100 running words).

Due to the complex nature of source use, developing a scoring rubric with key descriptors that can delineate distinct levels of source use was conducted as recommended by Plakans and Gebriel (2013). For assessing the timed summary-and-response task, two raters who both held a doctoral degree in TESOL first rated ten percent of the student essays using the Integrated Writing Rubric for TOEFL iBT test¹ as a starting point. They then worked together to revise the rubric into one version which fit our data better (see Appendix B for the complete rubric). The new rubric had ranks ranging from 0 to 5, which considered both the effectiveness of source use and the general text quality. The descriptors in the rubric pointed to (a) citation styles of explicit or

implicit format, (b) the connectedness of the source article with the students’ responses concerning their content, and (c) the quality of writing concerning organization, language use, and precise presentation of the content (related to the source comprehension).

Related to essay quality is how the students used vocabulary. To obtain the lexical frequency profile of all the essays, the vocabulary profiles of the two groups were compiled using the Compleat Textutor *VocabProfile* v.4 tool (Cobb, 2007, <http://www.lex tutor.ca>). *VocabProfile* yields percentages of the common 1,000 word-band, the 2,000 word-band, the academic word-band, type-token ratio, and other measures, given any input English text. The type-token ratio (total different words divided by total running words) is a measure of word richness (Laufer & Nation, 1995) where a higher value means more different words (word type, not repetitions of the same word) are used in each essay, meaning more sophisticated writing in the situation.

Results

Length, Lexical Profile, and Text Quality

As the amount of the corpus analyzed is important to know the nature of the data in our study, the length of 42 students’ essays and their average word counts are given in Table 2.

Table 2. Participants’ Groups and Lengths of Writing (in words)

Level	N	Mean of Essay Length	SD	Total Words
Group A	14	322.79	65.07	4,519
Group B	28	247.36	51.76	6,926

Note. $T=4.01$ (comparison of two group means); $p<.01$.

On average, each of the 42 essays is 272.5 words (altogether amount to a total of 11,445) where the mean essay length of the more proficient group is 322.8 words long and the less proficient group, 247.4. Mean differences of essay lengths were tested using the statistical independent T-test and it was found the essay lengths of the two groups differed significantly ($T=4.01$, $p<0.01$). As expected, the more proficient group

wrote much longer essays than the less proficient group. The finding confirmed to those in Cumming et al. (2005), Gebril and Plakans (2009, 2013), and Plakans and Gebril (2013) where they found higher score levels matched with higher writing fluency.

To understand lexical complexity in the corpus, Table 3 shows the contrasts of word bands from the output of *VocabProfile* in the two groups and the source text. The two groups did not make significant differences in terms of the common 1,000 or 2,000 words; however, the more proficient group used more academic words than the less proficient group (4.60% vs. 3.23%, mean difference=1.37%). Similarly their type-token ratios were also slightly higher (0.23% vs. 0.17%; mean differences=0.06%, 26.09% higher). We wondered if the academic words originally in the source text had been appropriated by students while they were completing the task. We compared all the academic words in the essays with those in the source text using the unit of word families. As a stricter standard than any running word in a text, a word family includes all its derived words under one-word root so *assumption*, *assumes*, *assuming*, *assumed* are counted as one word (family), *assume*, for instance. Twenty-eight word families were identified in the source text by the native speaker. Both groups used many more academic words than the 28 in the single source text. In Group A and Group B, 112 and 65 word families were found. Group A cited more academic words originally in the source for integrating them into their own writing than Group B, with 1.37% difference between the two groups' entire vocabulary profiles as shown in Table 3. The percentage difference of 1.37% means a difference of 47 academic word families between the two groups (112 minus 65). Meanwhile, the proportions of source-based academic words vs. non-source-based words in both groups were similar (non-source categories in the 2 groups: 77.70% vs. 76.99%). These mean that Group A knew and could use many more academic words than those originally in the source text. The past literature did not find a consistent pattern of source using writing as related to lexical richness perhaps due to different measures adopted in those studies (Cumming et al., 2005; Gebril & Plakans, n.d.; Plakans & Gebril, 2013). However, findings from the output of *VocabProfile* indicate that our more proficient group used a more complex lexical frequency profile in terms of more academic words and a slightly higher type-token ratio (word density, or different word types).

Table 3. Word Band Differences between the Two Groups (%)

Word Bands	Source Text	Group A	Group B	Group Difference (A-B) %	Expected or Not
K1 Words (1–1,000):	81.48	85.29	84.60	0.69	-
K2 Words (1,001–2,000):	5.74	5.12	5.76	-0.69	-
1k+2k	87.22	90.41	90.60	-0.19	-
AWL Words (academic):	<u>3.87</u>	<u>4.60</u>	<u>3.23</u>	<u>1.37</u>	V
Off-List Words:	8.91	4.99	6.41	-1.32	-
Type-Token Ratio	0.46	<u>0.23</u>	<u>0.17</u>	0.06 (26.09%)	V

Note. V indicated the comparison was expected as Group A was better than Group B; (A-B) %: 26.09%=0.06/0.23.

