

Source Distortion in Patchwritten EFL Academic Texts

Katalin Doró

University of Szeged, Hungary

dorokati@lit.u-szeged.hu

1. Introduction

Learning to develop appropriate source use is an essential part of academic writing. Efficient citation practice shows knowledge of the field, supports the line of argument, credits the author of the source texts and, even more importantly, indicates the place of the author in the argument. Novice writers, especially those using an L2, face considerable difficulties when asked to produce essays, summaries and literature reviews in which they should not only focus on textual mechanisms, but also integrate the ideas of others, evaluate sources and show authorial stance. Texts produced by university students, either in their L1 or L2, include forms of source misuse and inappropriate citation practices to an alarming degree. This is often the case even when students are familiar with general citation conventions and consequences of source misuse. Inappropriate citation can lead to misunderstanding, unclear authorial stance, or even charges of plagiarism. While previous research has extensively studied the types of source misuse, the reasons behind this behaviour, students' and instructors' attitudes and possible forms of intervention, less is known about how the source texts are changed inappropriately or kept untouched when integrated into students' own texts.

This paper analyses the types of source distortion found in patchwritten academic texts (seminar papers, thesis drafts and final versions) written by Hungarian undergraduate and graduate students of English. Compared to the paraphrasing error studies carried out under controlled circumstances with the rewriting of short paragraphs and a given set of sources, the texts under investigation integrate various sources to different degrees in order to form longer pieces of academic texts. Types of source misuse, such as the lack of appropriate and clear source indication, sentence structure errors, semantic distortion and other writing mechanism problems (e.g., numbering and referencing untouched), are reported and discussed.

2. Previous research on citation practices

Source text use by foreign or second language (L2) writers has received particular attention in the last few decades. The integration of the textual chunks of other authors has been referred to by different terms, including, but not limited to, *textual borrowing* (Baily & Challen, 2015; Keck, 2014, Petrić, 2012; Shi, 2004), *source text borrowing* (Weigle & Parker, 2012), *copy-paste plagiarism* (Haen & Molnar, 2014; Mozgovoy, Kakkonen & Cosma, 2010), *textual appropriation* (Shi, 2006, 2012), and *patchwriting* (Pecorari 2003; Howard, 1993, 1999; Howard, Serviss & Rodrigue, 2010; Li & Casanave,

2012). Most scholars and higher education instructors agree that inappropriate text reuse is most often not the result of a deliberate choice to deceive the reader. Rather it stems from a number of factors that may co-occur, such as inadequate reading skills, low language proficiency, lack of subject knowledge, misinterpretation of the sources, no clear expectations or misunderstandings about the appropriateness of close copying (Davis, 2013; Doró, 2014; Harwood & Petrić, 2012; Li & Casanave, 2012; McCulloch, 2012, Pecorari, 2003; 2008, Petrić, 2004; 2015; Shi, 2012; Stockall & Villar Cole, 2016). On top of these, the little writing practice accumulated during the few years of secondary and higher education and the limited time put into a writing task that is too difficult push many students to copy from the texts they read. Often task completion is boiled down to the accumulation of the required number of words and textual copying helps to fill the pages more quickly. When the sources are not clearly marked, there is an obscure ratio of what the student author is saying and what comes, often verbatim, from other sources. As Pecorari (2003) puts it,

the real nature of source use is only known to the writer, who uses conventional metatextual devices (citation, quotation marks, etc.) to signal the relationship. The reader decodes these signals and interprets the relationship between source and citing text. Ordinarily, these signals are the reader's only source of information about source use. If they are not used accurately, then the real nature of the source use is obscured. Transparency, then, means signaling the relationship between source and citing text accurately; its opposite is often termed plagiarism. (p. 324)

Researchers have highlighted different reasons behind textual plagiarism with debates over the role of cultural differences, language proficiency, task type and intentionality (Maxwell, Curtis, & Vardanega, 2008; Pecorari, 2003; Pecorari & Petrić, 2014; Setoodeh, 2015). Most researchers, however, agree that close copying from sources is a natural stage many novice writers go through while learning appropriate source use (McCulloch, 2012, Pecorari, 2003; 2008, Petrić, 2004; 2015; Shi, 2012). Pecorari introduced the term *patchwriting* to refer to the attempt to paraphrase with limited lexical or syntactic changes made to the source text without the intention to plagiarize. Extended patchwriting may also refer to the integration of longer textual chunks into students' writing with no or minimal changes and some connecting elements (e.g., reporting verbs and connectives) resulting in full paragraphs or even longer sections that really look like patches (Doró, forthcoming).

