

ESL Student Plagiarism: Ignorance of the Rules or Authorial Identity Problem?

Robert Lankamp, Associate Professor, University of Leiden, r.e.lankamp@hum.leidenuniv.nl

Abstract

One major cause of plagiarism by ESL students (and native speakers) is, unsurprisingly, ignorance of plagiarism conventions. Recent studies on ESL plagiarism have identified a second cause: absence of one's own voice, or, as it is also referred to, lack of authorial identity. So far, it appears that these two causes of ESL student plagiarism have been studied separately, with little reference to one another. Consequently, it is not clear whether these two factors are related, leaving an element of uncertainty in the body of knowledge about ESL student plagiarism and in the development of means to help students avoid it. In this paper, a qualitative study is presented where this relationship is investigated. Six ESL graduate student writers volunteered to take part in this study, without knowing its purpose. Plagiarism was detected in four of the in total twelve essays that the students wrote over a period of six to seven weeks. Data on their knowledge of plagiarism rules and on their authorial identity were elicited by means of interviews. Although some of the plagiarizing students turned out to be both ignorant of the rules and without authorial identity, no relationship was found between these two factors. It is in fact concluded that lack of authorial identity by no means always causes plagiarism. One implication of this study is that instruction in the avoidance of plagiarism should encompass more than just teaching the content of a style guide.

Introduction

In recent years, attention in plagiarism research has shifted away from putting the blame on the circumstances in which plagiarism may occur, such as access to the internet, uninterested teachers, lack of English and membership of a culture in which plagiarism is taken less seriously. Instead, plagiarism has come to be seen as a developmental issue, especially for ESL students who were the participants in most of the studies. One of the most often cited approaches to plagiarism as a developmental issue is the "patchwriting" approach developed in for example Howard (1993, 1995, 2007). According to this approach, beginning student writers plagiarize because they are not yet able to use their own voice to write about an idea. Forced to do so anyway, they resort to cutting and pasting the text of others, and they typically paraphrase it in an effort to merge their voice with that of others. According to the patchwriting approach, such plagiarism is unintentional because students genuinely believe that they are not doing anything wrong. The patchwriting approach thus claims that there are two main causes of student plagiarism: one is ignorance of the conventions, and the other is an absence of one's own voice to express the ideas that one is writing about.

It is not claimed in this study that ignorance of the conventions and an absence of one's own voice (or authorial problems as this is also called) are the sole two causes of plagiarism. For example, according to Pecorari (2003, p. 338) plagiarizing students may have "... had their own agenda, and the avoidance of plagiarism, while not unimportant to them, was overshadowed by other concerns." Alternatively, plagiarism may be a result of the confusion and inaccuracy in the notetaking process (p. 341). Ercegovic and Richardson (2004, p. 311) mention fear of failure and parent pressure to produce high grades as contributing factors. Nevertheless, this paper focuses on ignorance of the conventions and authorial problems as the major two causes of ESL student plagiarism.

Ignorance of the conventions of plagiarism is more than not knowing the correct format of citations or references. It is more than consciousness of the fact that the work of others should be cited. Most beginning academic writers are probably aware of these things. However, what they often do not know is where to put a citation to show that it is clear which text it covers. Often there is little or no information on how to insert citations in a reasonably elegant manner. How, for example, does one avoid inserting citations after every sentence? What to do if the same content is discovered coming from multiple sources? Often students receive little or no practical information about what needs to be cited and what does not need to be cited -- when is information common knowledge (material that is the intellectual property of no one and does not need to be cited)? Different teachers may issue conflicting guidelines on what "common knowledge" is. One teacher might demand citations for almost anything which appears in print, whereas another would expect more judiciousness from her students.

Authorial problems form the second cause of plagiarism covered by this paper. As pointed out by Abasi et al. (2006) and De Voss et al. (2002), authorial problems may arise when students are expected to introduce original ideas in their written work, while at the same time these supposedly original ideas must be backed up by existing materials. As Abasi et al. (2006) put it, often "even undergraduate students are expected to write like experts-in-training" (p. 110). Similarly, De Voss et al. claim that "asking a student to create original ideas encourages plagiarism in the sense that students often feel the need to consult sources for help. How many new insights are readily available for readers of *King Lear*, for example? We ask students not just for their insights, but for their original ideas, ideas that must also -- in some instances -- be 'correct'" (p. 195).