Writing teachers would like to know how well the two groups composed, given the task to show their source use. In the present study, quality of the essays came from the average scores of two raters who evaluated student writing based on the rubric with the inter-rater reliability of 0.89. The differences of score means between the two groups were examined using the statistical procedure, an independent T-Test. Without surprise, Group A wrote better essays than Group B (two group means: 3.79 vs. 2.86 out of 5.00, $T=3.0204$, $p<0.01$; if both group means converted to percentage, 3.79 divided by 5.00=75.8/100 vs. 57.2). The range of ratings in Group A was 2 to 5 out of a total of 5.00, while those in Group B, 1 to 5. Two lowest scores, 1.00, fell in the less proficient group members.

Textual Analysis of Source Use

Linguistic analysis on student texts allows writing researchers and teachers to obtain good understanding of the participants’ source use in detail. Coding and calculation of textual borrowing incidents concerning the 42 student essays yielded a total of 166 strings of source using text segments (3.95 sentences per essay) with 2,627 words in total, 22.95% out of total words (11,445). The proportion was much higher than that (9%) in Weigle and Parker (2012) but lower than those of Campbell (1990) with 5.7 incidents and Keck (2006) with 5 paraphrases in each student’s summary, where ESL students were all

involved. In the current study, 17 patchwriting incidents (one non-copying code with some incident of verbatim copying: NC, xNC, RC, or QC) were found, which amounted to 10.24% (17/166) of all the textual borrowing incidents. Seventy percent of them came from Group B (see details in Appendix C). Concerning the results of kinds and amounts of textual borrowing, high variations were evident as shown in its dispersion of as large as 12.10 incidents (SD of Group B concerning the semi-referenced summary category, xNS) with a mean of only 7.85, highest among all individual sub-categories. This suggests that incidents of source use in all the participants' writing fluctuated greatly, like the findings in Keck (2014).

As shown in Table 4, across the three major categories of referencing format, altogether both groups showed mainly referenced (41.89%) and semi-referenced incidents (52.09%), compared with those undocumented (with no reference, 6.02%). Given the variant forms distributed in both the undocumented and semi-referenced incidents of both groups (close to 60%), more instruction efforts are sorely needed in these two areas, so is instruction for patchwriting.

Respectively, Group B showed more referenced (11.06 vs. 8.48 of Group A) and semi-referenced incidents (13.73 vs. 10.57), and also the overall textual borrowing amount (25.70, 55.09%), 10% more than Group A (20.95, 44.91%). In spite of minor differences of non-referenced incidents between the two groups (less than 1.00, 1.90 vs. 0.91) in terms of raw numbers, Group A (59.52%), being more proficient, used a slightly higher proportion of non-referenced plus semi-referenced cases (2.56% difference) than Group B (56.96%) within their own total textual borrowing incidents. Non-referenced and semi-referenced cases were regarded as improper attribution, compared with referenced cases. Thus, Group B might have a slightly higher awareness that borrowing others' text without references was not a good practice in writing source-using essays, although they borrowed much more concerning referenced and semi-referenced incidents in terms of raw numbers. Three students in Group B did not borrow any text or ideas from the source; if they did, the contrast would become more drastic, when compared with the situation of Group A. The contrastive amounts and kinds of textual borrowing in the two groups provided new understanding of what these less proficient EFL students may need to improve in order to catch up with the more proficient group.

Table 4. Amount and Type of Textual Borrowing of the Two Groups

	Group A (n=14)	Group B (n=28)	Sum (C=Group A+B)
(I) With No Reference	1.90	0.91	2.81 (6.02%)
Semi-referenced	10.57	13.73	24.30 (52.09%)*
Sub-total	12.47 (59.52%) [#]	14.64 (56.96%)	27.11
(II) With Reference	8.48	11.06	19.54 (41.89%)
Total	20.95 (44.91%)	25.70 (55.09%)	46.65 (grand total)

Note. * Percentage: 52.09%=24.30/46.65; [#] 59.52%=12.47/20.95.

