Types of student intertextuality and faculty attitudes

Diane Pecorari a,*, Philip Shaw b

a Linnaeus University, Sweden
b Stockholm University, Sweden

Abstract

Intertextuality is a prominent feature of academic writing, and the ability to use sources effectively and appropriately is an essential skill which novice writers must acquire. It is also a complex skill, and student performance is not always successful. It is presumably beneficial for students to receive consistent messages about what source use is and is not appropriate, but some evidence suggests that university teachers and other gatekeepers may fall short of this consistency. This paper reports the findings of semi-structured text-based interviews aimed at understanding the basis of teacher attitudes and responses to intertextuality in academic writing. Teachers who were asked to evaluate the same examples from student texts differed in their judgments about whether the examples were appropriate, and provided different types of explanation for their judgments. These explanations enable us to develop a four-part typology of intertextuality which allows analytic discussion of differing judgments. The implications both of the teacher judgments and of the typology for second language writing instruction are discussed and an assessment of the relevance of our findings for the theme of this special issue is provided.

© 2012 Elsevier Inc. All rights reserved.

Keywords: Plagiarism; Intertextuality; Source use; Academic writing; Sweden; English; Faculty attitudes

Situating plagiarism

Intertextuality—the relationship between two or more texts—plays an important role in academic writing. Specifically, citation is one of the tools writers can use to frame questions, support answers, and establish a niche in the existing research (Swales, 1990). For novice academic writers, intertextuality can play an important additional role: The existing literature in a research area not only communicates the state of knowledge, it models the characteristics of good academic writing. The uses that writers make of sources vary widely according to their expertise. While the conventional aspects of citation in a number of academic disciplines have been documented from perspectives as diverse as bibliometrics and applied linguistics (see White, 2004, for a review), a large body of research (cf. Pecorari, 2008a; Roig, 1997; Shi, 2004) has established that novice academic writers experience difficulty in using sources in ways which comply with the demands of their discourse communities. This is due in large part to their being required to use unfamiliar language skills, in many cases because the language of study is a second or foreign language, and in most cases because academic register is new to them. As Bourdieu and Passeron noted, academic discourse is “jamais pour personne, même pour les enfants des classes cultivées, une langue maternelle” [never a first language, not even for the children of the cultivated classes—translation ours] (Bourdieu & Passeron, 1965, p. 18).

* Corresponding author.
E-mail address: diane.pecorari@lnu.se (D. Pecorari).

1060-3743/S – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jslw.2012.03.006
This situation creates serious challenges for L2 writers. When inappropriate intertextuality is diagnosed as plagiarism, the consequences can be serious. Yet it is widely understood that not all inappropriate intertextual relationships are properly classed as plagiarism. Novice writers may have limited understandings about when and how to signal the role that a source has played in a new text of their own. For example, Klitgård (2009) surveyed students and found that they overwhelmingly associated plagiarism with the act of copying, suggesting that they may not perceive other sorts of inappropriate source use as plagiarism. Cultural background has been widely (though not uncontroversially) associated with problematic source use, with explanations for the role played by culture ranging from differential experience with academic writing tasks (Timm, 2007a, 2007b) to varying attitudes (e.g., Shi, 2006). It has also been suggested that when learning to produce texts within a new discourse community, writers may adopt a close modeling strategy stitching together chunks of language from source texts with superficial changes. For such writers this is not only a virtually inevitable stage, it is one which may facilitate acquisition of writing skills (Howard, 1995).

Problems of source use are, therefore, both common among student writers and potentially risky, since inappropriate source use may be taken to be plagiarism, and plagiarism is a highly stigma-laden disciplinary offense. Yet despite the risks associated with misusing sources, simply avoiding using sources at all is not an option for novice academic writers. It is in the nature of academic texts that they are steeped in intertextuality; they paraphrase or quote, and—ideally—use citation as a tool to advance their arguments. The need to cite sources stems from the writer’s responsibility to make the influence that sources have had on her text transparent to the reader. The fact that there exist sources for virtually all academic texts stems from the nature of scholarly activity. It is incremental, and new work is not produced in a vacuum, it builds on existing work.

In addition to this very direct and visible type of intertextuality, academic writing, like all discourse, is subject to the sort of intertextual ties which inevitably arise when texts share features such as a common topic, purpose, or register. Texts on a common topic necessarily draw on a partially shared pool of words and phrases, both technical and metadiscoursal, and also to some extent have shared preferences for grammatical structures; for example, written academic discourse is more likely than conversation to feature impersonal devices such as the passive voice (Biber, Johansson, Leech, Conrad, & Finegan, 1999, p. 938).

Intertextuality is thus an unavoidable, and unavoidably complex, aspect of academic writing. In this situation students need guidance, and such guidance must be based on an understanding of the borders between legitimate and illegitimate types of intertextuality. Yet there is reason to think that the guidance students get from teachers may not be consistent, in part because the views held by their teachers are also inconsistent. In a study of attitudes and practices relating to source use, Roig (2001) asked university teachers to determine whether a set of examples constituted plagiarism or not, and found a degree of disagreement. In a later stage, he asked teachers to produce paraphrases of a text, and their paraphrases involved verbatim repetition to an extent which could be regarded as plagiarism. Several interview-based studies have also produced suggestions that such variation exists (Crocker & Shaw, 2002; Errey, 2002; Flint, Macdonald, & Clegg, 2005; Pecorari, 2008a; Sutherland-Smith, 2005). In a study of attitudes of teaching staff at a British university, Flint, Clegg, and Macdonald (2006) found differences in terms of how teachers defined plagiarism, and specifically the relationship they perceived between plagiarism and cheating, meaning that in this latter respect some teachers’ views deviated from university policy. Sensing the differences between their personal views and the official ones, teachers reported that they “would be more inclined to deal with plagiarism at an individual and informal level” rather than through the officially prescribed channels (p. 152). Borg (2009) was able to show that at least some differences in teacher response to what he terms “transgressive intertextuality” is accounted for by disciplinary variation.