Hyland (2012) rightly points out that acquiring field-specific academic writing is a long process and it involves more than simply learning the mechanisms of text production. Hamilton (2016) in his work in an Australian higher education context found that students are expected too early to apply referencing conventions perfectly, creating frustration both from the students' and the instructors' side. Academic writing courses do help students in learning about field-specific citation practices, but the application of this knowledge is difficult and requires massive reading of source texts to get familiar with the discourse conventions of the chosen research community. Tomaš (2011) concludes that instructors may miss valuable opportunities to address the functions of textual borrowing when they are explaining assignments. Studies have also addressed the positive influence and learning potential of computer-assisted forms of citation practices that may aid face-to-face writing support (Baily & Challen, 2015; Choi, 2012; Stockall & Villar Cole, 2016).

Studies on L2 writers' citation behaviour and source use have used both qualitative and quantitative methods and included participants from different language and educational backgrounds (e.g., Baily & Challen, 2015; Luzón, 2015; Yoke, 2013). The most recent research on textual borrowing has used corpus methods

to balance for the frequent assumptions about the nature and the degree of source misuse students have in their assignments (Keck, 2014). As Keck (2014) points out, larger corpus-based studies can counterbalance for the type of conclusions drawn by earlier research mainly done by hand with a limited number of student texts. While corpus methods give us exact frequency counts of certain types of source use, studies like that of Keck also limit researchers to the analysis of specific task types, such as the summary of assigned sources.

3. The study

3.1. Rationale and research questions

The present study looked at texts produced by undergraduate and graduate level EFL students to evaluate what types of source distortion happen when they integrate source texts into their own. As the main aim was to list and illustrate main distortion types, a quantitative analysis of source misuse was not included. Instead, a manual checking of student texts and, when available, the source texts was done not only to document the presence and level of patchwriting, but also to see how the original texts are changed when students rely very heavily on textual borrowing.

3.2. Participants

Participants were students of English at a Hungarian university where the language of instruction for their English course is English. They did not include first-year students who have minimal or no previous academic writing experience in English, low level English proficiency and often submit heavily plagiarized assignments until they understand the general requirements of source use. The English Study program of the participants includes an introductory writing course in their first year, an academic writing class in their second year and a thesis writing seminar in their third year. Additional academic reading and writing are practiced and discussed in disciplinary seminars in which they receive writing assignments that require source integration. Therefore, the participants had general knowledge about field-specific writing conventions and knew about academic integrity requirements. The applied linguistics texts produced by the students use in-text, integral and end of the paragraph citations all of which should clearly credit the sources used, mark the relationship between source and student texts and show authorial stance.

3.3. The corpus

The corpus consists of approximately 200.000 words of academic essays, summaries, comments, literature reviews and source-dependent introductory chapters of BA and MA level theses. All texts were produced on applied linguistics topics as regular course assignments or as a final graduation requirement. Unlike many studies that assign participants a set of sources, these texts included those that students searched for themselves and found relevant for their chosen topic.

4. Results and discussion

4.1. Incorrect source attribution

Although the participants had had some practice with academic writing and received instruction on source integration, extreme source misuse, such as verbatim copying of large sections, often without final editing, was detected in the corpus. This form of extended patchwriting, fortunately, is rare. However, much more often students identify key phrases, sentences or even larger chunks and partly reorganize them and paraphrase them to form what they consider to be their own texts. Reference to the source in most of these cases happens only at the very beginning or very end of the paragraphs, leaving the reader uncertain about the proportion of original and copied sections. In the examples 1 and 2 the student texts create the impression of a legitimate summary while these passages are, in reality, close copies of the original ones (source texts are provided below the student texts).