In addition to the different results of referencing and non-referencing cases, we also found patchwriting cases where some form of verbatim copying was patched with the students’ own sentence structures (numerical findings shown in Appendix C). The breakdown of patchwriting in the two groups demonstrates that Group B seemed to be struggling far more with efforts to avoid verbatim copying than Group A as over 70% of the 17 cases came from them. Overall, both groups seemed to know using documentation (58.82%) was safer than using the common nouns as referents as shown in the illustrated sentence below, that is, using the so-called semi-referenced categories (41.18%) (see Appendix A for a text illustration of RS+(RC)). More instruction efforts are sorely needed in semi-referenced and undocumented areas so is instruction for patchwriting

Original source:
I would hope that in the near future when a plant manager is touring the factory ... his comment after watching them work will be, “Which one is disabled?”

Student’s textual borrowing of a patchwriting incident:
Just like the article says: “I would hope that one day the manager would ask which one is disabled?” (from a member in the more proficient Group A, A5, with the code of xNP (19) + xNC (8) indicating a semi-referenced paraphrase incident of 19 words long with verbatim copying of 8 words from the source)

Across the seven non-referenced and semi-referenced sub-categories, both groups relied heavily on semi-referenced summaries the most often (6.36 vs. 7.85) as illustrated in Table 5. Following the dominant use of semi-referenced summaries, whereas Group A used more semi-referenced

paraphrases (2.95 vs. 1.84) to fulfill the task demand, Group B resorted to semi-referenced verbatim copying more often (3.79 vs. 1.26). This is similar to the ESL group in Gebril and Plakans (2013) where they found indirect source use (summaries and paraphrases) was used less often by the lower scorers. Indirect source use such as paraphrase and summary requires more syntactic and lexical resources from the students, a sign of more mature proficiency, than direct copying.

Table 5. Comparison of Non-referenced Textual Borrowing

		Group A		Group B	
		M	SD	M	SD
With No Reference	Copied (NC)	<u>0.64</u>	1.15	0.20	0.52
	Paraphrased (NP)	0	0	0.47	2.13
	Summarized (NS)	1.26	3.13	0.24	<u>1.28</u>
Subtotal		1.9		0.91	
Semi-referenced	xNC	<u>1.26</u>	1.77	3.79	8.55
	xNM	0	0	0.25	1.33
	xNP	2.95	5.42	1.84	4.44
	xNS	<u>6.36</u>	8.36	<u>7.85</u>	12.10
Subtotal		10.57		13.73	
Grand Total		12.47		14.64	

Table 6. Comparison of Referenced Textual Borrowing

		Group A		Group B	
		M	SD	M	SD
With Reference	Copied (RC)	0	0	0.22	1.17
	Modified (RM)	0	0	0.65	1.75
	Reformulated (RR)	0	0	0.29	1.54
	Paraphrased (RP)	0	0	1.68	5.94
	Summarized (RS)	8.02	10.89	7.58	10.33
With Quotation	Copied (QC)	0.46	1.17	0.64	2.05
Subtotal		8.48 (40.48%)		11.06 (43.03%)	
Total Textual Borrowing		20.95		25.70	

Note. Percentage: 40.48%=8.48/20.95.

In Table 6, altogether, Group B adopted more in terms of raw numbers of most of the seven referenced categories except those of referenced summary, compared with the more proficient Group A (40.48% vs. 43.03% in Group B). This indicates there is room for Group A to learn better referencing skills because close to 60% of their textual borrowing incidents were either non-referenced or semi-referenced ones. Referenced summaries were used most often by the two groups, like the finding in the semi-referenced category in Table 5. This may be attributed to the task effect as the in-class timed writing task designed in the study required all students to summarize either one aspect or the entire article of the source text, and then respond to it. Summaries seemed to become the major sub-task in order to fulfill the goal of writing, similar to the paraphrasing study when ESL students had to paraphrase ideas or words in a source in order to complete summary writing (Keck, 2006). In our study, Group A relied on only strategies of summary and direct quotes with proper referencing, but Group B demonstrated a wide variety of strategies including direct copying, modification of the original text by substituting synonyms, reformulation of the original structure as well as paraphrasing the entire sentences, besides the two strategies used by Group A. This finding demonstrates their various attempts using available linguistic resources at their disposal in order to avoid verbatim copying as also found in the paraphrasing study of Keck (2010). Group B, perhaps due to more limited syntactic abilities, would resort to a wide variety of adding or deleting words in the source, substituting synonyms, or changing partly the original sentence structures in the source in order to construct their own texts by avoiding verbatim copying (or thus plagiarism). The contrastive findings between how the two groups improperly borrowed source texts into their own writing to fulfill task requirement provide concrete illustration of weaknesses between the two groups which can be converted into teaching materials for pedagogical intervention.