In the same vein, Ercegovac and Richardson (2004) argue that one of the difficulties faced by beginning writers is controlling the authorial voices that speak through them, i.e. their own voice and the voices of the sources they use. Hyland (2001, p. 380) noted that for "many ESL students, plagiarism was an act of desperation. After they mentally compare their texts with target 'expert texts', they may feel so overwhelmed by the distance between what they are expected to achieve and what they feel capable of doing, that plagiarism seems the most realistic strategy." Paxton (2007) found that students had great difficulties in incorporating their own voices in their work, because they considered the textbook they were working with to be very authoritative. As a result, students' essays tended to borrow heavily from the textbook.

Noting that in many academic institutions originality of student writing is expected, Abasi et al. (2006) concluded that plagiarism was "essentially the result of students' failure to represent themselves as writers who should make a novel contribution, however modest it might be, through critically engaging with sources" (p. 114). In Abasi et al., it is argued that "student textual plagiarism can best be viewed as an issue of authorial identity construction" (p. 102).

In the study performed by Abasi et al. (2006), the more experienced ESL graduate student writers exhibited three representations of their authorial selves: as intertextually knowledgeable, as self-aligned with the professor's particular research perspectives and interests, and as having a stance with regard to the topics of their papers and having something worthwhile to say (which Abasi et al. (2006) term "representation of self as author" (p. 109). (In this paper, "representation of self as author" will be termed "authorial identity".) The less experienced subjects in Abasi et al. (2006) were assigned the textual identities of potential plagiarists rather than authors, since they primarily depended on the texts of others for what they had to say.

The notion of authorship proposed by Abasi et al. (2006) makes it possible to look at plagiarism in a way that goes beyond attributing it to ignorance (in the sense in which it was discussed above). In terms of the framework presented by Abasi et al. (2006), plagiarism can be explained as a search for authorship where the students' own authorship is deemed impossible or unreliable.

Method

The purpose of this study was to ascertain whether there is a relationship between two causes of student plagiarism: ignorance of the rules and lack of authorial identity in as defined by Abasi et al. (2006).

The participants in this study were six ESL students in a Masters course in linguistics (taught by myself, in English) at an English department in The Netherlands. All of the students were women. Four (A1, A2, B1 and B2) were Dutch. C1 and C2 were of different nationalities. C1 and C2 possessed Bachelors degrees in linguistics, obtained in their countries of origin. They had no prior experience with academic writing in English, but they did have experience with writing linguistics papers in their L1s. Both reported that they were used to writing papers where they were expected to voice their own opinions. All four of the Dutch students had completed a Bachelors degree in English, obtained in the Netherlands. Consequently, they had all had extensive experience in writing academic texts in English. Two of the Dutch students had additional higher-education degrees: A1 had a Bachelors degree in European studies (which was taught in English), and A2 had a Bachelors degree in English translation.

All of the students were between 22 and 25 years old, except for A2, who was in her early forties.

All of the students reported that they were aware that it was not allowed to quote or paraphrase text without attributing it.

These six students had responded to my call for volunteers to participate in my research. They were not told what the purpose of my research was, apart from the fact that it related to the essays they would write for my course. They had not been promised any kind of reward for their participation, nor would there be penalties. They were told that they could pull out of the project at any time with no consequences.

The following definitions of plagiarism were used. Pecorari (2003) used the term "transparency" for the accurate signaling of the relationship between source and citing text; the term for the opposite would be plagiarism (p. 324). According to Pecorari, readers "make a number of assumptions based on the principle of transparency" (p. 324). The assumptions related to plagiarism were: (1) "language which is not signaled

as a quotation was original to the writer"; (2) "that if no citation is present, both the content and form are original to the writer"; (3) "that the writer consulted the source which is cited" (p. 324) unless the source is clearly marked as secondary (Pecorari, 2006, p. 9). The violation of these assumptions results in three basic types of plagiarism: unattributed language, unattributed content, and unattributed secondary citations and references. The first two types of plagiarism are familiar.

As for the third type of plagiarism, Yamada (2003) has called attention to the phenomenon that beginning writers often focus on one or two primary sources and then cite secondary sources as if they were primary sources. According to Yamada, students do this because they do not know how to deal with secondary sources, and they are unaware that they might be doing anything wrong.