(1) Nation adds that most English content words can change their form by adding prefixes or suffixes. According to him the affixes can be divided into two types such as inflectional and derivational ones.

Cf. Most of the content words of English can change their form by adding prefixes or suffixes. These affixes are typically divided into two types: inflectional and derivational (Nation, 2001, p. 263).

(2) One of the most important messages of this study is that it is never too late to pursue an interesting, challenging career based on passion. Unfortunately it is still a problem that people with dyslexia do not get the appropriate help. Naturally, it would be the ideal state if dyslexic students needs were satisfied by the educational system but we could not reach this aim. Fink however draws the attention to the possibility that if individuals with dyslexia lack the necessary background and training but are interested in pursuing challenging fields, career counselors can brainstorm with them to help figure out how to obtain the necessary training (Fink, 2002).

Cf. It is never too late to imagine and pursue an interesting, challenging career based on passion. If individuals with dyslexia lack the necessary background and training but are interested in pursuing challenging fields, career counselors can brainstorm with them to help figure out how to obtain the necessary training (Fink 2002, p. 7).

It often happens that the sources are distorted when the authors' names are misspelled or wrong publication dates or numbering are provided. In most cases this can be a simple typing error or a sign of a copy-paste writing methods, but may not influence the source text adaptation. The same happens when a source is missing from the reference list, but is indicated in the text. However, if the source is not easily traceable from an in-text reference or references are mixed or randomly provided, we can talk about real source distortion. In example 3 below, one of Nation's 1989 articles is indicated as the source, but the information provided, namely error correction, is not discussed in this Nation 1989 source.

(3) In dictionary usage words can also be checked if they are existing or not. In dictionaries the difference between the known and the new words can be found, and it is useful in error correction, too... (Nation 1989)

Example 4 illustrates a student's effort to summarize in one single paragraph not only a source article, but also some of the further sources that appear in this one. While the primary source is referenced incorrectly, both in the text and in the reference list, the secondary sources are missing from the reference list. The last sentence in

the passage refers to “the researchers”, but it is unclear which study and authors are mentioned here. This is a frequent problem and a serious form of source distortion in which primary and secondary sources are mixed. This practice makes it unclear how many of these sources the student has actually read and to what extent the summary is his or her own integration and interpretation of the referenced readings and not a condensed version of the research background sections of a few key articles.

(4) Kata Csizér, Judit Kormos and Ágnes Sarkadi also mention in their study that dyslexia as a term is difficult to define for two reasons. ... The authors prefer the definition of the British Dyslexia Association (1998), which is widely accepted in Europe. According to this statement, dyslexia is “a complex neurological condition which is constitutional in origin. The symptoms may affect many areas of learning and function, and may be described as a specific difficulty in reading, spelling and written language.” They mention studies conducted by Sparks, Ganschow, and their colleagues, in which they examined students with dyslexia who scored significantly lower than students without dyslexia on all of the components of the Modern Language Aptitude Test, which measures language aptitude, an important determinant of language learning success. Moreover, the researchers found that a number of other linguistic skills considered relevant in L2 learning, such as word recognition, spelling, pseudoword reading, word and nonword repetition, were in dyslexic language learner (Kata Csizér, 2010).

Previous studies have also pointed out the frequent incorrect source attribution of student writers, including the absence or under-use of in-text sources, the overuse of integral citations and excessive quotations (for a review see Luzón, 2015). In our corpus, the focus was on patchwritten texts with excessive textual borrowing. The four examples above show different referencing problems, but interestingly all of them indicate sources in the text, even though these are sometimes missing from the final reference lists. This could be an indicator of an alarming tendency of close-copying source texts not only for the ideas they contain, but also for the references they include and the texts themselves which already draw on a number of sources. This saves both time and effort to search for, read and integrate various sources, when all of these are readily available in some selected articles. If the students are not making very obvious referencing and syntactic errors in their texts, their writing can easily pass as assignments that meet the requirements.