Comparison of the findings in previous ESL or EFL studies (Campbell, 1990; Shi, 2004; Wiggle & Parker, 2012; Yeh & Liou, 2000), and the current study indicates that most NNS-ESL groups, when compared with their NS groups, borrowed more non-referenced strings, like the less proficient group in the current study, when compared with the more proficient group. With or without referencing, quotations were used oftener by Group B than Group A (NC+ $\times$ NC+RC+QC, 4.85 vs. 2.36), confirming to previous findings (Campbell, 1990; Plakans & Gebril,

2013; Yeh & Liou, 2000). Although direct quotation might be regarded as a quite straightforward citation skill, compared with paraphrase or summary, the EFL students in the current study still demonstrated their problems. Students copied directly from the source by following the correct format at our first sight; however, it turned out that the entire stretch of strings was directly taken from the source where quotation marks already existed. Even doing verbatim copying, one student changed it slightly or perhaps made mistakes unintentionally such as “He said” for “He says” in the following example.

Original:

“He says he doesn’t need an interpreter,” the orderly reported to the woman (Wyrick, 2007, paragraph 14).

Student’s text:

The last sentence that the orderly speak to the old woman was “He said he doesn’t need an interpreter.” (A13, QC, direct copying with quotation marks)

In another case where the three-word string with quotes and the comma “He is blind,” already existed in the source, the student did not copy the comma with its correct location, inside of the single quotes.

Original:

“Our next-door neighbor, aged five, wandered over into our yard with a playmate. ‘He’s blind,’ our neighbor whispered to her friend in a voice that could be heard distinctly by Dad and me.” (Wyrick, 2007, paragraph 19).

Student’s text:

And then his next-door neighbor who is a little boy said “He is blind”, the little boy also told to his friend. (B20, xNS+(xNC))

Lengths of Strings of Words Borrowed

Like Shi (2004) and Weigle and Parker (2012), we also calculated how both groups used various lengths of strings of words for proper attribution or not. If more proficient students assuming having more awareness of proper attribution show less textual borrowing or borrowing in shorter strings than in longer ones, it is a sign of more mature writing performance. The lengths of borrowed strings were classified into five groups ranging from 2–3 words to 20 words or longer. Raw frequencies

of each major category were tallied with group means calculated. The shortest incident was found in the data of Group B and Group A; however, the longest incident was found in the data of Group B, the semi-referenced category, 59 words long. Both groups borrowed texts more often as the lengths of strings increased, peaking at the range of 10-19 words as shown in Appendix C. This is different from the ESL data in Shi's study (2004) and an EFL study involving instructional effects (McDonough, Crawford, & Vleeschauwer, 2014) where their students seemed to have more awareness of better attribution as the string lengths increased. Across the three main categories in our study, the semi-referenced categories showed the highest frequencies. Thus, the EFL students in this study require more instruction on proper attribution as related to linguistic-syntactic strategies.

Discussion

In this study, we looked into 42 EFL college students' discourse features in general writing (text length, text quality, and lexical frequency profiles), and those of academic writing (source use). Both groups used summaries and other strategies to document source use, but close to 60% of textual borrowing were not documented properly. Taking different L2 proficiency into account, we found that Group A, the more proficient learners, wrote longer essays than Group B, a finding similar to previous studies (Cumming et al., 2005; Gebril & Plakans, 2009, 2013; Plakans & Gebril, 2013) where the researchers confirmed fluency increased with score levels. Additionally, their lexical frequency profiles and rating of text quality were also found to be better. While the more proficient student group wrote longer and better with lexically richer essays using more academic words and higher lexical richness to show better source-using performance, both groups demonstrated similar as well as different citation features. With success or failure, these developing writers adopted a wide variety of (13) strategies in order to incorporate the source text into their writing. They borrowed up to 23% from the source text into their own essays. As observed by Plakans and Gebril (2013, p. 227), "the reading text may be a lifeline for these [lower proficient] writers, providing not just content but also wording". Close to one fourth of textual borrowing by our EFL students indicates that their use of verbatim copying, pathwriting, and semi-referenced cases of various kinds (e.g. NC, xNC, xNM+(xNC), xNP, xNS, RM(+RC), RR(+RC), RS(+RC))

were attempted to avoid plagiarism due probably to their lack of mature lexico-grammatical knowledge or source-using skills (cf. Keck, 2010).

Different from previous literature, we have found that both groups in our study similarly used a great amount of semi-referenced incidents with common nouns of “the article,” “the author,” or “he” for acknowledgement using strings of ten to nineteen words. Shorter strings were not matched with better documentation, unlike the findings in Shi (2004). These indicate more attribution awareness on nuances while working on source-using writing is sorely needed for the EFL students.