In an analysis of postgraduate second language writing, Pecorari (2006) showed how primary sources might be misused by copying or near copying references to secondary sources. Secondary citation "refers to a report of a source based not on the source itself, but upon an account of it from another, later text" (p. 9). In other words, a secondary citation is someone else's citation, and if that is not attributed, it is plagiarism.

Unattributed references are items in lists of works cited which had not been consulted by the writer of the essay. Since these items must have been provided by someone else's text, including them in a list of works cited also amounts to plagiarism.

The texts used to search for instances of student plagiarism were the two essays of the course.

The first essay, written halfway the course (week 6), counted for half as much as the second essay. The topic of the first essay was to select one of the approaches to Second Language Acquisition covered so far. The assignment was: "Write an introductory paragraph where you summarize the claim(s) made by the approach that you have selected. Next, write a number of paragraphs where you set out what the support is for these claims, and what the benefits are of this approach for language teaching." The word counts of the first essays ranged from 553 to 1,926. The number of items in the lists of works cited ranged from four to twelve.

The second essay, written at the end of the course (weeks 13-14), was a critical assessment (on a larger scale than in the first essay) of any second-language acquisition topic or approach. The word counts of the second essays ranged from 2,252 to 4,425. The numbers of items in the lists of works cited ranged from six to 42.

Possible unattributed language was detected by means of the university's online plagiarism detector. Unattributed content was identified as such when it was unlikely to have been original to the writer. Possible unattributed secondary citations were often identified by the fact that these did not have references in the list of works cited -- students seemed to have developed an apocryphal convention where secondary citations were not included in lists of works cited to avoid giving the impression that they have been actually read. Possible unattributed references (and thus unattributed secondary citations, if they appeared in the text) were identified as those that were not easily available to the writer, for example because they were not available at the university library (including online resources).

The possibility of plagiarism could not be ruled out with certainty in texts that appeared to be free of it.

The textual identities (Abasi et al., 2006) projected by the students of this study were identified through data provided by them. The students were invited to reply to questions on a separate forum on the course discussion board or by email. The purpose was to ascertain to what degree authorial identity or other projected textual identity were at the forefront of students' consciousness. They were asked, after having received the grades for each essay, whether they had preferred quotations or paraphrases and why; what they thought the role was of citations in their essays, and why they had chosen their particular topic.

Next, students' knowledge of plagiarism rules was ascertained as follows. The students who had been thought to plagiarize were invited to verify that they had indeed plagiarized. In most cases, students suspected of having used unattributed text agreed that they had not applied the rules, after the rules had been explained to them. They were then asked to account for their use of unattributed text, in order to assess whether they knew the rules or not. It was unlikely that students would know about the finer points such as unattributed secondary citations or even unattributed content, because they had not received instruction about these types of plagiarism.

Not all cases of suspected plagiarism were verified as such. For example, in the case of a reference unavailable at the university library, thus possibly an unattributed secondary reference, the student turned out to live in another town and had used her boyfriend's library card to find her material in the university library there.

After the Essay 1 data had been gathered, all of the students in the course received instruction in the avoidance of all three types of plagiarism (see above). They were told that it was expected of them that these would not recur in the second essay. For the second essay, the one student who did appear to have used unattributed text was censored for doing so. She was not, however, reported because she had been promised that participation in the research project would incur no penalties.

Any relationship between the two causes of plagiarism was first ascertained by looking at the data of the individual participants and checking for co-occurrences of plagiarism, knowledge or ignorance of the rules and type of textual identity (authorial or not authorial). Next, it was ascertained whether these co-occurrences could be seen as some sort of relationship.

The drawback of qualitative studies such as this one is that the results are not generalizable. However, the advantage is that the data permit an understanding of what is happening in a group. Moreover, a series of such qualitative studies does allow for insights that cover a wider range of situations. Finally, it is doubtful whether the essentially qualitative data of studies such as these could be converted into the types of numerical variables required for generalizable statistical analysis.

Results

Table 1 shows the Essay 1 and Essay 2 data for plagiarism and the ways in which students accounted for it. There were no cases of language plagiarism.