It is potentially dangerous for students that teachers hold significantly diverse views about the sorts of intertextuality that are, and are not, acceptable. The instruction or feedback on their written work that they receive from one teacher may not prepare them for the expectations of the next teacher. Similarly, if their use of sources takes them over the line into the unacceptable, the responses may be varied, opening up the possibility that potential instances of misconduct may be dealt with inconsistently.

In order to ensure equitable responses to student work, it is necessary to go beyond understanding what common ground exists in faculty responses to student intertextuality. This entails not only investigating diversity of views but also discovering the ways that teachers account for intertextuality which they do not believe to be deceptive plagiarism. An aim of this study was therefore to document teacher responses and the readings they give student texts, that is, to document the variety of responses to intertextuality which are held. A further aim of this paper is to put forth an analytic framework for discussion, university teacher education, and policy development which can accommodate potential differences in individual and disciplinary perspectives, with the ultimate goal of more nuanced and
transparent teaching. Formulation of such a framework is a prerequisite for further investigation of the nature of intertextuality and only after that can reliable guidance be given. This paper thus aims to contribute to scholarship in this area by using text-based interviews to elicit the perceptions of instructors as a basis for a more abstract framework to guide their practice.

**Methods**

**Participants and data sources**

During the spring of 2010 semi-structured interviews were conducted with eight university teachers at several Swedish universities. The participants varied in academic rank and length of teaching experience. All were from the areas of natural sciences and engineering or medicine. Six were Swedish and two had moved to Sweden from other European countries. The interviews with three participants—Victoria (this and all other names are pseudonyms), Ulrich, and Kjell—were conducted in English; the remainder were conducted in Swedish, and where quotations from the latter set of interviews appear below, they have been translated into English (see Table 1 for an overview of the participants).

The interviews were structured around a set of five pairs of texts. Each pair included an extract from a piece of student writing, and an extract from the source which the student had used. The five extracts were from the introductory/literature portions of thesis drafts written by two (L2 English) master’s students in biology and were taken from a research corpus gathered for an earlier study (Pecorari, 2008a).

The five examples used in this study were selected to cover a range of intertextual relationships, and varied in the following ways: length, referencing, degree of secondary citation, and type of content reproduced, as illustrated in Table 2. Thus, for example, in the student text extract in the first pair less than a third of the words are also found in the source, while in the other extracts much larger proportions seem to derive from the source (with varying degrees of adaptation). Two student texts fail to refer to their source, while three have (apparently) standard referencing. In some of the examples the repeated language describes work clearly signaled as having been done by other researchers while others amount more to generalities about the topic. In order to make these relationships easier for participants to see, words which appeared in both texts were printed in red. Examples 1–4 appear in Appendix A, while Example 5 is too long to reproduce here without issues of copyright and confidentiality arising. Examples were presented to the participants one at a time.

In selecting examples which differed from each other, our intention—given our aim of documenting a range of responses and attitudes—was to create a research instrument which would elicit a broad range of participant responses, and to avoid steering responses into a narrow channel. In order to stimulate discussion, we were also careful to choose examples which appeared to leave some scope for criticism. That is to say, we deliberately avoided including any examples which appeared to be flawless models of citation behavior. The examples were also chosen to be broadly representative of the sorts of source use found in the corpus from which they were drawn, rather than to illustrate specific types of clearly inappropriate source use. Thus, with the possible exception of number five, we do not believe that there are clear right and wrong answers.

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Academic rank</th>
<th>Sex</th>
<th>Discipline</th>
<th>L1 Swedish?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Victoria</td>
<td>PhD student/tutor</td>
<td>Female</td>
<td>Environmental science</td>
<td>Yes</td>
</tr>
<tr>
<td>Kjell</td>
<td>Reader</td>
<td>Male</td>
<td>Zoology</td>
<td>Yes</td>
</tr>
<tr>
<td>Gregorija</td>
<td>Reader</td>
<td>Female</td>
<td>Computer science</td>
<td>No</td>
</tr>
<tr>
<td>Karin</td>
<td>Senior lecturer</td>
<td>Female</td>
<td>Health care</td>
<td>Yes</td>
</tr>
<tr>
<td>Lars</td>
<td>Full professor</td>
<td>Male</td>
<td>Biology</td>
<td>Yes</td>
</tr>
<tr>
<td>Mikael</td>
<td>Professor emeritus</td>
<td>Male</td>
<td>Biology</td>
<td>Yes</td>
</tr>
<tr>
<td>Urban</td>
<td>Full professor</td>
<td>Male</td>
<td>Biology</td>
<td>No</td>
</tr>
<tr>
<td>Stefan</td>
<td>Full professor</td>
<td>Male</td>
<td>Biology</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Table 2
Characteristics of the textual material.

<table>
<thead>
<tr>
<th>Extract</th>
<th>Total words</th>
<th>Words from source (usually with omission, reordering, etc.)</th>
<th>Longest unbroken string from source</th>
<th>Signaled quotation?</th>
<th>Reference given to source?</th>
<th>Nature of content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>12</td>
<td>4</td>
<td>No</td>
<td>No</td>
<td>General</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>22</td>
<td>11</td>
<td>No</td>
<td>Yes</td>
<td>Results</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>29</td>
<td>29</td>
<td>No</td>
<td>Yes</td>
<td>General</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
<td>26</td>
<td>13</td>
<td>No</td>
<td>No</td>
<td>General</td>
</tr>
<tr>
<td>5</td>
<td>311</td>
<td>285</td>
<td>121</td>
<td>No</td>
<td>Yes</td>
<td>Results</td>
</tr>
</tbody>
</table>

The student texts all came from biology, and our participants were all in disciplines which would be classed as “hard,” so while some were clearly closer to the writers in background than others, it could be assumed that they had some degree of familiarity with the norms according to which the student writing would be assessed. Importantly, as Borg (2009) notes, policies about plagiarism tend to be university-wide, so despite the real existence of disciplinary variation in conventions for writing, gatekeepers who determine whether inappropriate source use has occurred often come from outside the student writer’s background.