The linguistic analysis of textual patterns on students’ source use provides illuminating findings on how the two groups differed or struggled similarly with misattribution phenomena due probably to lack of mature lexical, syntactic, and knowledge in English or proper source-using knowledge in academic writing. To illustrate, the semi-referenced cases of using “the article” or “the author” may come from not knowing specific referents were needed in proper sourced writing. Systematic training on source-using writing is needed for both groups, as Group A, being more proficient in general writing and overall English proficiency, does not show fewer problems. For example, they showed more semi-referenced summaries but Group B had more semi-referenced verbatim copying. Additionally, Group A tended to use summaries and paraphrases of complete sentences, but Group B would add or delete words in the source, substitute synonyms, or change partly the original sentence structures in the source for textual borrowing cases. The lower proficient Group B borrowed 10% more than the more proficient Group A, and also adopted more incidents of verbatim copying. They also patched more and relied more on the source for the given task. They demonstrated a wide variety of misattribution glitches ranging from mechanics issues of misplacing the comma inside of the quotation mark, patchwriting cases of miscellaneous problems, to direct copying without referencing (up to 59 words).

Compared with prior seven studies reviewed in Table 1, only the current study provides evidence which indicates that the more proficient students used more sophisticated vocabulary including higher type-token ratios and more academic words than the less proficient students. It seems that good source using skills required in a summary-and-response task grow with the increase of the general language proficiency which involves more sophisticated lexico-grammatical features in better writing performance.

Task differences are often claimed to be responsible for L2 writing performance (Petrić & Harwood, 2013; Shi, 2004). The current study designed an hourly summary-and-response task where all participants mainly adopted the summary citation strategies. Yet, in Campbell (1990) where a similar task was designed, her NNS students relied on quotations more. Even in Yeh and Liou (2000), in a similar EFL context as that of the current study where a take-home extended research paper was required, quotations were found to be the dominant citation format for their EFL students. Task effects on citation format would require further studies to pinpoint how EFL learners of different proficiency levels cope with source-using tasks of various types.

Conclusion

To contextualize this EFL writing story, 42 second-year college student-writers from two writing programs in an Asian country participated and were classified into two groups (A or B) based on their different English proficiency levels. They were asked to read a text, summarize its content, and provide a response. The more proficient group were found to write longer essays with better rated text quality and more complex vocabulary. They also demonstrated a slightly better skill in incorporating the source text into their essays. It seems that good source using skills required in such a summary-response genre grow with increase of the general language proficiency which involves more sophisticated word use as well as better and more fluent writing performance. However, a variety of misattribution problems were found in the two groups and point to needed instruction for these EFL students in spite of their different proficiency levels.

The limitations of using a single source in the current study mainly lie in not being authentic in the context of real-world (academic) writing tasks where often writers would be required to integrate various source texts or multi-modal materials (including audio or even video materials), and construct their final written pieces. Design of the writing assignments requires more thoughts from EFL writing instructors. In our study, one source text seems challenging enough for the EFL students as such arrangement also facilitates the research procedures for identifying textual borrowing. Several other lines of future research can be pursued. The task effects in EFL settings warrant future research to understand more why different citation formats are preferred in various source-using

writing tasks. Our study does not address the issue of syntax-related problems; future research can work toward that direction by revealing how the attribution of EFL students of different proficiency levels is impacted by their diversified grammatical performance. Although we conducted careful analysis on detailed textual borrowing by counting the length of sentences involved, citation purposes in context are not fully considered. The purpose of source-using determines how a writer chooses proper citation format to merge others' ideas with one's own thoughts in a sophisticated way. This is the type of knowledge EFL writing teachers need to provide in their instruction but also a crucial dimension L2 source-using researchers need to take into account in dealing with students' attribution behaviors. Last, both the descriptors and the rubric for classroom-based assessment can be communicated to the EFL students for acquiring good source skills and to writing teachers in order to understand more about various dimensions of source-using performance when other different source-using tasks and proficiency come into play.

Some studies have found resistant problems related to source use concerning vocabulary and language use still exist even after instruction on summary writing was given (Chen & Su, 2012; McDonough, Crawford, & Vleeschauwer, 2014), and suggest that much longer instructional time is needed. We echo their calls for needed pedagogical efforts in a systematically designed writing program of more than one semester long (common in Taiwan). Specifically, we find academic words are an area less proficient EFL students need to devote their efforts to while learning to write source-based texts.

Second, explicit source-using instruction is needed not only for the weak group but also for the strong group in this EFL study. The more proficient students showed in their essays close to 60% improper attribution cases (including non-referenced and semi-referenced ones), which requires instructional efforts to scaffold their learning of adequate source-using skills. The findings in the current study help us pinpoint both general and academic writing features; in particular, how specifically less proficient EFL writers need to work toward better source-using writing: more academic words, skills in both indirect and direct citation formats, and stronger general writing abilities (concerning length and quality). Proper attribution takes a great amount of writing practice through various tasks as these developing writers move to mature source use.

