Citation and plagiarism: Undergraduates writing from sources

Yongyan Li, Christine Pearson Casanave
Faculty of Education, University of Hong Kong; Graduate College of Education, Temple University, Japan Campus

Citation of sources is a hallmark of academic writing. However, in novice writers’ texts citation often goes hand in hand with plagiarism. Although student plagiarism has been a problem raising wide concern, not much research, in particular in the context of Hong Kong, has been conducted to investigate how students’ understanding of proper citation may connect to their writing practices and how teachers respond to students’ citational practices when Turnitin.com is involved in the assessment stage. Using a case study approach, the present study triangulates multiple sources of data to illuminate several students’ citational practices and the invisibility of the problematic aspects to their lecturers. We end the paper with a few pedagogical recommendations, addressing the role of subject professors and peer learning, noting the importance of using Turnitin in ways that provide teaching-learning opportunities, and advocating pedagogy that goes beyond the teaching of referencing skills.

Keywords: Citation, sources, plagiarism

Introduction

We draw upon other texts implicitly or explicitly when we write in academia. Citation of source texts is an explicit form of drawing upon other texts, or, a form of “manifest intertextuality” as Fairclough (1992) put it. However, citation has also been described as an “occluded” feature of academic writing, in the sense that “[t]he real nature of source use is only known to the writer, who uses conventional metatexual devices (citation, quotation marks, etc.) to signal the relationship between source and citing texts” (Pecorari, 2003, p. 324). In this regard, transparency of the relationship is expected so that the reader can decode the relationship accurately (ibid.). The lack of transparency (i.e. failing to accurately signal the relationship) apparently goes hand in hand with plagiarism. However, the occluded nature of citation also means that plagiarism can be hard to detect. In the literature, we find scenarios where plagiarism (which can range from whole-sale copying to local failures of source attribution) passes the purview of examiners (Currie, 1993; Pecorari, 2006), or gets caught and interrogated or punished (Chandrasoma, Thompson, & Pennycook, 2004; Valentine, 2006). With the use of plagiarism (or more accurately, text-matching) detection software such as Turnitin.com, detection of plagiarism does become a lot easier. However, according to Howard (2007, p. 12), “[i]n place of the pedagogy that joins teachers and students in the educational enterprise, plagiarism-detecting software offers a machine that will separate them.”

1 Another version of this study is under review with the Journal of Second Language Writing.

2 Corresponding author. Email: yongyan@hku.hk
Anti-plagiarism policies seem to have nowadays become a common feature among English-dominant universities. Indeed, they are probably found at every university in Hong Kong. In pamphlets and on websites plagiarism is defined and warned against and sanctions are spelt out, typically in a “discourse of morality” (Abasi & Graves, 2008, p. 228). Here is a brief excerpt from such a code of conduct:

*Plagiarism is regarded as a very serious offence in the academic world. It constitutes academic theft - the offender has 'stolen' the work of others and presented the stolen work as if it were his or her own. (...) In this University, plagiarism is a disciplinary offence. Any student who commits the offence is liable to disciplinary action. (...) (The University of Hong Kong, 2002, http://www.hku.hk/plagiarism)*

An anti-plagiarism pamphlet usually also provides guidance, through the illumination of examples, on how to avoid plagiarism and make proper acknowledgement of sources. However, while a student upon registration at a university would normally receive a copy of such a pamphlet in the package of enrollment materials, it is questionable whether they would spend time perusing the text, beyond catching a glimpse of the warning message at the beginning. The teaching of referencing skills seems to be a normal part of an English for Academic Purposes (EAP) course offered to first-year undergraduate and postgraduate students. Yet it is commonly noted that such preparation courses do not eliminate citational problems from student texts.

The present study was designed to investigate how students’ understanding of proper citation may connect to their writing practices and how instructors respond to students’ problems in citation when Turnitin.com is involved in the assessment process. We asked three research questions:

1. How do the students understand “plagiarism” and to what extent do they perceive plagiarism as an issue of concern in their writing?
2. How do students use sources in their writing?
3. How do the lecturers respond to the students’ use of sources?