The interviews were audio recorded with the consent of the participants. Lengths varied from 25 minutes to over an hour. We started each interview by asking the participants to evaluate each of the five examples, using a set of questions that aimed first at establishing whether the example was appropriate, and then probing further for specific explanations. In addition to responding to the examples, participants also spontaneously provided accounts of plagiarism and source misuse and observations and generalizations about plagiarism and the forms in which it occurs. Where needed we followed up with more general questions to elicit participants’ view of related issues broadly (see Appendix B for interview questions).

Data coding and analysis

The resulting recordings were transcribed and coded using a recursive process to which both researchers contributed. Through the coding process we extracted themes, i.e., ideas, tropes and key words, which were particularly salient to participants and which were taken up either within or across interviews. The recursive coding process meant that when a new theme was identified in later interviews, the earlier transcripts were again studied for that theme (Smith & Eatough, 2006).

In the analysis below we indicate whether comments address a specific case, and if so which case, only where relevant. As noted above, participants’ comments are presented not for the purpose of establishing any true status of the example (i.e., whether it is appropriate or not); that question is not the focus of the present study. Rather our interest here is in documenting the factors which our participants took into consideration in making those determinations.

Results and discussion

In this section we present the two major themes which emerged from the interviews: that our participants had diverse and conflicting views on what constitutes appropriate and inappropriate intertextuality; and that they again were different in terms of the sorts of factors which they weighed up in coming to their judgments.

In responding to the request to comment on intertextual relationships in the examples, participants had in principle two basic options: to approve or to disapprove of the source use. In fact, in many of the 40 judgments (eight participants commenting on five cases each), the response was cautious and hedged, involving a certain amount of reasoning aloud, rather than a direct and unambiguous answer.

Conflicting views

Not only were the teachers’ judgments often hesitant, they were mutually contradictory. On seeing each case, the teachers ordinarily commented first on whether or not it was acceptable, and then on whether it was plagiarism, and
often needed to be prompted to comment on the latter aspect. Only the final example attracted clear disapproval from all of the teachers. While three and four were unacceptable to the majority of teachers, three was fully acceptable to Lars and on the borderline for Mikael and Urban, and number four was a borderline case for Lars (see Table 2 above for an outline of differences among the texts). On the other hand, the majority found examples one and two to be acceptable, although in each case one teacher disagreed, and in addition one teacher found both cases to be only borderline acceptable.

As to whether the unacceptable cases were specifically plagiarism, or unacceptable in other ways, not all teachers were able to provide an answer, and even when they did, it was often qualified and hedged. Classifying an instance of intertextuality as plagiarism was, however, the response in only a minority of cases. A total of 40 judgments were made (five examples judged by eight teachers), and 27 of them to the effect that the examples were unacceptable or of questionable acceptability. However, in only ten cases was the example classed as plagiarism or on the borderline with it (see Table 3), and it is worth noting that those ten came from four of the participants only; in other words, half of the participants did not directly classify any of the examples as plagiarism, even in a qualified way.

Decisive factors

In deciding whether the five cases were appropriate or not, the teachers discussed their reasoning and named the factors to which they were attending in making their determinations. The themes that they brought up in this process are of interest because they illustrate that not only do individuals respond differently to the same specific cases, they differ in their conceptions about which intertextual features confer legitimacy on texts generally.

Legitimate intertextuality

Appropriate intertextuality, as conceptualized by these teachers, is mainly characterized in negative terms through comments on features which problematic examples lacked rather than as comments on features which acceptable examples possessed. This is very likely due to the nature of the task with which participants were presented. That is not to say that they all disapproved of all examples, but rather that those features which did attract disapproval presented more scope for elaboration. Thus, in commenting on the first example, which he saw as unproblematic, Mikael simply answered “I think that’s okay.” Another participant, Karin, initially identified no problem with one of the examples and commented “I think we can say that it’s okay,” but when she looked more closely and decided that she did not approve of it, she was able to expand that judgment with a rationale: “but there isn’t, in the first one there isn’t any reference, is there? In that case it’s not right.”

As a result, the evidence for how the teachers conceptualized legitimate intertextuality is primarily indirect and indicated through comments about what was wrong rather than what was right. Rather than describing the characteristics of more direct but acceptable intertextuality, participants described the perimeter of it by defining what was not acceptable. They identified two parameters: the extent to which source content was expressed in what they called the writer’s “own words” and the presence of a reference to the source. Although, as we see below, individual
participants meant different things by these criteria, failure to meet either criterion as the participant sees it results in illegitimacy. Thus Gregoriija described a problematic group of students as one in which “people didn’t refer to their sources,” while a few students “took entire articles, entire articles from beginning to end.” Karin described a case in which a student had quite appropriately included a reference, but when she examined the sources, she found that the student had copied from them verbatim.

Participants also recognized as legitimate some types of intertextuality which do not require ascription of ownership, involving the use of content or text which is somehow in the public domain. In these cases no direct source is identifiable, and so no reference is needed. One prominent instance was content of the sort which is not claimed as new but is so old that it is common property (as Kjell said, “there are some things that you don’t have to, you know, cite Watson and Crick every time you mention DNA”). However, it is not always self-evident what knowledge falls into this category of common property, and one may be forced to reason about whether something is common knowledge or not, as Victoria was when she looked at the first example:

(1) I think he should have given a, a reference... so that people can go back and actually check the source and because there may be... debates about [whether] this belonged to this genus... I’m not so sure actually about the second [a discussion of essential oils in the plant]... that’s almost common knowledge I think.

By contrast with Victoria, Lars had a remarkably generous concept of knowledge that is common property, saying “this could be in a textbook somewhere as well, and what do we do then, we don’t have to cite textbooks.” If something is in a textbook, therefore, it counts as common knowledge.