If the growth of source using skills happens at the same time with general writing skills, then both students and L2 writing teachers should know general writing indeed sets a solid foundation for good academic writing performance. EFL writing program directors should be informed about the finding of the current study in order to design individual programs with adequate proportions of both general writing and academic writing tasks or assignments in their own context to reach feasible goals, given different types of learners, course/curricular objectives, and resources.

Note

1. https://www.ets.org/Media/Tests/TOEFL/pdf/Integrated_Writing_Rubrics_2008.pdf.

Acknowledgements

The paper is funded by a research project of Ministry of Science and Technology (Taipei, Taiwan) under NSC102-2410-H-035-051-MY3.

References

- Ädel, A. (2014). Selecting quantitative data for qualitative analysis: A case study connecting a lexicogrammatical pattern to rhetorical moves. *Journal of English for Academic Purposes*, 16, 68–80.
- Astika, G. (1993). Analytical assessment of foreign students writing. *RELC Journal*, 24(1), 61–72.
- Baba, K. (2009). Aspects of lexical proficiency in writing summaries in a foreign language. *Journal of Second Language Writing*, 18(3), 191–208.
- Campbell, C. (1990). Writing with others' words: Using background reading text in academic compositions. In B. Kroll (Ed.), *Second language writing: Research insight for the classroom* (pp. 211–230). Cambridge: Cambridge University Press.
- Chen, Y., & Su, S. (2012). A genre-based approach to teaching EFL summary writing. *ELT Journal*, 66, 184–192.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning & Technology*, 11(3), 38–64.
- Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, 10, 5–43.

- Donohue, J. P., & Erling, E. J. (2012). Investigating the relationship between the use of English for academic purposes and academic attainment. *Journal of English for Academic Purposes*, 11(3), 210–219.
- Gebril, A., & Plakans, L. (2013). Toward a transparent construct of reading-to-write tasks: The interface between discourse features and proficiency. *Language Assessment Quarterly*, 10, 9–27.
- Gebril, A., & Plakans, L. (2009). Investigating source use, discourse features, and process in integrated writing tests. *Spain Working Papers in Second or Foreign Language Assessment*, 7, 47–84.
- Gebril, A., & Plakans, L. (n.d.). Borrowed words: The impact of source texts, proficiency, and topic on the lexical diversity of integrated writing. Unpublished manuscript.
- Hu, G., & Lei, J. (2012). Investigating Chinese university students' knowledge of and attitudes toward plagiarism from an integrated perspective. *Language Learning*, 62(3), 812–850.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.
- Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing*, 15, 261–278.
- Keck, C. (2010). How do university students attempt to avoid plagiarism? A grammatical analysis of undergraduate paraphrasing strategies. *Writing & Pedagogy*, 2(2), 193–222.
- Keck, C. (2014). Copying, paraphrasing, and academic writing development: A re-examination of L1 and L2 summarization practices. *Journal of Second Language Writing*, 25, 4–22.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307–323.
- McDonough, K., Crawford, W. J., & Vleeschauwer, J. D. (2014). Summary writing in a Thai EFL university context. *Journal of Second Language Writing*, 24, 20–32.
- Pennington, M.C. (2010). Plagiarism in the academy: towards a proactive pedagogy. *Writing & Pedagogy*, 2(2), 147–162.
- Petrić, B., & Harwood, N. (2013). Task requirements, task representation, and self-reported citation functions: An exploratory study of a successful L2 student's writing. *Journal of English for Academic Purposes*, 12(2), 110–124.
- Plakans, L. (2009). Discourse synthesis in integrated second language writing assessment. *Language Testing*, 26, 561–587.
- Plakans, L., & Gebril, A. (2013). Using multiple texts in an integrated writing assessment: Source text use as a predictor of score. *Journal of Second Language Writing*, 22, 217–230.

- Polio, C., & Shi, L. (2012). Perceptions and beliefs about textual appropriation and source use in second language writing. *Journal of Second Language Writing, 21*(2), 95–101.
- Shaw, P., & Pecorari, D. (2013). Source use in academic writing: An introduction to the special issue. *Journal of English for Academic Purposes, 12*(2), A1–A3.
- Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication, 21*, 171–200.
- Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integrated reading/ writing assessment. *Journal of Second Language Writing, 21*(2), 118–133.
- Wette, R. (2010). Evaluating student learning in a university-level EAP unit on writing using sources. *Journal of Second Language Writing, 19*, 158–177.
- Vleeschauwer, J. (2007). *Steps to writing well*. Boston: Wadsworth.
- Yeh, H. Y., & Liou, H. C. (2000). A study of using the Web articles to support college English students' ideas in writing. In S. S. Young (Eds.), *Proceedings of ICCE/ICCAI (International Conference of Computers in Education) 2000* (pp. 309–316). Taiwan: National Tsing Hua University.
- Zhang, C. (2013). Effect of instruction on ESL students' synthesis writing. *Journal of Second Language Writing, 22*, 51–67.