**Context**

As part of a larger project, this study was conducted at a research-intensive university in Hong Kong (referred to as XU in this paper). A majority of the undergraduate students in Hong Kong universities are graduates from the local CMI (Chinese as a Medium of Instruction) and EMI (English as a Medium of Instruction) schools; some are from international schools where International Baccalaureate (IB) Diploma Programmes are increasingly adopted. In addition, a growing number of students are coming down from mainland China from Year 1 on competitive enrollment schemes whereas a relatively small number of undergraduates are exchange students from prestigious universities in the mainland.

Students’ diversified backgrounds imply that the students may approach the academic literacy tasks in the university at different levels of readiness and may interpret the university’s code of conduct, such as the plagiarism policy, differently. Students graduated from local CMI/EMI
schools and those from the mainland are probably similar in terms of their lack of experience in writing from sources with proper citation, as this is generally not part of their previous study programme. In the case of exchange students from the mainland, although they have had some experience of disciplinary writing in Chinese, it can be noted that generally little training is offered to university students in mainland China on how to cite sources, nor have stringent policies of plagiarism (as embodied in the form of a pamphlet on what is plagiarism and how to avoid it) become a common phenomenon in mainland universities. Therefore, in terms of the level of readiness for writing from sources, exchange students from the mainland, despite their more advanced year of study, cannot be expected to have reached a more advanced stage. In fact, given the relatively short duration of the exchange programme, these exchange students may be far less ready to understand proper citation than the regular students at the universities in Hong Kong. Distinguished from all the above-named categories of students are those who have studied the IB syllabus at some international schools in Hong Kong, where emphasis upon academic honesty and the teaching of how to use and cite sources is an unambiguously part of the IB ethos (see e.g., Wallace, 1999). These students are likely to be much better prepared in writing from sources upon entry to the university.

Methods

A case study approach was adopted in the present study. The data presented in this paper were collected by the first author with the help of a research assistant (RA) from four students (Iris, Yumin, Fanny, and Jenny) and two of their lecturers (Mark and Betty). Iris, Yumin and Fanny worked on the same Year 1 assignment and their work was graded by Mark, with Iris receiving the lowest and Yumin the highest grade among the three. The assignment required presentation of three piece of evidence on a given point of view and was supposed to be an essay ranging from 500-700 words. Jenny worked with a fellow student in a pair on a Year 3 research paper assignment. This assignment was meant to be a 3000-word research paper on a self-selected topic that should link technology to ethics. Their paper was graded by Betty. The grades received by Iris, Yumin, Fanny, and Jenny were C, A, B, and C respectively. Both lecturers are expatriate staff; one has some fluency with Cantonese and has worked in Hong Kong for a number of years and the other is relatively new in Hong Kong. The students represent different backgrounds and levels of expertise in writing: Iris has graduated from a local CMI school, Jenny is an exchange student from a prestigious university in the mainland, Yumin is a regular student at XU and has graduated from a mainland Chinese high school, and Fanny has graduated from a local international school that adopts the IB curriculum. The data collection took place at the time when the students were doing their final assignments at the end of an academic semester. The sources of data drawn upon in this paper include the following:

- Interviews and text-based interviews with the students before and after they received grades

---

3 Jenny’s fellow student (male) had a similar background with her, being also an exchange student from the same mainland university. Although the fellow student joined Jenny in the final interview with the first author and the RA, in this paper we will only use the interview data obtained from Jenny and examples of her writing in the paper that she wrote jointly with the fellow student.

4 All names given here are pseudonyms.
- Interviews with Mark and Betty with a focus on the participating students’ cases, only after they have graded students’ work
- Conference between Mark and Iris
- Students’ texts (with the lecturers’ comments) and most of the source texts cited
- The Turnitin Originality Report on Iris’s essay (obtained from Mark) and the Turnitin Originality Report on the paper by Jenny and her fellow student (obtained from Betty)

Analysis of the data involved iterative reading of the transcriptions of the audio data, analysing the student texts against source texts, and examining the case profiles separately constructed for the individual students, with a view to looking for patterns and contrasting across the sets of data. In the following we present evidence to address some of the issues raised by the research questions.

**Students’ understanding and perception of “plagiarism” in relation to their own writing**

Of the four students, three had a fair understanding of what plagiarism is according to the university policies:

Iris: Plagiarism – is that a student copy some opinion or some original thought when it is not think by him- or herself without proper citation.

(Interview in English, April 22, 2010)

Yumin: Simply put, this thing is not yours, but you intentionally or unintentionally make others feel it’s yours.