Another variation on this theme was similarity of phraseology, and participants believed that stretches of language can become common property because, as they say “there are only so many ways to say the same thing.” Mikael touched on this notion with respect to terminology: “we know that within a subject area there are specialist terms... this means at the same time that texts can sometimes look very similar.” Gregoriija agreed with this, noting that “sometimes one uses long, really long expressions, for example when one speaks about Popper’s view of science.” She also regards not just specialized registers but all language as restricted, saying “if someone were to write ‘hello, hello, how are you,’ and if one were to check how many people in the world have said ‘hello,’ or say ‘this is a quotation’...” Lars echoes this, asking rhetorically, “cows have four legs, how should I write that in a different way?”

A less common but related topos is that repetition may be acceptable not because there are no other options for expressing the same idea, but because there is no value in attempting to do so. Stefan commented that Example 2 had “only very short extracts of sentences and not full sentences” in common with the source. That made the student text “approximately what it would have been like if you had read a text and tried to write using your own words.” In fact, if the copied text is relatively short (and “short” and “long” were indeed relative concepts for the participants), a forced attempt to reword it may be ill advised, according to Lars, who asked “how could you write it without it becoming noticeably worse? But if we’re talking about [a longer chunk] you can surely write it in a different way.”

This point is also recognized by Kjell, although evaluated differently. Making a more general comment, he reports saying to students “you have to somehow really write this and put it in your own words. I know that it isn’t always so easy and sometimes it’s silly but that’s how it is.” It is presumably silly because the paraphrasing does not add any value to the text, but unlike Lars, Kjell is not prepared to loosen referencing criteria to accommodate that element.

Also related to this idea of a restricted range of choices is the notion which arose repeatedly that reused text is more legitimate in some parts of the paper—methods and perhaps the opening phrases from the introduction—than in others, because of the repetitiveness of technical language or of the rhetorical moves involved. Stefan said, “people rarely react equally strongly if you copy a description of methods, it’s not possible to write it in so many different ways and there you want to be quite careful that it’s exactly the way it appeared in the source.” Similarly, Mikael said of the methods section “I think people accept less reworking.”

Several participants raised the potential value of copying phraseology, where it was appropriate to do so. Lars believed that his students would produce better writing if they engaged in more copying, saying

(2) they’ve read quite a lot but that’s reflected surprisingly little in their own [texts]—they come up with their own strange formulations when they’ve read descriptions of this a hundred times for sure, why do they write such funny stuff themselves? So I’m surprised that people don’t... make use of functional text, so to speak.
Victoria suggests that mastery of the target register comes from actively copying strings that inevitably recur in the source material, although she has in mind shorter strings: "in a way you know you... make a mixture of several people’s, you know you quote more people, you sort of create your own library of ways of expressing it.” Similarly, Kjell says “I think that it’s not used enough by the students that they in fact should read ‘I like this, why is it good,’ and try to learn from those examples, that’s how I learned.”

To sum up, our participants thought that in academic writing there is both information that does not need a reference and text strings that are freely available for re-use, although they do not agree on the boundaries of the information or on the length or type of strings that fall into this category. Several point to the value of re-use of strings in raising the quality of students’ textual output, and in their efforts to acquire the register of the discipline (see also Pecorari, 2009). Appropriate textuality could then involve not only referenced and paraphrased stretches of text but also unreferenced and unparaphrased public domain material.

Illegitimate intertextuality

The participants were also able to identify (in the examples we showed them) or describe (from their own experiences) instances of inappropriate intertextuality. However, even though these instances did not meet the criteria for good source use which the participants believed to be applicable, they were usually judged to have some explanation other than intentional deception. Ulrich said that new students in particular are overloaded with information, which can lead to unintentional problems with source use:

(3) there are so many new things, so many new bits of information, they can’t really judge what’s important and what’s less important... They take some of the things on board and not others and you know if it happens to be this bit of information about plagiarism... it’s simply they forgot it...

Victoria believed that this was an issue for experienced students as well as new ones. She said that she had never reported a student to the disciplinary board, preferring instead to speak to them about possible misunderstandings. In her experience they “tend to think that plagiarism is when you don’t do the reference or when you copy another student’s work... or if you copy like a whole piece of work.”

As those comments illustrate, the explanations for non-deceptive illegitimate intertextuality often related to students’ knowledge of the rules, to declarative knowledge. One subcategory of this was the cultural explanation, which has so often been put forward, and contested, in the literature (e.g., Ha, 2006; Shi, 2006). Lars was one of several teachers who raised this issue. He had observed that problematic source use was more likely to arise with postdoctoral researchers than with PhD students, and explained this by saying “I think it’s cultural. It’s Asian postdocs I mean.” His PhD students, by contrast, were usually Swedish. Ulrich was originally from Germany but had taught in Australia before moving to Sweden, and had experience of international students from both contexts. He identified a number of issues and pressures which he thought were particular for these students, ranging from “how much they value the written word” to how students respond to feedback, and to the fact that visas might depend on attaining a certain grade. Although he did not attribute problems with plagiarism directly to these causes, he did indicate that national background needed to be taken into account in responding to source misuse: “as you know we have quite strict rules about this... and I try not to enforce them too early in particular for a student who’s come to the country, a student who’s come for a masters, you know, what can you do?” On the other hand, Mikael presented the opposite point of view: “We also have some PhD students from India but we haven’t had the feeling that there was a culture clash or that we had very different views on how text material should be used.” At one level, given the prolific discussions of plagiarism vis à vis international students in the EAP literature, this is not surprising. However, it is interesting to note that the perceived differences in source use between international students and home students were salient even when the home students come from outside the English-speaking world. If Mikael’s observation is shared by other teachers in this context (and anecdotal evidence suggests it may be) then there are two possible explanations. The first would confirm the rather contentious idea that orientations toward plagiarism are culturally bound (see, for example, Ha, 2006; Liu, 2005; Sowden, 2005). The second would be that the tendency to create an us/them dichotomy, with the risk of “constructing an exoticized prototype of the rhetoric of the Other” (Kubota, 2004, p. 11), is not exclusively a phenomenon of the English-speaking inner circle (Kachru, 1985). Our data do not permit us to establish which, if either, of these explanations may be correct, but it is a question which would merit investigation.
While for some teachers the issue was one of knowing the rules of referencing, others saw language/writing skills as at least part of the problem. For Stefan this was an issue at word level, and students’ limited ability to find synonyms created a risk that their texts could become “too like the source if they haven’t really mastered English, because they have no synonyms etc. and they don’t dare to choose a synonym if they’re not sure what it means.”

