Try it on: Voice, concordancing and text-matching in doctoral writing

Cally Guerin
School of Education, University of Adelaide
cally.guerin@adelaide.edu.au

Michelle Picard
School of Education, University of Adelaide
michelle.picard@adelaide.edu.au

Abstract

Appropriate use and acknowledgement of sources continues to be a central concern of academic integrity. A major challenge for research students in using sources is the development of a confident authorial voice that matches disciplinary expectations in language use, yet manages to do so without plagiarising through inappropriate text-matching or recycling of language. This is a daunting challenge for all research writers, particularly for English as an Additional Language (EAL) researchers who are still grappling with English grammar and syntax. In order to develop novice research writers' understanding of acceptable use of sources and mastery of disciplinary language, we have developed a process called "Try it on" that uses concordancing software alongside text-matching software (*Turnitin*). Here we present textual analyses of two cases using this process: in one, the student's percentage of matches decreased as he developed his authorial voice; in the second, the percentage of matches increased as the student's language choices came to reflect more closely the expected usage in the discipline, thus replicating the expected authorial voice for that particular audience. These cases demonstrate how "Try it on" can be used to help students write in an appropriate authorial voice while also avoiding plagiarism.

Introduction

Appropriate use and acknowledgement of sources continues to be a central concern of academic integrity. However, this is more complex than simply avoiding the copying associated with plagiarism and patchwriting (Chandrasoma, Thompson, & Pennycook, 2004; Howard, 1995; McCulloch, 2012; McGowan, 2008a, 2008b; Pecorari, 2003; Pennycook, 1996); sources also contain the disciplinary and academic language that writers must necessarily use. Thus novice research writers are forced to navigate a complex set of contradictory expectations before achieving a confident authorial voice. On one hand, they must make an original contribution in their "own words" and their own voice (Stolley & Brizee, 2010); on the other hand, they are required to acknowledge previous research using the standardised, highly formulaic language of their disciplinary discourse community. That is, they must "be original, but not too original" (Picard & Guerin, 2011, p. 221). Most novice research writers find this challenging, but it is particularly difficult for those English as an Additional Language (EAL) researchers who are not yet entirely confident or

comfortable in manipulating English grammar and syntax. In order to develop novice research writers' understanding of acceptable use of sources and mastery of disciplinary language in relation to authorial voice, we have developed a process called "Try it on" that uses concordancing software alongside the text-matching system *Turnitin*. The text-matching capacity of *Turnitin* is increasingly being used as an educational tool (see, for example, Davis & Carroll, 2009; McCarthy & Rogerson, 2009; Rolfe, 2011). However, our process of "Try it on" is innovative in its addition of concordancing software as an extra step in this pedagogy.

In this paper we present textual analyses of two cases where research students used *Turnitin* in conjunction with concordancers to develop an appropriate voice in academic writing. In the first case the student's percentage of matches decreased as he developed his "authorial presence" (Zhao & Llosa, 2008, p. 159) and disciplinary and "situational voice" (Ede, 1992). In the second case, the percentage of matches increased as the student's language choices came to reflect more closely the expected usage in the discipline. We demonstrate how our process of "Try it on" helps students understand how to avoid plagiarism while developing a suitable authorial voice in their writing that recycles the expected disciplinary phrases and expressions.

To match or not to match?

For many novice research writers, the difficulty lies in knowing how much text needs to match other research writing (to meet disciplinary/academic expectations) and how much is too much (and therefore constitutes plagiarism). We demonstrate how *Turnitin* plus concordancers can illuminate this distinction for EAL research students who are in the process of discovering how to write in their own voice within the conventional language use of their discipline. The notion of voice, though, is a slippery one, and requires some exploration.

Voice

Discussions of voice tend to fall into two general categories: those that focus on voice as self-representation and expression of identity (see, for example, Flowerdew, 2000; Gale, 1994; Grobman, 2009; Hirvela & Belcher, 2001; Matsuda & Tardy, 2007; Ramanathan & Atkinson, 1999; Thompson, 2005); and those that focus on the linguistic and rhetorical strategies mobilised to create authorial voice in a text (see, for example, Chang & Schleppegrell, 2011; Harwood, 2005; Ivanic & Camps, 2001; Taş, 2010; Zhao & Llosa, 2008). For our purposes it is the intersection of these approaches that is of interest in terms of supporting doctoral candidates in their writing practices, and allows us to explore how identity is constituted in writing (Aitchison, 2009, 2012; Baker & Lattuca, 2010; Boud & Lee, 2005; Caffarella & Barnett, 2000; Catterall et al., 2011; Kamler & Thomson, 2008; Lee & Boud, 2003). Indeed, it is precisely in order to highlight the academic identity or "academicity" (Petersen, 2007) that is coming into being during doctoral writing that we draw students' attention to the concept of voice.