(Interview in Mandarin Chinese, May 14, 2010)

It is OK use others’ opinion and sometimes it’s necessary but you must give reference. …

(Interview in English, April 23, 2010)

Fanny: It’s if you use someone else’s words and try to make it seem like your own, or use someone’s resources and say it’s your own.

(Interview in English, April 27, 2010)

Jenny, the exchange student from the mainland, seemed to be both unfamiliar with the English word “plagiarism” and was unsure how to define it:

Jenny: Betty also told us … criteria of citations, what is cheating, I don’t know how to pronounce that word -- /ˈpleɪdʒə/ /ˈpleɪgərɪzəm/

First author: you mentioned the word plagiarism. What’s your understanding of it?
Jenny: you mean in a quantitative way?
RA: your own understanding
Jenny: well, I think, this is quite complicated
First author: Betty […] I assume she introduced this concept
Jenny: […] it’s very hard …you have read other’s …then you can come up
with what you’re gonna write. I don’t think…I think…how to pronounce that word --
RA: /ˈpleɪdʒərɪzəm/
(Interview in English, April 27, 2010)

Judging from the interviews with Jenny, it seems she primarily interpreted plagiarism as the use of others’ thoughts as one’s own. Yet it is a “complicated” issue (from a “quantitative” point of view at least, as she noted in the interview excerpt shown above). Because students like her had to learn from others’ texts, she felt this put her in a dilemma, and hence she was worried about unintentional plagiarism. She seemed to perceive a contradiction between “having her own thinking” (which she preferred) and following a “route” related to the course:

First author: Are you ever worried about plagiarizing?
Jenny: I worried, of course every student worried about that. I don’t know what I’m going to say about this, this is really complicated. As for myself, I’d like to, I’d rather have my own thinking; but you cannot talk ungrounded, because you have to talk …related to where the course is going, you have to follow the route and have your own thinking. You can’t talk ungrounded, no destination.
(Interview in English, April 27, 2010)

Iris was similar to Jenny in her belief that as students they must read in order to write and in her concern about committing unintentional plagiarism. And likewise, she did not know how to present her own argument while integrating sources properly. Talking from her memory of being called into Mark’s office on an earlier assignment for the problem of plagiarism, she said:

I do not intend to copy thought really, I always want to use my own words to interpret and express it, but maybe the situation is not allowed it, or I just don’t know the proper way of dealing with it, so I was you know… I think if I know how to present my own argument, it will be far better.
(Interview in English, April 22, 2010)

In contrast to Jenny and Iris, Yumin said he had no problem citing sources and he actually enjoyed giving citations (which is indeed clear from his citation-loaded paper) because that “shows to your lecturers that you have read a lot for doing the assignment” and “it can make your essay more professional.”

Fanny is apparently the most experienced writer among all the student participants, largely due to her previous training in the IB curriculum. She noted that toward the end of the two-year IB programme in high school, they had to submit a 3,000-word extensive essay, which was run through Turnitin: “If you plagiarize, you’ll get a fail for every single subject and won’t be allowed to do exam at the end of two years.” She described her general approach of citation:

I think most of my essays do the same thing, even if I find something in a textbook, I mostly quote it, and I say why it’s relevant to the thing I’m talking about.
(Interview in English, April 27, 2010)
Of the four student participants, Fanny seemed to be the only one who both understood what plagiarism is and was able to successfully avoid it in writing.

**How the students use sources in writing**

The texts of Iris, Yumin and Jenny contain evidence of “patchwriting,” which, according to Howard (1993), can be characterized as “copying from a source text and then deleting some words, altering grammatical structures, or plugging in one-for-one synonym substitutes” (p. 233).

By searching the library e-resources, Iris found two research articles which she felt she could cite for her assignment. From the second article (an empirical report of experiments), she cut and pasted its Abstract into a Word file and then on the basis of that chunk worked out one of her paragraphs (underlining indicates overlaps with the source):

Thirdly, there are 2 experiments conducted which are called repetitive transcranial magnetic stimulation (rTMS)(Luigi, 2005). In the first experiment, 2 left frontal lobe sites are stimulated. They are a motor site (left posterior site) and a nonmotor site in correspond to the posterior part of the inferior frontal gyrus. The second experiment focuses on the right hemisphere in which the motor site and nonmotor site of the right hemisphere are stimulated. The result showed that rTMS can induce a covert SA when applied to areas over the brain that are pertinent to language. It further shows that both the left posterior and the left anterior site are critical to language elaboration. (a complete paragraph from Iris’s essay)

Although Iris felt this relatively “direct” way of using a source would give her “a low grade,” she felt it was acceptable, because “these are the ‘definitions’ which I couldn’t create by myself.”