At least two others drew attention to difficulties with higher-level writing discoursal skills. Lars noted that his PhD students usually cannot write on their own in the beginning, and require extra support in producing articles. Lars has found a creative solution to this; he asks students to write long and detailed descriptions of their figures and tables, thus forcing them to produce a prose account of their results which can then form the core of a “results” section. For Victoria, the discoursal problems were symptomatic of reading comprehension problems, since “if you can’t take in information and somehow process it you have to just sort of take direct snippets of text then I don’t think you’re… skilled enough in what you’re doing really.”

Notably, while the intertextual relationships which fall under this category were labeled as unacceptable by the teachers we interviewed, had they encountered the student texts outside the framework of this investigation, they may not have been aware of the source–student text relationship and may have adjudged the textual products to be satisfactory. Karin recounted the story of a student working on an end-of-degree project whose dependence on sources had been discovered by text-comparison software. When Karin met the student to discuss this, the student was shocked to learn that her source use was inappropriate. “She was completely shattered and said ‘I’ve been doing this all through my nursing training, I wrote my undergraduate dissertation just like this.’” The student’s surprise was compounded by her perception, based on earlier comments from the teaching staff, that they had been “so pleased with the way I handled the language!”

All of the teachers, then, identified some form of intertextuality which was not appropriate, but which was due to something other than intentional deception, though they differed in whether they believed the problem was lack of declarative knowledge about “rules” governing source use or the process of learning to be an academic writer.

Plagiarism was, however, also a relevant concept for the teachers we interviewed, although not an easy one with which to come to grips. When shown the five discussion texts, teachers found it relatively easy to decide whether each was appropriate (though they did often appear to need to qualify, reason through, or explain their determinations). However, it was more difficult to determine from textual evidence whether the inappropriate cases constituted plagiarism. For example, when Gregorija looked at the fourth, and second-most extensive, example, she quickly answered that she did not think it was all right. When asked whether it was plagiarism, she hesitated before answering “it’s unacceptable, I don’t know what you’d call it.”

Similarly, Victoria’s initial response to the third example was disapproving, but when asked if it constituted plagiarism, her answer was the following rather uncertain chain of reasoning:

(4) Well I suppose because, I mean... he did give the source, hmmm but I suppose it is plagiarism in the sense that he doesn’t show it as a direct quote... But somehow, I don’t know. I suppose it’s always easier to be a bit more kind of forgiving when after all you’ve shown that you’ve... But then on the other hand it’s in a way it’s just really almost as bad because you haven’t made the text your own really. You haven’t sort of added anything or, so... But actually considering it, I mean it would be plagiarism because he’s not showing it as a direct quote, but then I suppose because he’s given a reference at least he’s shown it’s not his own.

It has been seen that these eight teachers accounted for intertextuality in different ways. Some examples were found to be appropriate; those which were not appropriate could sometimes be explained by various features of the writer’s process or knowledge. Only a relatively small portion of the examples we investigated, and of the cases which interviewees provided, were believed to be deliberately deceptive instances of plagiarism. In addition, it is noteworthy that plagiarism as an explanation tended to be proposed last by the participants, after they had tested but been forced to discard other explanations.

**A typology of intertextuality**

Our participants’ comments as reported above reveal striking diversity, both in the judgments they made about specific instances of intertextuality, and in that they recognized a number of factors as relevant to the question of
intertextuality. We would like to propose that those views can be understood if they are accommodated within a four-way typology of intertextual relationships which emerges from further elaborating the dichotomy presented in the introduction between legitimate and illegitimate intertextuality. As can be seen in Table 4, these four categories can be distinguished by reference to four criteria: whether a given intertextual relationship is a direct one, with a specific text, or indirect, and more generally to a larger discourse; whether the relationship is perceived as legitimate; whether it is signaled by the writer; and whether it is visible to the reader (see Fig. 1 for examples).

**Table 4**
Typology of intertextuality.

<table>
<thead>
<tr>
<th></th>
<th>Indirect</th>
<th>Conventional</th>
<th>Unconventional</th>
<th>Deceptive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the textual relationship specific or generalized?</td>
<td>Generalized</td>
<td>Specific</td>
<td>Specific</td>
<td>Specific</td>
</tr>
<tr>
<td>Is the relationship legitimate?</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Is the relationship signaled by the writer?</td>
<td>Rarely</td>
<td>Yes</td>
<td>Not adequately</td>
<td>No</td>
</tr>
<tr>
<td>Is there intention to deceive?</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Is the relationship visible to the reader?</td>
<td>Not foregrounded; visible as appropriate register</td>
<td>Yes</td>
<td>Sometimes</td>
<td>Possibly</td>
</tr>
</tbody>
</table>

Indirect intertextuality

Indirect intertextuality is the relationship that arises between texts which have commonalities in areas such as topic, purpose, or readership, creating structural and phraseological similarities. Because such indirect intertextual ties are so pervasive, they are not ordinarily signaled by the writer (although on occasion they may be; a small stockpile of phrases exists for this purpose, including as is well known, as the saying goes, and to coin a phrase). For the same reason—because it is so pervasive—it is easy for readers not to note this sort of intertextuality, although its absence, characterized by departures from generic and register conventions (e.g., the use of the first person in a research article in the natural sciences), may be quite obtrusive.