For all its imprecision as a term (DiPardo, Storms & Selland, 2011), we choose to evoke the notion of voice because "as a metaphor [voice] has to do with feeling-hearing-sensing a person behind the written words, even if that person is just a persona created for a particular text or a certain reading" (Bowden, 1999, quoted in Hirvela & Belcher, 2001, p. 85). That is, it helps our students to notice this aspect of writing that requires them to adopt a particular identity or persona for a given situation (in this case, thesis writing). In their influential paper on mature EAL writers, Hirvela and Belcher (2001, p. 89) take up Ede's (1992) concept of the "situational voice", investigating how writers adopt different voices for different texts and purposes, much as they might put on different outfits for different occasions. A related metaphor is employed by Ivanic and Camps (2001, p. 21), who posit that the choices of "voice types" available in academic writing are pre-determined by the "disciplinary discourse

communities they are entering, like second-hand clothes waiting to be selected and given new life when worn by someone new". Academic writing, then, can be understood as an act not only of putting on different outfits to suit particular occasions, but also of choosing that outfit from a selection of second-hand clothes that have been worn by academics before us. The value of these metaphors lies in their capacity to draw our attention to the ways in which academic texts borrow and recycle language, bringing together our concerns here with both voice and plagiarism.

How is voice constituted in the text?

Ivanic and Camps (2001) map out a detailed framework to explain the concept of voice by building on systemic functional linguistics, delineating between ideational, interpersonal and textual positionings of the authorial voice and articulating the linguistic realisations of each of these categories. On closer inspection of their material, it becomes apparent that "voice" is created out of all the enormous range of language choices made in writing any given text – vocabulary, verb tenses, modality, evaluation, linking, references to other texts and individuals (whether general or specific), classification and structuring of information, and choice of pronouns.

In an attempt to pin down the elements of voice in a manner that could be effectively used to assess voice in student writing, Helms-Park and Stapleton (2003) developed a "Voice intensity rating scale". Working from this, Zhao and Llosa (2008, p. 160) assessed four main components of voice in their study:

1. assertiveness – hedging language, intensifiers;
2. self-identification – pronoun use, active voice;
3. reiteration of central point – frequency and explicitness of presentation of central ideas; and
4. authorial presence and autonomy of thought – presentation of alternative viewpoints, and "reader's impression of the overall authorial presence in a particular piece of writing".

As students start to make progress in developing their own voices, this list can provide a relatively straight forward focus for examining voice in their writing.

Important for our purposes is an awareness of the relationship between author, text and reader. Our EAL students soon discover that their readers and discourse communities can hold "rigid prohibitions against allowing different voices using nonconventional and nonnative forms of language and rhetoric" (Cho, 2004, p. 50). Attempts to publish or submit theses for examination in those "different voices" are rarely accepted on their own terms. While one might well be troubled by the politics and power relations this implies, we would argue that we still owe it to our students to help them understand the conventions in order for them to make informed decisions about how they choose to present their writing. There are certainly examples of successful manipulation of new and different voices in academic writing (see, for example, Viete & Ha, 2007), but this is a complex and often risky business. In terms of the clothing metaphor outlined above, this might be equivalent to wearing a ball gown and tiara to a small, informal academic seminar – beautiful undoubtedly, but sending to onlookers a rather different message about one's purpose from the jeans and shirt more usually adopted for the occasion.

It is useful to put this unreceptiveness to different voices alongside Ivanic and Camps' (2001, p. 31) reminder that academic writing is in fact a "creative recombination of voices". As academic developers we can harness this concept for our students' benefit: "Learner-writers can discuss the aspects of voices they encounter in source texts that they would like to adopt and those they would like to avoid...the role of source texts in learning to write can be acknowledged and used creatively rather than dismissed with the moral outrage associated with 'plagiarism'" (Ivanic & Camps, 2001, p. 31). Thus, it is through a focus on voice that

students can develop aspects of their writing that in turn allow them to participate effectively in their disciplinary and discourse communities.

Why does voice matter?