Hsien-Chin LIOU is Professor of Department of Foreign Languages and Literature at Feng Chia University, Taichung, Taiwan. She has taken a long-term interest in L2 writing, EAP/ESP, and the use of technologies for literacy skills. She has published mainly in CALL-related and general ELT journals.

Appendix A
Coding Scheme for Textual Borrowing and Source Use

Major Categories	Subcategories	Codes	Example From Student's Writing	Source Text
With No Reference (N)	Strings of words exactly copied (C)	NC	By the way, besides <u>mandating equal-employment opportunities for the</u> disadvantaged, it could not be more emphasized that our attitudes to the disadvantaged should be enormously altered. (A5, 5)	... issued regulations that mandate equal-employment opportunities
	Paraphrase	NP	Therefore, the Department of labor legislated a policy to stop these unfairness. (B12, 12)	for the handicapped
	Summary	NS	The story is about a blind person talking about his experiences in his life. (B20, 14)	... the Department of Labor issued regulations that mandate equal-employment opportunities for the handicapped.
Semi-referenced (x) -those with the article... the author, <u>he</u>	Strings of words exactly copied (C)	xNC	He had been <u>told</u> that it was impossible <u>for a blind person to practice law</u> . (B10, 12)	... told that it was impossible for a blind person to practice law...
	String of words slightly modified by adding/ deleting words or using synonyms for content words (M)	² xNM +(xNC)	... he told use that he <u>asked</u> the <u>ticket agent</u> for help in the airport, and the person <u>picked up the phone</u> immediately, <u>calling</u> for the <u>hostess</u> . (B20, xNM(26)+xNC(11))	when I go to the airport and ask the ticket agent for assistance to the plane, he or she will invariably pick up the phone, call a ground hostess.
	Paraphrase	xNP	Much employer have misconception on him that he is the blind, so he cannot work. (B22, 15)	The toughest misconception of all is the view that because I can't see, I can't work
	Summary	xNS	This article is about the author who is blind. (B10, 9)	

(Cont'd Appendix A)

Major Categories	Subcategories	Codes	Example From Student's Writing	Source Text
With Reference to the Author (R)	Strings of words exactly copied (C)	RC	Harold Krents said, "I have been completely dependent on the image I create in the eye of the observer". (B30, 15)	Blind from birth, I have never had the opportunity to see myself and have been completely dependent on the imagine I create in the eye of the observer.
	String of words slightly modified by adding/deleting words or using synonyms for content words. (M)	RM(+RC)	The author, Harold Krents, he got the qualification included a cum laude degree from Harvard College and a good ranking in my Harvard Law School class. (B19, RM (29)=+RC (22))	...my qualifications included a cum laude degree from Harvard College and a good ranking in my Harvard Law School class.
	String of words slightly modified by reformulating syntax or structure of the original text (R)	RR(+RC)	People assume that Krent cannot see, so obviously he cannot hear. (B5, 7)	There are those who assume that since I can't see, I obviously also cannot hear.
	Paraphrase	RP	However, the turning point was when in 1976, the Department of Labor finally issued laws that allow handicapped people to have equal opportunities to be employed as a wellbeing person would. (B14, 31)	On April 16 [1976], the Department of Labor issued regulations that mandate equal-employment opportunities for the handicapped.
	Summary	RS	Darkness at Noon was written by Harold Krents. (B22, 8)	

(Cont'd Appendix A)

Major Categories	Subcategories	Codes	Example From Student's Writing	Source Text
		RS(+RC)	Just like what Mr. Krents referred in this article, while he was in the hospital, the staff there asked the orderly who had been wheeling him as if he was not a person who was able to speak. (A5, 7)	"What is his name ?" the women asked the orderly who had been wheeling me.
With Quotation (Q)	Strings of words exactly copied (C)	QC	"Those who assume that since I can't see, I obviously also cannot hear." (B26, 13)	There are those who assume that since I can't see, I obviously also cannot hear.

Note.

¹ Two symbols are in the parenthesis where the first symbol indicates the ID of the participant whose essay shows the illustrative source-using incident, and the second symbol indicates how many words from the source are taken for source use.

² xNM+ xNC means the incident was mainly regarded as a semi-referenced modification but it was patched with verbatim copying.