Structurally, many research articles use the familiar divisions of introduction, methods, results and discussion, and research article introductions have been shown to consist of a predictable pattern of moves and steps (Swales, 1990). Our concern here is the similarities among texts within a single discourse which arise at the level of phraseology, rather than text schema. A great deal of research has been directed at establishing the extent to which language follows the idiom principle (Sinclair, 1987) rather than being compositional. The pervasive nature of formulaic language means that texts written in the same language not only share a common pool of words, they share a stock of phrases, multi-word units, and of patterns with open slots with limited possibilities for completion. Hunston and Francis (2000) illustrate this with the example of ‘afraid’ (p. 37 ff.). Used as part of an apologetic phrase, ‘afraid’ can be completed with or without ‘that’:

*I’m afraid that I just wasn’t ready
I’m afraid I’ve only got three*  
( Hunston & Francis, 2000, p. 37)

The choice to include or exclude ‘that’ is a binary one; what follows is an open slot.

Hoey (2005) explains this phenomenon with reference to the psychological notion of priming. As experienced users of a language, we have had repeated exposure to apologies beginning “I’m afraid that,” so that we are primed to produce that phrase in an appropriate situation. It is important to note that this sort of intertextuality therefore comes about indirectly: two apologies resemble each other because the speakers have a shared familiarity with the language of apologies, and not (usually) because one speaker has deliberately copied an earlier apology. The effects of priming apply also to the structure of a text, so that phraseological units are especially likely—or unlikely—in certain portions of texts (Hoey & O’Donnell, 2008).

Somewhat obviously, another sort of similarity among related texts is content: two texts on the same topic are likely to present some of the same ideas, without one drawing directly on the other. There appear, in addition, to be interactions between structure, phraseology, and content which lead to indirect intertextuality (and constitute a discourse). While most studies of phraseological patterns suggest that they are ordinarily short units, Pecorari’s (2008b) investigation of subject-specific texts identified strings as long as 12 words in length, suggesting that similarity of content leads to a
greater likelihood of phraseological similarity. In addition, many chunks were so closely tied to a particular rhetorical function that they would be most likely to occur in certain parts of a text only; for example, “the plates were incubated at” (p. 102), followed by a temperature and an amount of time, is so closely associated with a description of how the research was performed that it would naturally tend to appear in the research methods section of an article.

**Conventional intertextuality**

The other type of legitimate intertextuality is what we are calling conventional intertextuality, and it is distinguished from the first category by virtue of the fact that the relationship with a particular earlier text is a direct one. Writers ordinarily signal these relationships clearly; indeed, it is widely accepted that it is an academic writer’s responsibility to do so. Fully appropriate instances of this category are therefore usually visible to the experienced academic reader.

The most obvious members of this category involve direct reference to another scholar’s work, presented as quotation or paraphrase. This category also accommodates, for example, allusion, which, while not very common in

<table>
<thead>
<tr>
<th>Source</th>
<th>Re-use</th>
</tr>
</thead>
<tbody>
<tr>
<td>no direct source</td>
<td>INDIRECT</td>
</tr>
<tr>
<td></td>
<td>it remains to be seen</td>
</tr>
<tr>
<td></td>
<td>there is little doubt that</td>
</tr>
<tr>
<td></td>
<td>or conventional structure such as</td>
</tr>
<tr>
<td>Stilton, 1985, p. 137: The lunar substrate is undoubtedly caseous in nature, although the precise proportion of curds to whey has yet to be determined.</td>
<td>CONVENTIONAL</td>
</tr>
<tr>
<td></td>
<td>or On the subject of the moon's composition, Stilton states that it is &quot;undoubtedly caseous in nature&quot; (1985, p. 137).</td>
</tr>
<tr>
<td>Stilton, 1985, p. 137: The lunar substrate is undoubtedly caseous in nature, although the precise proportion of curds to whey has yet to be determined.</td>
<td>UNCONVENTIONAL</td>
</tr>
<tr>
<td></td>
<td>or According to Stilton, the moon is made of green cheese.</td>
</tr>
<tr>
<td>Stilton, 1985, p. 137: The lunar substrate is undoubtedly caseous in nature, although the precise proportion of curds to whey has yet to be determined.</td>
<td>DECEPTIVE</td>
</tr>
</tbody>
</table>

![](Fig. 1. Types of intertextuality.)
academic writing, is successful only to the extent that it is understood as allusion. The rather special case of allusion serves as a good example of a general point which is true for all four of these categories, namely there may be uncertainty about which category a specific instance of intertextuality belongs to, but such uncertainty is due to a lack of clarity surrounding specific cases, rather than a lack of theoretical distinction between the categories. Allusion can certainly be attested in academic writing, but is presumably more common in some fields than others. Consider, for example, this opening from a journal article on a literary theme: “It is a truth universally acknowledged that minor writers in search of material sometimes lift passages from their predecessors” (Patrides, 1963, p. 217). The author presumably assumed that his intended readership would recognize the allusion to the opening line of Pride and Prejudice (“It is a truth universally acknowledged, that a single man in possession of a good fortune, must be in want of a wife;” Austen, 1980). Such an opening to an article in (say) nuclear physics would be much more unexpected. The extent to which it is conventional, and therefore legitimate, thus depends in part on the community of readers in which it is produced, but also in part on the tastes and perceptions of the individual reader. The fact that individual readers may hold different views about whether to assign allusion to the category “conventional” is indicative of a degree of uncertainty around the status of allusion, rather than the category.