4.2. Sentence structure errors and basic grammar problems

Students who draw too closely on source texts, yet try to make an effort to paraphrase to a minimal extent, soon run into the problem of sentence structure errors. Those who have low language proficiency which does not allow them to fully process the source texts, or who do a quick, careless job, may also leave basic grammar mistakes in their writing. Example 5 was most likely the result of the misunderstanding or simple misreading of the source text, which was a lecture handout the student had received. Sg was interpreted as someone instead of something, and the sentence this way makes little sense, which should have been noticed if the text had been proofread.

(5) Face threatening act (FTA) appears when someone is said that represents a threat to another individual's expectations while face saving act appears when someone is said in order to lessen the possible threat (Brown and Levinson, 1987).
Cf. Face threatening act (FTA) - when sg is said that represents a threat to another individual's expectations regarding self-image. Face saving act - when sg is said in order to lessen the possible threat (lecture handout).

Students with lower language proficiency level often produce very uneven texts. Borrowed chunks and partially paraphrased parts are mixed and sometimes connected in an ungrammatical way. These cases are relatively easy to notice, unlike the examples in the previous section in which ready-made sentences are lifted verbatim from source. Probably the student author of example 6 did not understand the source text or lacked language skills to paraphrase.

(6) According to Harmer teaching listening skills shares similarities with reading, but difficulties arouses, most importantly, the text itself is different. While the written text is static and can read again at the speed of the reader, the spoken text being on tapes can be repeated, but only at its speed not the reader's (Harmer, 1991).

4.3. Semantic distortion

Various types and degrees of meaning distortion can happen while sources are borrowed, summarized or paraphrased. Reasons may vary including, among others, inadequate reading and note-taking skills, lack of careful proofreading with the sources at hand, low proficiency level and the misinterpretation of the sources used. Very few studies have addressed the semantic distortion of source texts that are borrowed and integrated into students' writing. Also, when these excessive forms of textual borrowing (also called near copies and close paraphrases by Shi [2004]) are detected either in research or in educational settings, little further analysis happens concerning the way the source text is changed. A careful, word-for-word comparison of the source and the student texts can reveal some unexpected solutions students have. Examples 7–9 illustrate the minimal changes the authors made to the source texts. All three students referenced their sources and all seem to be legitimate paraphrases or summaries at first glance. Interestingly, the three examples below exchanged one single word in the original text, resulting in meaning distortion: *ambiguous* is changed to *vague*, *valuable* to *valid* and *fairly* to *very*. We do not know whether these word changes were attempts to paraphrase or whether the students simply felt more comfortable using the new words without noticing the differences in meaning.

(7) This view is supported by Lance (1977) who states that the meaning of grammar is not vague, but polysemous, having more meanings. Cf. When an individual uses the word *grammar*, he/she may be referring to any one (or more) of a variety of theoretical constructs. The term *grammar*, in other words, is polysemous—not ambiguous, but “having more than one basic meaning... (Lence, 1977, p. 43).

(8) Mufwene claimed that the subject's grammar supported the structural claims of Bickerton's hypothesis so the UG orientation of Bickerton's LBH seems to be valid. Cf. Mufwene asserted that the subject's grammar supported the structural claims of Bickerton's hypothesis, though genetic claims were not similarly supported. The UG orientation of Bickerton's LBH was deemed valuable, though Mufwene suggested that UG features of acquisition are also available to adults and hence would afford them agency in the creolization context (Barikmo, 2007, p. 23).

(9) First she speaks about *real-life listening* situations such as listening to the news, weather forecast, sport reports, chatting at a party, receiving instructions on how to do something, watching a film or a television programme. Moreover, she follows with hearing a speech or a lecture, attending a formal occasion etc. She adds that this list is naturally rough and incomplete, but very representative.

Cf. Now this list is naturally rough and incomplete, but fairly representative (Ur, 1984, pp. 2-3).