Yumin’s essay displayed a similar style of patchwriting, as shown below:

More evidently, another technique called the “Wada test” [Footnote 8] has an anesthetic called sodium amytal injected into the candidate’s artery leading to one side of the brain or the other. After the drug being delivered to the language side of the brain, a temporary paralysis of language function is experienced. Simultaneously, the arm opposite the patient’s “language hemisphere” gradually loses senses due to the suspended operation of that hemisphere [Footnote 9]. (from the main text of Yumin’s essay)

Paul Pierre Broca reported impairments in two patients who had lost the ability to speak after injury to the posterior inferior frontal gyrus of the brain. (from a Footnote in Yumin’s essay)

Patchwriting is one of the three forms of “plagiarism” described in Howard (1995) (the other two forms being cheating and non-attribution). Howard pointed out that often patchwriting is “a form of writing that learners employ when they are unfamiliar with the words and ideas about which they are writing” (p. 799). In the excerpts shown above, it can be seen that the source texts the
students used were quite technical and not easy to process. Thus it was perhaps not surprising that the students resorted to patchwriting. Studying the students’ complete essays and the interviewing with them would reveal, however, that they did try to digest and integrate the patchwriting into their texts as far as they could, although apparently Yumin did so more successfully than Iris.

Unlike Iris and Yumin, Fanny’s comment indicates that she aimed to avoid reproducing strings of words, even when the source text was highly technical:

> I would still try to paraphrase, because I think it’s quite obviously if you suddenly take all these difficult words and put in your essay but it doesn’t flow well with your other words. If you don’t cite properly, it’s very obvious that you took words from the research. (Interview in English, May 17, 2010)

Fanny’s use of sources in her essay illustrates the point:

> Studies on the localization certain brain functions have suggested that in normal circumstances, the left hemisphere is responsible for speech – namely the “back part of the left frontal lobe” (Klawans, 2001). I say “normal circumstances” here because the rationale for this view is based on studies of aphasia. (from Fanny’s essay)

> ...it was found that there was generally a “right ear advantage for linguistic sounds” (Yule, 2006), suggesting that language signal input into the right ear has a faster and more direct path to the language processing centre, the left hemisphere. Therefore because of this, subjects can identify the sound from the right ear more often and also quicker than they can process the sound from the left ear. (from Fanny’s essay)

On the whole the passages indicate a clear authorial voice (Baynham, 1999; Ivanič, 1998), prominently with the use of first-person pronoun I and the way Fanny linked sources with interpretations of her own (e.g. namely, suggesting that, and Therefore because of this...). Fanny herself termed this way of “putting it [a quote] into a sentence, making it flow grammatically” an “embedded quote.” Asked how she had developed a preference for this style of citation, Fanny referred to her high-school learning:

> Mainly when I was doing the IB, in our English literature class, for any literature class, we had to analyse quotes and write essays on it, so I guess I learned how to do that through English literature. …but before that we learned how to express our own opinions. (Interview in English, May 17, 2010)

So her experience of expressing her own opinions in high school had cultivated her ability to use sources with an authorial voice. Perhaps the only visible flaw in the extracts is the missing page numbers for the two quotations. When queried, Fanny regretfully noted that it was an oversight, probably resulting from having confused this with some other situation where a page number is not required.
To turn to Jenny’s citation in her paper, as noted above, she seemed to link plagiarism primarily to the use of others’ thoughts as one’s own. She thus did not consider what she did in the following problematic:

The importance of privacy is partly a matter of psychological health and comfort.

In Jenny’s paper:
Firstly, the protection of privacy is partly a matter of psychological health and comfort.
(Hugh 2003:490)

(2) The same chapter in the same book noted above, p. 490
According to Ruth Gavison (1980: 428), ‘in perfect privacy no one has any information about X, no one pays attention to X, and no one has physical access to X’. [a three-line sentence omitted here] So conceived, privacy can function as an ‘umbrella’ concept, encompassing subordinate concepts, each of which denotes a particular form of limited accessibility to others.