Table 1 Plagiarism and Accounting for Plagiarism in Essays 1 and 2

Subject	Essay 1	Essay 2
A1	no plagiarism	no plagiarism
A2	no plagiarism	no plagiarism
B1	plagiarism: unattributed secondary citations because: multiple primary citations, not clear who to attribute text to	no plagiarism
B2	no plagiarism	no plagiarism
C1	plagiarism: content because: multiple primary citations, not clear who to attribute text to	plagiarism: unattributed secondary references because: I mislaid my notes on plagiarism
	plagiarism: unattributed secondary references because: read bits of unattributed references in other works	
C2	plagiarism: content because: multiple primary citations, not clear who to attribute text to plagiarism: unattributed secondary citations because: multiple primary citations, not clear who to attribute text to plagiarism: unattributed secondary references because: read bits of unattributed references in other works	no plagiarism

According to Table 1, there was a marked decrease in plagiarism for Essay 2. For Essay 1, three of the students were identified as having plagiarized. For Essay 2 this was down to a single student. It appears that instruction did have an effect – note that the single student who did plagiarize for Essay 2 said that she had mislaid her instruction notes.

All of the students who had plagiarized were able to account for that. Content plagiarism and unattributed secondary sources were accounted for in essentially the same way. For instance, B1 thought the unattributed secondary citation was permitted because she was summing up the theory, and this was not attributable to any specific source. As another example, C1 agreed that a paragraph in her essay that lacked clear citation could be construed as content plagiarism. Her reason for not including a citation was that the

paragraph was a "combination of two or three sources that said the same thing" and that it was thus impossible to attribute it to any specific source.

Unattributed secondary references were also accounted for in the same way. For instance, the list of works cited in C1's essay contained multiple instances of works that were not available at the university library. C1 said she had not read all of these works, but she had included them in her list of works cited because she had read passages from them in other works.

The similarities in the ways students accounted for plagiarism are probably due to apocryphal attribution conventions adopted by students in the absence of instruction. These similarities also suggest that the plagiarism was not intentional, since the students did not know the real rules.

Table 2 shows the Essay 1 and Essay 2 data for textual identities projected by the students. The data included in Table 2 are student responses which come closest to being relevant to textual identity.

Table 2 Projected Textual Identities in Essays 1 and 2

Subject	Essay 1	Essay 2
A1	I think I prefer not to have too many quotes in my work as it may look like that I just put a bunch of quotes together to produce an essay. Furthermore, paraphrasing shows that you understand what has been said.	When I paraphrase I show that I understand the material I use and that I am capable of interpreting findings of previous research to make a point. [...] I referred to other research and their findings to support the point I was trying to make. I believe it shows that what I am trying to prove is not just something I made up.
A2	The attractiveness of citation is that it builds on your creativity to use it and combine various sources and hence, despite giving the proper credit, make it your own paper after all. Leave your own hallmark on it.	In general, using a citation, quotation or paraphrasing is done by me to strengthen my essay. If your own research shows a certain outcome which is supported by theories or outcomes of previous surveys it is an indication that your results are solid.
B1	I prefer to paraphrase sources. This way the essay contains my own voice rather than that of several other people.	I preferred to paraphrase. The citations showed where I got certain angles to the topic from and they sometimes functioned as an authority for certain statements.
B2	I also used citation to show the audience that it is not something I claim, but that it is someone else's opinion [...] I do not think anyone really cares what I have to say about the subject matter, because I am a nobody in the field of language acquisition. However, if I cite someone who has studied the subject or who can be seen as an authority, the statement gets more value.	I think I used paraphrasing more than quotations. I used citations to refer to previous research or to emphasize the authority of the source.
C1	I used citations because [...] if I tried to explain them in my own words there was a possibility to paraphrase them wrongly and not make my point clear.	I preferred to paraphrase than to quote, because I think it is easier for the reader to understand something that is explained this way (I believe that quoting is ok for the reader only when the reader is an advanced reader of the topic).
C2	When I finished the first version I realized that 65% of 5 pages that I wrote [...] was purely my thoughts, and that was not what we supposed to do. So I deleted everything that was about my experience, my opinion [...] I try to paraphrase the source, giving a "touch" of my own opinion.	In most cases I preferred to paraphrase rather than quote. I think this way helps me to think more about certain ideas or statements I have read and search for own words helps to understand the essence of the idea. I tried to use citations, that I found very meaningful for myself.