Student in example 9 selectively copied some items from a list of thirty provided in the source, working with a close copy for most of her paragraph. Reference to Ur (1984) is made at the end of the paragraph, in parenthesis. The reporting structures at the end of the paragraphs indicate that the student was familiar with academic writing mechanisms and conventions and the importance of referencing. Nevertheless, most of the text is borrowed.

The above examples also support the view of previous research which concludes that the teaching of the basic rules of academic writing may not produce assignments with legitimate source use. All samples reference their sources and most show the knowledge of field-specific integral citation. Yet, this remains a surface level knowledge if the rest of the text is closely copied from the sources and only a skeleton is provided for the textual borrowing. This is an alarming tendency as students easily deceive their readers, and may also feel they are doing the right job, either intentionally or unintentionally.

5. Conclusion

This study reported on some frequent textual borrowing mechanisms of students of English in their applied linguistics assignments. Different levels and forms of source misuse were identified, focusing on cases in which large textual chunks were lifted from the sources. Patchwriting, near copies and close paraphrases were found to be common forms of textual borrowing also in previous research (Shi, 2004; Jahić, 2011). Reasons for borrowing extensively from the sources may vary, but the examples of incorrect source attribution, syntactic and semantic distortions discussed in this study suggest that students with heavy source reliance aim to produce texts that look good on the surface. In-depth interview studies in the future may shed more light on what academic writing strategies push students to adopt the strong textual borrowing reported in this study and whether they are aware of source distortion.

It is important to raise both students' and instructors' awareness to the excessive textual borrowing and source integration problems through the examples of student (peer) texts and highlight the importance of good reading, note-taking skills and writing practice. Through this it is hoped that the frequency of unintentional plagiarism and source misuse can be lowered.

6. References

- Baily, C., & Challen, R. (2015). Student perceptions of the value of Turnitin text-matching software as a learning tool. *Practitioner Research in Higher Education*, 9(1), 38-51.
- Barikmo, H. (2007). Perspectives on creole genesis and language acquisition. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, The 2007 APPLE Award Winning M.A. Essay in Applied Linguistics*, 7(2), 1-25.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

- Choi, Y. H. (2012). Paraphrase practices for using sources in L2 academic writing. *English Teaching*, 6(2), 51–79.
- Csizér, K., Kormos, J., & Sarkadi, Á. (2010). The dynamics of language learning attitudes and motivation: Lessons from an interview study of dyslexic language learners. *Modern Language Journal*, 94(3), 470–487.
- Davis, M. (2013). The development of source use by international postgraduate students. *Journal of English for Academic Purpose*, 12(2), 125–135.
- Doró, K. (2014). Why do students plagiarize? EFL undergraduates' views on the reasons behind plagiarism. *Romanian Journal of English Studies*, 11(1), 255–263.
- Doró, K. (forthcoming). Extended patchwriting in EFL academic writing – signs and possible reasons In M. Chitez, C. Doroholschi, O. Kruse, L. Salski, & D. Tucan (Eds.), *University writing in Central and Eastern Europe: Tradition, transition and innovation*. Springer.
- Fink, R. P. (2002.). Successful careers: The secrets of adults with dyslexia. *Career Planning and Adult Developmental Journal*, 18(1), 118–135.
- Haen, J.J., & Molnar, K.K. (2014). Exploration of the ethical maturity of an undergraduate student cohort. *Journal of Academic and Business Ethics*, 9, 1–10.
- Hamilton, J. (2016). Attribution, referencing and commencing HE students as novice academic writers: Giving them more time to 'get it'. *Student Success*, 7(2), 43–49.
- Harmer, J. (1991). *The practice of English language teaching*. London: Longman.
- Harwood, N., & Petrić, B. (2012). Performance in the citing behavior of two student writers. *Written Communication*, 29(1), 55–103.
- Howard, R. M. (1993). A plagiarism penitence. *Journal of Teaching Writing*, 11(3), 233–246.
- Howard, R. M. (1999). *Standing in the shadow of giants: Plagiarists, authors, collaborators*. Stamford, Conn.: Ablex Pub.
- Howard, R. M., Serviss, T., & Rodrigue, T. K. (2010). Writing from sources, writing from sentences. *Writing and Pedagogy*, 2(2), 177–192.
- Hyland, K. (2012). *Disciplinary identities: Individuality and community in academic writing*. Cambridge: Cambridge Applied Linguistics.
- Jahić, A. (2011). An insight into textual borrowing practices of university-level students in Bosnia and Herzegovina. *Journal of Academic Writing*, 1(1), 173–179.
- Keck, C. (2014), Investigating textual borrowing in academic discourse: A need for a corpus-based approach In V. Cortes & E. Csomay (Eds.), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber* (pp. 177–196). Philadelphia: John Benjamins.
- Lance, D. (1977). What is “Grammar”? *English Education*, 9(1), 43–49.
- Li, Y., & Casanave, C. P. (2012). Two first-year students' strategies for writing from sources: Patchwriting or plagiarism? *Journal of Second Language Writing*, 21(2), 165–180.
- Luzón, M. J. (2015). An analysis of the citation practices of undergraduate Spanish students. *Journal of Academic Writing*, 5(1), 52–64.
- Maxwell, A., Curtis, G.J., & Vardanega, L. (2008). Does culture influence understanding and perceived seriousness of plagiarism? *International Journal for Educational Integrity*, 4(2), 25–40.
- McCulloch, S. (2012). Citations in search of a purpose: Source use and authorial voice in L2 student writing'. *International Journal for Educational Integrity*, 8(1), 55–69.
- Mozgovoy, M., Kakkonen, T., & Cosma, G. (2010). Automatic student plagiarism detection: Future perspectives. *Journal of Educational Computing Research*, 43(4), 511–531.