Although many doctoral candidates take a considerable length of time to develop an appropriate academic voice in their writing, this is an even greater hurdle for EAL students; indeed, it is argued elsewhere that this might in fact constitute a "threshold concept" in doctoral education (Guerin & Green, 2012). Selection of the appropriate levels of formality and assertiveness, signposting of central ideas in the argument, employment of jargon and technical terms, and assertion of authority in texts are not always obvious or transparent to novice writers and require experience and experimentation. Getting this right is crucial, however, in that it facilitates access to a desirable identity as a legitimate member of the targeted discourse community (Cho, 2004; Flowerdew, 2000). Ivanic and Camps (2001, p. 31) explain that, "For those learning to write in a second language, there may be a double demand for critical awareness: firstly, recognizing the extent to which the voice types supported by the new language are culturally acceptable or culturally alien to them, and secondly, recognizing differences between voice types associated with the range of genres and discourses to which they are exposed in the new culture."

These challenges are borne out in the study conducted by Hirvela and Belcher (2001, p. 83), which recognises the importance of existing authorial voices in "mature multilingual writers". Their subjects are much like our EAL research students, who also find themselves in a situation where they are expected to adopt new identities that can in fact be at odds with the self they have previously successfully represented in their academic writing in other languages. For example, as lecturers in their home countries, they presented their ideas with ease and authority to a student audience, but now as doctoral candidates they are expected to justify themselves to examiners; or their previous suitably deferential attitude towards established experts is suddenly regarded as lacking in critical awareness and judgement. These kinds of difficulties in establishing the appropriate academic voice can be particularly irksome when it comes to writing the discussion and analysis sections of the doctoral thesis (Bitchener & Basturkmen, 2006). Thus, EAL research students frequently need direct instruction in how to write about their discipline and their research within the field in ways that match examiners' and reviewers' expectations. Concordancers can be harnessed in particular ways to aid this instruction.

Concordancers

The use of concordancers in corpus study was introduced by English for Academic Purposes teachers in order to explicate lexical and grammatical patterns, and more recently it has been used in direct explorations of disciplinary language patterns (Cargill & Adams, 2005; Cheng, 2008; Conroy, 2010; Yoon, 2011). In highly specialised fields, corpora (bodies of written works) need to be tailor-made for the precise discipline (Lee & Swales, 2006). For novice writers working towards an appropriate academic voice, corpora and concordancers can provide an invaluable source of information about idiomatic or disciplinary use of English. Concordancers are designed to show words in context, and are thus useful for identifying collocations (words that go together) and grammatical patterns. This is particularly important for EAL writers seeking to enter discourse communities that are unsympathetic to the "different voices" mentioned above. At the same time, these writers need to develop a better understanding of the appropriate (that is, unplagiarised) recycling of language elements in the sources they employ. As Eira (2005) has demonstrated, research writers are engaged in a process of "obligatory intertextuality". The difficulty lies in knowing just how much text needs to match (to meet disciplinary/academic expectations) and how much is too much (and therefore constitutes plagiarism).

Our Process – "Try it on"

In preparing ways for our students to learn how to work with disciplinary voices, we are guided by the principles of "learner autonomy" and "scaffolded learning". Learner autonomy is well established as an important element of language learning (Palfreyman & Smith, 2003), and is closely linked to other kinds of independence required of doctoral candidates. However, such autonomy requires initial nurturing if it is to prove robust. A scaffolded approach to writing instruction has been demonstrated as effective for university students in academic English language programs (see, for example, Cotterall & Cohen, 2003), even at doctoral level (Kamler & Thomson, 2006). In breaking down the steps and working with students at each stage, the expectations and requirements of academic writing can be more easily understood, integrated into the writing produced, and later applied in new situations. Together, these approaches provide guided steps towards independent research writing.

To provide a scaffolded approach to encouraging learner autonomy in research writing, we have developed a process that uses *Turnitin* in conjunction with concordancers to help EAL research students develop an appropriate voice in academic writing. At present it is tentatively known as "Try it on", reminding us of the range of matching choices and their recyclability (as well as the persona/clothing being tested out to see how it looks and feels). Below is a brief summary of the process we have described in detail elsewhere (Picard & Guerin, 2011).

Step 1: Text-matching
The first step of Try it on involves writers running their text through plagiarism detection software (e.g., *Turnitin*), and then categorising the text-matches according to type (see Table 1). The document is revised in response to these decisions. For most students, some initial guidance in this process is useful as they learn to interpret the *Turnitin* report. Most importantly, they need to understand that the actual percentage match is not the issue; rather, it is the type of match that is significant. In some cases we have seen a considerable percentage of matches to student papers on unrelated topics which our students have clearly never read. In such cases, the matches are to standard academic phrases of the kind commonly taught to EAL students in English for Academic Purposes classes. For example, students learn some standard strings of words such as "has been used to develop a theoretical framework", "although there are a number of studies", "the existing body of knowledge" and "to the best of my knowledge". A comfortable manipulation of standard, idiomatic phrases is precisely the kind of intertextuality that Pecorari and Shaw (2012) identify as being integral to scholarly writing.