Unconventional intertextuality

Unconventional intertextuality involves a direct relationship with a given source text, and is not adequately signaled by the writer and in that sense is not legitimate. However, the relationships that fall under this category are not the result of an effort on the part of the writer to be deceptive. Pecorari (2003) has identified three aspects of source use about which a writer needs to be transparent. The first of these is the identity of the source; references should be used to indicate which sources have informed a text. The origin of the language should also be identified. Specifically, the reader should understand whether the language of the source has been adopted (i.e., quotation) or whether ideas from the source have been formulated in a way that the writer arrived at herself (i.e., paraphrase). Finally, the propositional content should be transmitted clearly, and not distorted, so that the reader gains an accurate impression of what the source said. If one or more of these three aspects are not effectively signaled to the reader, then the intertextual relationship is not one which is conventionally appropriate. Such unconventional relationships may easily be missed by the reader, precisely because they are not signaled by the writer. However, it is also possible that an experienced reader may be alerted to the problem by features such as register shift (as, for example, when a copied chunk in appropriate register is inserted into a less polished student text).

Examples of the processes underlying this type of intertextual relationship include poor note-taking strategies, or the rare phenomenon of cryptomnesia (Marsh & Bower, 1993), also called source-memory forgetting. Writers may also omit references or paraphrase insufficiently because they are unsure of the conventions (as our participants pointed out). A related phenomenon arises when writers who are unsure of their abilities to produce appropriate-sounding language mine sources for elegant forms of expression and ideas. As noted above, some gatekeepers consider some instances of this to be not only appropriate, but productive and beneficial, both to the learning process and to the text. However, the examples on which we elicited comments establish that there exists a threshold beyond which textual reuse is not acceptable. (Kerstin, for example, was hesitant to classify one example as plagiarism because ‘small bits’ had been reused, while Lars said it was when ‘longer’ extracts, ‘entire pages’ were taken that text reuse became potentially problematic.) The diversity of response appears, based on the comments of our participants, to be due at least in part to uncertainty about the boundary between unconventional and indirect intertextuality.

Deceptive intertextuality

Deceptive intertextuality also involves a direct relationship with a specific source text. Textually this category is often not easily distinguished from the previous one; the key distinguishing factor here is that the writer intends to deceive the reader about the real relationship between her text and its source. In the present typology, this category thus excludes unintentionally inappropriate source use, paralleling the views of our respondents. Victoria, for example, explained that “plagiarism to me is something when you’re a little more aware of what you’re doing.” In principle any of the three aspects of source use mentioned above may be the focus of this deceptive effort. As a result, the intertextuality is often not visible to the reader, although the reader’s suspicions may be triggered by (for example) a mismatch between the quality of the text and the abilities the writer is perceived to have.
Common examples of deceptive intertextuality include the prototypical forms of plagiarism: buying an essay on the internet, collusion with a classmate jointly to write an assignment and hand it in to separate markers, etc. Other deceptive intertextual practices, such as padding a reference list to give the impression that a text has been informed by a wider range of sources than is actually the case, also belong here.

Summary

Our four categories are designed to allow the diversity of perceptions which is the focus of this issue to be accommodated in a single discourse. We propose that they enable us to account for most if not all instances of intertextuality (Shaw & Pecorari, in preparation). The extensive disagreement among our participants and those in other studies is not due to permeability in these categories, but rather to variability in the ways that judgments are made. There are two factors in particular which lead to differences in judgments. The first of these is the matter of personal evaluation about which textual practices are appropriate and/or conventional. A somewhat trivial example of this is the question of common knowledge, which is widely said, in writing guides and plagiarism policies, to be an example of content for which no citation is needed. As some of our participants noted, though, it is not always easy to distinguish common knowledge. A reader who believes that the uses to which peppermint is put are common knowledge would believe that naming those uses, without a reference, is appropriate, while a reader who thought they were not common knowledge would take the opposite view. In either case, though, they would be using the criterion “common knowledge” to distinguish between the categories indirect and unconventional.

A similar example is the case of citation at second hand. There is broad consensus that it is better to cite primary sources, but that second-hand citation may be unavoidable. Within this consensus there is ample room for disagreement about whether, in a given instance, secondary citation is or is not acceptable. Here, again, disagreement about whether a given instance is conventional does not call the category into question.

The other problematic factor is the question of intention to deceive. It is that which distinguishes deceptive from unconventional intertextuality. Although the criterion itself is straightforward, it can be very difficult to discern the intent by examining the textual product.

Implications for teaching and learning

The participants in our study were all university teachers (at various levels) and teachers are inevitably gatekeepers, both in the sense that they dispense or withhold the academic rewards which permit students to advance to the next level, and in the sense that it is from the ranks of teachers that members of institutional bodies such as disciplinary boards and curricular committees are drawn. As noted by Polio and Shi in this issue, the variety in their reactions to source use can mean that students receive uneven guidance and especially uneven treatment by various powerful people.1

The response needed from the education community to this situation must come not primarily from the classroom, but from higher levels. University teacher education should aspire to make members of the academic community aware of the differences that exist in views across the curriculum. This is important generally, but especially so for those, like many teachers of second-language writing, who teach students from across a range of subject areas and levels. Further, training should aspire to encourage university teachers to problematize their conceptions about this aspect (and others) of academic writing. It is inevitable that students will encounter teachers who vary individually in the features in the work they assess. What is essential is, first, that whatever rules, standards and conventions are applied, they have a solid pedagogical ground, and are not arbitrary; and, secondly, that they are made transparent, to avoid the sense of an “institutional practice of mystery” (Lillis, 1999).

At the levels where policy is made there should, again, in our view, be an awareness of two aspects of the issue: first, of course the tacit assumptions held by particular individuals may not be shared by all members of the university community, as amply demonstrated by our findings and others in this issue; but also, these are assumptions about the assignment of cases to categories of intertextuality, not about the definitions of the categories. This dual awareness

---

1 It should be kept in mind that our participants came exclusively from the “hard” disciplines. Given differences in orientations to many aspects of academic writing across disciplines, including views of intertextuality (Borg, 2009), it seems likely that the variety of responses could be greater still across the university as a whole.
should both shape policy and, most importantly, serve as an impetus to an academy-wide discussion about which criteria are shared, which need not be, and what values underlie them.