- Nation, I.S.P. (1989) Dictionaries and language learning. In M.L.Tickoo (Ed.), *Learners' Dictionaries: State of the ArtRELC Anthology Series 23* (pp. 65–71).Singapore: SEAMEO Regional Language Centre.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Pecorari, D., & Petrić, B. (2014). Plagiarism in second-language writing. *Language Teaching*, 47(3), 269–302.
- Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing*, 12(4), 317–345.
- Pecorari, D. (2008). *Academic writing and plagiarism: A linguistic analysis*. New York: Continuum.
- Petrić, B. (2004). A pedagogical perspective on plagiarism. *NovELTy*, 11(1), 4–18.
- Petrić, B. (2012). Legitimate textual borrowing: Direct quotation in L2 students' writing. *Journal of Second Language Writing*, 21, 102–117.
- Petrić, B. (2015). What next for research on plagiarism? Continuing the dialogue. *Journal of Second Language Writing*, 30, 107–108.
- Setoodeh, K. (2015). The effect of proficiency and task type on the use of paraphrase type in writing among Iranian EFL university students. *Theory and Practice in Language Studies*, 5(12), 2480–2489.
- Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication*, 21(2), 171–200.
- Shi, L. (2006). Cultural backgrounds and textual appropriation. *Language Awareness*, 15(4), 264–282.
- Shi, L. (2012). Textual appropriation and citing behaviors of university undergraduates. *Applied Linguistics*, 31(1), 1–24.
- Stockall, N., & Villar Cole, C. (2016). Hidden voices: L2 students' compensatory writing strategies. *Teaching in Higher Education*, 21(3), 1–14.
- Tomaš, Z. (2011). *Textual borrowing across academic assignments: Examining undergraduate second language writers' implementation of writing instruction*. Unpublished doctoral dissertation, University of Utah.
- Ur, P. (1984). *Teaching listening comprehension*. Cambridge: Cambridge University Press.
- Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integrated reading/writing assessment. *Journal of Second Language Writing*, 21(2), 118–133.
- Yoke, L.L. (2013). Errors in paraphrasing and strategies in overcoming them. *Journal of Creative Practices in Language Learning and Teaching*, 1(1), 4–17.