In Jenny’s paper:
According to Ruth Gavison, “in perfect privacy no one has any information about X, no one pays attention to X, and no one has physical access to X.” So, privacy can function as an umbrella concept, encompassing subordinate concepts, each of which denotes a particular form of limited accessibility to others. (Hugh 2003:490)

Jenny’s report of the following understanding in an interview (Interview in English, May 19, 2010) may explain her practices illustrated above:

a) She thinks what she did in (1) above “is a kind of quotation” but quotation marks in this case were not needed; she followed a “principle” of “whether the language style is similar to mine or not” to decide whether quotation marks are needed or not: if not similar, it is “weird to see a sentence without quotation mark, then see no consistency in your paper”, meaning quotation marks should be used; but if “this kind of language style is the same to what I’m thinking about”, there is no need to use quotation marks.

b) She did not know how to make citations when an author was quoted or cited in another source; but she thought what she did in (2) above was the right way.

Compared with the patchwriting in Iris’s and Yumin’s texts, Jenny’s text exhibits more extensive copying as well as inability to do proper secondary citation. It is fair to say, however, that all these three students have gaps in their understanding of how to use sources in writing, even when they seem to be able to recite the tenets of the official definition of “plagiarism” (Iris and Yumin) or at least have a partial understanding of it (Jenny).
The lecturers’ responses to the students’ citation of sources

Fanny’s skillful citation of sources did not strike Mark as outstanding; to Mark the main fault of her essay seemed to be the lack of a clear structure. With the other two students, Mark immediately picked on Iris’s problematic use of sources and checked her text with Turnitin (finding a similarity index at 25%). He then had a conference with Iris to discuss her problems of source use in her essay. In the 40-minute conference, he used a number of strategies to try to drive home the message that “you must acknowledge,” and that she should get into the habit of “over-acknowledging rather than under-acknowledging,” for she “will get respect, once the rules are there.” The following illustrates the use of an analogy:

You’ve got to get into the habit of feeling good about acknowledging someone, that’s the law; it’s like you know when you live in Hong Kong you’ve got to pay taxes, ok, it’s up to you, you can choose to pay taxes, and uhhh — [indicating disgust and unwillingness] … I don’t pay taxes, and be caught …or I’ll pay taxes and a lot of people will be benefiting, and I’ll benefit if I get sick or whatever. In the same way you’ll get more, more acknowledgement from me as assessor. (Conference in English, May 14, 2010)

In contrast, to Yumin’s essay Mark gave an “A.” The voluminous footnotes used by Yumin in his essay did seem to have successfully impressed Mark (as he wished) that he had worked hard on the assignment. To Mark, Yumin’s text was coherent and “integrated.” Mark did not check Yumin’s paper with Turnitin as he would only use the software if his suspicion was raised in going through a student text (as was the case with Iris’s text). He said he did not check everyone’s paper with Turnitin both because of time and because of the trust established between himself and the “good students” like Yumin (“you know they are not going to copy”).

The paper by Jenny and her fellow student also went through a Turnitin check (as Betty routinely ran all student papers through it). Most of the comments that Betty made by track-change on a soft copy of the students’ text relate to the use of sources, either pointing out the lack of citation (“no citation?” “you need a cite [sic] for these stories and there is a citation needed for that quote”, “that you need to cite”), or a failure to integrate sources and make a clear line of argument (“I don’t follow”, “I’m all for bring in Political theorists, but what does […] have to do with this?”)

The Turnitin Originality Report indicated the similarity index of the paper by Jenny and her partner to be 12%. When queried on a few citational problems (including those shown in the previous section) in the students’ text, Betty noted: “Part of it has already been reflected in the grade that I gave them for citations here, that I found their citation system was wanting.” Yet despite a link between the grade and the perceived “wanting” of the students’ “citation system,” it can probably be suggested that some citational problems in Jenny’s writing (such as those illustrated in the previous section) have escaped Betty’s attention.
Discussion and recommendations