According to Table 2, for Essay 1 A1, A2 and B1 seemed to have projected authorial identities. However these bear no similarity to those discussed in Abasi et al. (2006). The three components of authorial identity in Abasi et al. (2006) were, in short, knowledge of the topic, awareness of teacher expectations and commitment. However, what A1, A2 and B1 seemed to focus on was their use of paraphrase (rather than quotation) to give their essay their own "voice" (see for example Ivanic and Camps 2001 for a definition and discussion of "voice").

This datum is not a function of the questions the students had been asked. For instance, as a reply to a question about the use of citations a student might have answered that she had used citations to show her knowledge of the topic (no students gave this kind of response).

B2, C1 and C2 also seemed to be concerned with "voice". In their case however they preferred to let other voices speak. C2 actually thought that that was the requirement.

As for the relationship between projected textual identity and plagiarism, C1 and C2 were identified as having plagiarized in Essay 1 (Table 1), and they also lacked authorial identity (Table 2). B1 had also plagiarized, but she did appear to project an authorial identity.

For Essay 2 almost no one had plagiarized, and almost no one reported projecting any kind of textual identity. Instead, all students except A1 expressed concerns that they had for themselves rather than for readers, so that there did not seem to be any audience awareness at all.

Discussion, Conclusion and Implications

One of the questions posed in Abasi et al. (2006) was whether students were aware of the identities they projected, whether these were psychologically real, or whether "it is an imposition on the part of writing researchers" (p. 107). In this study, judging by the responses in Table 2, there is every reason to suppose that students were conscious of the identities they projected, although of course they would not have used this terminology to refer to that.

However, the real question of this study was whether plagiarizing students exhibited ignorance of plagiarism conventions as well as lack of authorial identity, and whether students who did not plagiarize exhibited authorial identity. The three plagiarizing student writers of Essay 1 could all be considered ignorant of the rules (see above), but only two did not appear to project an authorial identity. The single plagiarizing student of Essay 2 could be considered to be ignorant of the rules, and she did not project an authorial identity. Of the three students who did not plagiarize in Essay 1, two appeared to project authorial identity. However, of the five students who did not plagiarize in Essay 2, only one projected authorial identity. In other words, students who plagiarized tended to be ignorant of the rules and tended not to project authorial identity, whereas students who did not plagiarize only sometimes projected authorial identity.

The question remains whether there was a high co-occurrence of ignorance and lack of authorial identity, indicating a relationship between these two factors. The data in Table 1 or Table 2 provide no evidence for such a relationship. Not one student accounted for their plagiarism by anything else than a peculiar view of the rules. In addition, it should be recalled that Abasi et al. (2006) and other studies cited above argued that plagiarism was an authorial problem when novice students were required to provide original ideas. No plagiarizing participant in this study indicated that her plagiarism was linked to a requirement for originality.

In fact, there was no originality requirement in either of the two essay assignments. The Essay 1 assignment required the students to interpret and to summarize, and the Essay 2 assignment asked for critical assessment but did not require originality. Table 2 shows that the three students who did not project authorial identity in Essay 1 were aware of their lack of voice, while the five students in Essay 2 who did not project authorial identity were concerned with other things. It may be that the authorial concerns of B2, C1 and C2 for Essay 1 were the product of uncertainty (see especially C2's text in Table 2) of whether originality, mostly in terms of personal "voice", was required. By the time six weeks later that students were writing Essay 2, it may have been clear to them that this kind of originality was not required, and that the teacher was far more interested in whether students understood what they were writing about. Hence the issue of authorial identity may have become far less important to the students, as shown in Table 2.

Concluding, it could be that students' concerns about (lack of) authorial identity were linked to their interpretation of the requirements of the writing assignment, and were not linked to plagiarism at all. Put in a slightly different way, the students who did not project authorial identity for Essay 1 chose to project the authorship of others instead, and were not found to have plagiarized intentionally. It remains to be seen what the results would have been in the case of some kind of requirement of originality, whether in the content of the text or in expression.