What is needed, in other words, is a better understanding of the issues by both teachers and students so that issues can be argued rationally, and parameters can be made explicit. It is painfully clear by now that the simple “plagiarism versus acceptable use” dichotomy on which instructions to students are often based is ineffective. Even a somewhat more elaborated description allowing for patchwriting as a developmental stage seems to us to fail to represent the criteria teachers actually use. As Polio and Shi say (this issue) “plagiarism can be defined only through perceptions and beliefs about what types of textual borrowing are transgressive” but these perceptions and beliefs are elusive and differences can lead to accusations of opening floodgates or victimizing the weak. Our four categories may enable the criteria to be verbalized accurately, and thus discussed within agreed boundaries.

Ultimately, students would benefit from conceptualizing the question of source use at a more abstract level than permissible numbers of repeated words, and in ways nearer their own experience. They would be able to explain and defend their practice and have a vocabulary for being told what was wrong with drafts. The four categories would be especially appropriate as a heuristic in staff training, since they formalize what teachers actually use to make decisions. The varied settings on the familiar parameters create a valuable opportunity for discussion which could in turn lead to insight at the levels of disciplinary expertise that are appropriate for university students.

Future research

This special issue has dealt with intertextuality in academic contexts. But students are not isolated in an academic genre system; in fact they may not clearly perceive the boundaries of academic writing where norms imposed by their teachers apply. Many have experienced business genre systems where boilerplate text is a norm, even a legal requirement. Others may have experienced popular-culture discourse communities in which sampling, allusion, interdiscursivity, or plain pirate-copying are normal.

In general a step forward would be to set teachers’ and students’ perceptions in the context of their entire repertoire of discursive and intertextual practices. In particular, in our view it will be easier to discuss reuse of text with students if we can show how greatly different perceptions among discourse systems can be represented within a uniform framework; that is the aim of our current research.

Appendix A. Example texts

During the research project, these examples were printed in larger type, with one example (student text and source) per page. Participants were presented with them in the order below, one at a time. Words which appear in both texts were printed in red in the examples presented to participants, and appear in bold type here.

Example 1.

Erden

The genus *Mentha* (Lamiaceae) is composed of 19 geographically widespread species and 13 named natural. Peppermint (*Mentha × piperita*) and spearmint are grown world-wide as perennial herbs, and produce different essential oils which are used as flavorings.

*Source*²

*Mentha* is a genus of wide distribution and considerable economic importance. Shoots and leaves of several species are often used as a condiment. The essential oils, which are steam distilled from the herbage, are processed into flavourings for food, medicine, mouthwash, toothpaste and powder, chewing gum, and candy.

---


Example 2.

**Erden**

Reed (1999) reported that 50% of the mint cultures in National Clonal Germplasm Repository (NCGR) under slow growth conditions were lost due to the fungal or bacterial contamination.

**Source**

Some clonal crops are kept in slow-growth storage as in vitro cultures for germplasm conservation (Ashmore, 1997; Engelmann, 1991; Withers, 1991; Withers et al., 1990). Previously, mint cultures held at the National Clonal Germplasm Repository (NCGR) were stored at 4 °C in darkness in 13 mm × 100 mm glass tubes on MS (Murashige and Skoog, 1962) medium. Under these storage conditions, 50% of the cultures were lost to fungal or bacterial contamination (Reed, unpublished data).

Example 3.

**Erden**

The global and world’s largest Musa germplasm collection, 1141 accession, is currently stored at the INIBAP Transit Centre (ITC) under slow growth conditions at reduced temperatures and light intensity (Van den Houwe, I, et al., 1998).

**Source**

The global and world’s largest Musa germplasm collection (1141 accessions) is currently stored at the INIBAP Transit Centre (ITC) (Laboratory of Tropical Crop Improvement, K.U. Leuven, Belgium) under the auspices of FAO. In vitro proliferating shoot-tips are stored under slow-growth conditions at reduced temperatures and light intensity.

Example 4.

**Ingrid**

Brassicas are of major international agricultural significance, for example, as vegetables, animal fodder and oils. These crop species are also diverse in their morphology and development regulation.

**Source**

Brassicas are of major international agricultural significance, for example, as vegetables (cauliflower, Brussels sprouts, cabbages, broccoli), animal fodder (kales and swedess), oils (rapeseed), etc. They are also known to have important anti-cancer activity (BEECHER 1994). These crop species are also diverse in their morphology and developmental regulation.

Appendix B. Interview protocol

The following questions were used to stimulate participants’ discussions of the text examples, though not necessarily in this order, as participants’ responses to early questions often anticipated later ones.

Is this an example of a good way to use and refer to sources?

Why/why not?

(If the answer was that it is problematic, the following questions were asked as well.)

What would you do about it if you found it?

How would you explain to the student that this was a problem?

How would you tell the student what to do to fix it?

Would you take any actions against the student?

Probe specifically about disciplinary action.

Is this plagiarism?

Why/why not?

After all the passages were looked at:

If you imagine a continuum between fully appropriate and entirely unacceptable, where would you place these examples?

What is it that makes number x more/less acceptable than number y?
And follow up with other comparisons.

Probe for these if they don’t emerge:
- quantity in one passage
- quantity overall
- what section it’s in
- only so many ways to say the same thing
- the fact of writing in a second language
- the fact of being new to academic discourse

Now you’ve heard what I’m interested in, is there anything I haven’t asked but should have?
Do you have any questions for me?

References


**Diane Pecorari** is Professor of English Linguistics at Linnaeus University in Växjö, Sweden. Her research interests include discourse analysis and second language writing. She is a member of the editorial boards of *TESOL Quarterly* and *English for Specific Purposes*, and is the author of *Academic Writing and Plagiarism: A Linguistic Analysis*.

**Philip Shaw** has taught English and linguistics in Thailand, Germany, England and Denmark, and is currently a professor in the English Department of Stockholm University. He is interested in the uses of English in a variety of contexts and is currently involved in research projects on aspects of the “parallel-language” situation in Swedish academia. He is co-author (with Gunnel Melchers) of *World Englishes: an Introduction* (second edition 2011).