Appendix B

Integrated Writing Rubrics

Score	Task Description: An essay incorporates the source text "Darkness at Noon" into the subsequent response
5	<p><i>An essay at this level accomplishes all of the following:</i></p> <ul style="list-style-type: none"> • The essay effectively and accurately presents the information from the source text, and has no exact copy of more than 3 words from the source text. • The presented ideas in both the source text and the response are highly interrelated. • The essay is well-organized, and may have occasional language errors that do not result in inaccurate or imprecise presentation of content.
4	<p><i>An essay at this level mostly accomplishes all of the following:</i></p> <ul style="list-style-type: none"> • The essay accurately presents the information from the source text despite some information being unclear and underdeveloped; OR the essay may have occasional cases of exact copy of words from the source text. • The presented ideas in the source and the response are understandably interrelated. • The essay is generally organized and may have more frequent or noticeable minor language errors that do not result in anything more than an occasional lapse of clarity.
3	<p><i>An essay at this level is marked by one or more of the following:</i></p> <ul style="list-style-type: none"> • Although the response is definitely oriented to the task, it conveys only vague, global, or unclear points made in the reading. • Some key points taken from the source may be incomplete, inaccurate, or imprecise; the ideas presented in the source and the response are sometimes not well connected, although the essay may copy some words from the source. • The essay is partly organized, and with errors of usage and/or grammar may be more frequent or may result in noticeably vague expressions or obscured meanings in conveying ideas.
2	<p><i>An essay at this level is marked by one or more of the following:</i></p> <ul style="list-style-type: none"> • The response significantly misrepresents the reading. • The response significantly omits or significantly misrepresents important points taken from in the source. Or the ideas presented in the source and the response are somewhat disconnected, although the essay may copy quite a few words from the source. • The essay is barely organized, and with language errors or expressions that largely obscure meaning, or that would likely obscure understanding of key ideas for a reader not already familiar with the reading. • The essay contains too much information from the source (about 60%) and lacks the author's own ideas.

1	<p><i>An essay at this level is marked by one or more of the following:</i></p> <ul style="list-style-type: none">• The response provides little or no meaningful or relevant coherent content from the source, which may or may not come from misunderstanding of the content.• The language level of the response is so low that it is difficult to derive meaning.• The entire essay contains too much information from the source and lacks the author's own ideas.• If the author ever copies words from the source, they are not connected to the response.
0	<p><i>An essay at this level merely (a) copies sentences from the reading, (b) rejects the topic or is otherwise not connected to the topic, (c) is written in a foreign language, (d) consists of keystroke characters, or (e) is blank.</i></p>

Appendix C

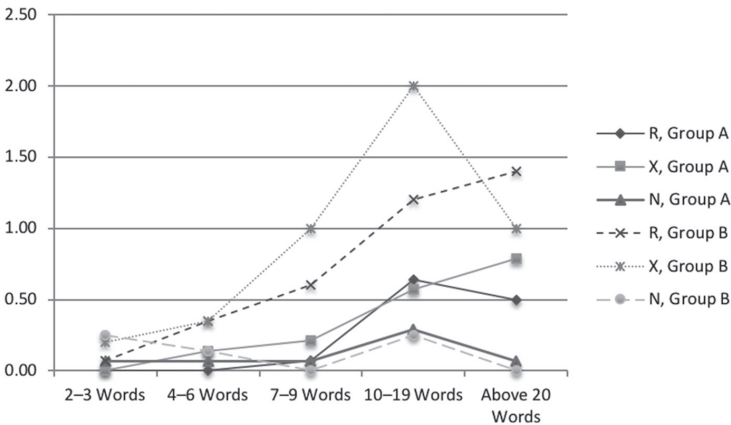
Additional Tables for Findings

Comparison of Patchwriting Incidents between Two Groups (Raw Number)

	Group A	Group B	Sum (Group A+B)
(I) With No Reference	0	0	0
Semi-referenced	3 (xNP)	4 (xNM, xNP, xNS)	7 (41.18%)
(II) With Reference	2 (RS)	8 (RC, RM, RR, RS)	10 (58.82%)
Total (Column)	5 (29.41%)	12 (70.59%)	17 (grand total)

Note. 3 (xNP): raw number of 3 cases distributed among 3 sub-categories of xNP, a semi-referenced paraphrase incident.

Length of Strings of Words Borrowed Concerning Three Referencing Categories from Two Groups (Group Means)



Lengths of Borrowed Strings of Words on Average between Two Groups

	R, Group A	X, Group A	N, Group A	R, Group B	X, Group B	N, Group B
2-3 Words	0.00	0.00	0.07	0.04	0.07	0.11
4-6 Words	0.00	0.14	0.07	0.14	0.14	0.04
7-9 Words	0.07	0.21	0.07	0.29	0.50	0.04
10-19 Words	0.64	0.57	0.29	0.57	1.00	0.07
Above 20 Words	0.50	0.79	0.07	0.68	0.50	0.00

Note. 3 main categories—R: referenced; X: semi-referenced; N: non-referenced.