An alternative explanation for the fewer cases of projected authorial identity in Essay 2 is that the instruction on plagiarism rules (after Essay 1) had somehow resulted in a shift of attention away from audience

awareness, perhaps as a result of the focus on correct rendering of citations and references. Another possible explanation would be that Essay 2 was larger and required more effort – perhaps the increased requirements of Essay 2 (including the grade that counted for twice as much) also resulted in a shift of attention for students away from audience awareness. Whatever the case, given the lower frequency of plagiarism in Essay 2, it seems that instruction works.

These conclusions have provided material for implications for further research as well as implications for teaching writing.

The most obvious implication for further research is to ascertain whether writing assignments that do require some sort of originality would contain plagiarism that is caused by lack of authorial identity. If such plagiarism is not also caused by ignorance of the rules, it would by definition be intentional and this of a wholly different order than the type of plagiarism encountered in this study.

Another implication for further research would be to further investigate apocryphal student beliefs about plagiarism conventions as shown in Table 1.

One implication for teaching would be that teachers need to understand what kinds of knowledge and skills are involved in avoiding plagiarism and take the time to instruct their students on this matter. Another implication is that it seems important to say explicitly and clearly in writing assignments whether or not originality is required. Finally, even if the concept of authorial identity proves to have no relationship with plagiarism, it might still prove to be a useful tool in teaching. A link between teaching style, grammar and lexical choice and how the students “come across” to their readers might well provide students with motivation that is now lacking. But that is another story.

References

- Abasi, Ali R., Akbari, Nahal & Graves, Barbara (2006). Discourse appropriation, construction of identities, and the complex issue of plagiarism: ESL students writing in graduate school. *Journal of Second Language Writing, 15*, 102-117.
- De Voss, Danielle & Rosati, Annette C. (2002). "It Wasn't Me, Was It ?" Plagiarism and the Web. *Computers and Composition, 19*, 191-203.
- Ercegovac, Zorana & Richardson, John V. Jr. (2004). Academic Dishonesty, Plagiarism Included, in the Digital Age: A Literature Review. *College & Research Libraries, July 2004*, 301-318.
- Howard, Rebecca Moore (1993). A plagiarism penitence. *Journal of Teaching Writing, 11*, 233-245.
- Howard, Rebecca Moore (1995). Plagiarism is, Author ships, and the Academic Death Penalty. *College English, 57*, 788-806.
- Howard, Rebecca Moore (2007). Understanding "Internet Plagiarism". *Computers and Composition, 24*, 3-15.
- Hyland, Fiona (2001). Dealing with plagiarism when giving feedback. *ELT Journal, 55*, 375-381.
- Ivanic, Roz & Camps, David (2001). I am how I sound: Voice as self-representation in L2 writing. *Journal of Second Language Writing, 10*, 3-33.
- Keck, Casey (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing, 15*, 261-278.
- Paxton, Moragh (2007). Tensions between textbook pedagogy and the literary practices of the disciplinary community: A study of writing in first year economics. *Journal of English for Academic Purposes, 6*, 109-125.
- Pecorari, Diane (2001). Process citing: avoiding plagiarism in student writing. In K. Gray and M. Leedham (Eds.), *The Japanese Learner: Context, Culture and Classroom Practice*. Oxford, UK: Department of Continuing Education, University of Oxford.
- Pecorari, Diane (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing, 12*, 317-345.

- Pecorari, Diane (2006). Visible and occluded citation features in postgraduate second-language writing. *English for Specific Purposes, 25*, 4-29.
- Rilling, Sarah (2005). The development of an ESL OWL, or learning how to tutor online. *Computers and Composition, 22*, 357-374.
- Shi, Ling (2004). Textual Borrowing in Second-Language Writing. *Written Communication, 21*, 171-200.
- Sowden, Colin (2005). Plagiarism and the culture of multilingual students in higher education abroad. *ELT Journal, 59*, 226-233.
- Sunderland-Smith, Wendy (2005). Pandora's box: academic perceptions of student plagiarism in writing. *Journal of English for Academic Purposes, 4*, 83-95.
- Swales, John M. & Feak, Christine (1994). *Academic Writing for Graduate Students: A Course for Nonnative Speakers of English*. Ann Arbor: The University of Michigan Press.
- Yamada, Kyoko (2003). What prevents ESL/EFL writers from avoiding plagiarism?: Analyses of 10 North-American college websites. *System, 31*, 247-258.