In this study, three of the four student participants seemed to demonstrate a fair understanding of what plagiarism is. One student (Jenny), an exchange student from the mainland, had difficulty with the term and her citational practices seemed to be the most deviant case if compared with the other two students whose texts contained patchwriting (Iris and Yumin). The only student (Fanny) who was able to align an understanding of proper citation with appropriate textual practice was the one who had the most previous experience of writing from sources among all four students. Despite the involvement of Turnitin, citational problems in the students’ texts were to a large extent hidden to the lecturers. Mark picked up on the patchwriting in Iris’s text and used Turnitin to confirm his suspicion, but Mark did not spot the similar issue in Yumin’s paper; nor did Betty pick on the problematic source use in Jenny’s writing as illustrated in this paper. A number of factors can be enumerated to explain what may have caused such invisibility.

Iris’s text gave Mark an impression that “someone else is speaking,” but Yumin’s text was “integrated” well. Nonetheless, it might be suggested that it was easier for Yumin to appear “integrated” so that his problem was hidden from Mark, than it was for Iris. This is because Yumin was patchwriting from a book, a secondary source written in popular scientific prose, while Iris relied on extracts from empirical reports written in highly specialized language. In the case of Betty assessing the by Jenny and her fellow student, a feeling of the general “wanting” of the citational system in the paper (confirmed by the Turnitin Originality Report) helped to explain a low grade she gave to the paper, but apparently Betty was not clear about the nature of inappropriate source use in the paper. Perhaps like many teachers who use Turnitin during the assessment stage, both Betty and Mark used it as a reference without necessarily going through each instance of plagiarism thrown up by Turnitin with an examining eye; and of course, copying from book sources (as illustrated by the extracts from Jenny’s writing) may simply not be revealed in the Turnitin report.

In addition, where problematic source use escapes the lecturers’ notice, a contributing factor would be the demand on time on the lecturers’ part. Betty observed that she did not feel she could afford the time to meticulously track problems down. She bluntly admitted she had no “incentive” to “implement a high standard” (with large classes, the lack of time, and the pressure to do research placed on professorial staff like herself). Indeed the demand on time is a realistic issue that can account for academics’ reluctance to pick on or report plagiarism (McCabe, Trevino, & Butterfield, 2001; Pickard, 2006). In this light, it is admirable that Mark, as he admitted, was very much “front line” with students and was devoted to “turning them around,” using Turnitin as an aid in providing a teaching opportunity (e.g. having a conference with Iris and using analogies as a strategy to teach the latter proper citation).

All the student participants in the study, except Jenny (an exchange student from mainland China), had taken an EAP course at the time of the study, which should have contributed to their being able to recite a definition of plagiarism in the spirit of the university policy and having a basic mastery of referencing skills. However, as the study indicates, this cannot necessarily rule out inappropriate source use in their writing, and the problem may be hidden to their lecturers, even when Turnitin is used.
A few pedagogical recommendations can be derived from the findings from the present study. Firstly, subject professors could consciously try to discover hidden problems in their students’ writing in terms of the use of sources. Secondly, rather than letting Turnitin separate teachers and students (Howard, 2007), the software should be incorporated into a teaching/learning process that aims to bring instructors and learners together, although a big challenge is perhaps how to do this fruitfully and efficiently (Emerson, 2008). Thirdly, with the students’ diverse backgrounds and different levels of previous experience in writing, promoting peer learning may be particularly fruitful: for instance, more experienced student writers can be identified and encouraged to share with peers how they use and epistemologically position sources in their writing. Finally, pedagogy on using sources needs to go beyond the teaching of referencing skills and the technical aspect of searching for electronic resources or a general discussion of the reliability of electronic resources, and discuss issues such as how to investigate topics in the spirit of knowledge discovery rather than fact regurgitation, how to select sources that may be appropriate for a particular writing task, how to engage with source texts in reading, and how to integrate sources into one’s text while maintaining an authorial voice (Burton & Chadwick, 2000; McDowell, 2002). In a word, from the message “what is plagiarism and how to avoid it,” there is much to be taught and learnt.

References


University of Hong Kong (2002). What is plagiarism. [http://www.hku.hk/plagiarism/page2s.htm](http://www.hku.hk/plagiarism/page2s.htm)


Copyright © 2010 <Yongyan Li, Christine Pearson Casanave>. The authors assign to CETL, HKU the right to publish this document in full in the conference proceedings. Any other usage is prohibited without the express permission of the authors.