Table 1

Text-match categories (Picard & Guerin, 2011, p. 228)

Type of match	Reason and Response
Too close	The text is too similar to the source and needs to be paraphrased or rewritten.
Not relevant	Some other text has been highlighted, e.g., a formula or a bibliographic reference.
Discipline-specific phrase	This is the way that concept must be expressed in this context.
Unsure	Anything else the student does not know how to categorise.

Concordancers can be used to confirm the author's belief that a phrase does in fact match discipline-specific language and that a particular example can indeed be found in numerous published articles on the subject. It is the last category, "Unsure", that is particularly beneficial in developing the student's voice when taken to the next step where the concordancer programs come into play.

Step 2: Concordancing
Writers can use an existing corpus (a body of written works) such as the Springer Exemplar (available at http://www.springerexemplar.com), or create their own corpus using the articles they have cited in their literature reviews. In the latter situation, students will also need to download concordancer software such as ConcApp© (available for purchase) or AdTAT© (the latter is a freeware program that has been developed at the University of Adelaide: http://www.adelaide.edu.au/red/adtat/). Students then search the corpus for instances of the phrases that have been identified as unoriginal by text-matching software such as *Turnitin*, but which they suspect are legitimately recycled (that is, they might be discipline-specific terms or standard academic expressions). If the phrase appears a number of times, it is likely that, rather than constituting plagiarism, it is the accepted language used by the discipline. On the contrary, if it appears only once, it may indeed be a case of plagiarism.

Developing authorial voice

In what follows, we present two examples of how "Try it on" has been used to respond to different kinds of text-matching issues. In the first, there is too little sense of the student's own voice; in the second, there is too little sense of the required disciplinary language.

Student 1. Too much matching = 48%
The first situation involved a student who exhibited the common errors that result from patchwriting (Howard, 1995) and poor note-taking, in which large chunks of text are copied directly from the source document and there is very little evidence of the author's own contribution to the discussion. The major challenge here was to create a sense of authority in the writing through developing the student's own voice; that is, the student must be encouraged to offer confident opinions and critical assessments of the published ideas, to present him/herself as a researcher entering into dialogue with the discipline (Chanock, 2008; Wisker, 2005). In Zhao and Llosa's (2008) terms listed above, these elements draw on the first of their categories (assertiveness) and the fourth (authorial presence and autonomy of thought).

For this particular student, the first attempt at preparing a doctoral research proposal resulted in a 48% match in the *Turnitin* report, 26% of which was attributed to one article which served as the basis for establishing the research gap to be filled by the doctoral project. However, the other 22% comprised a range of matches to academic publications.

In establishing the research gap, this student relied heavily on a single text that was reproduced in the initial research proposal with only minimal changes or interventions from the student's voice. A major task was for the student to consider what precisely he himself needed to take from this article, and to interweave the voices from other texts into the discussion as well. We advised the student to step back from the secondary sources, put on the "cloak" of authoritative researcher in the discipline, and to prepare an outline of his own proposal that allowed him to identify the central points he wanted to present in his discussion. This process stripped away the reliance on the plagiarised text and created a new structure for the argument that allowed for a more overt authorial presence value-adding to existing debates on the topic.

The revised text reduced the match to the plagiarised article from an initial 26% to 16%. While still recycling more material than is acceptable, the student is gradually

developing an ability to organise his own ideas about the field and to comment on the material presented, as well as going to the original sources, rather than simply copying what the main article had to say about these other studies.

Table 2
Text-matches for Student 1

48% matched text with *Turnitin* – entire paragraph copied	16% matched text with *Turnitin* – student's own voice starting to appear
Ishima et al. (1993a) studied the effect of particle residence time in the layer on dispersion. Three different size classes (42, 72 and 135 um glass beads) were injected in to the gas flow at varying relative particle velocity for change their residence time within the layer. Particle concentration was measured by recording the rate of particles <u>passing through ten LDA measurement volume</u>. The results indicate that increased relative velocity reduces the effect of the large scale vortices on the particles due to smaller residence time within the vortex. Increasing the relative velocity was seen to have the same effect on dispersion as increasing the particle size which led to a modified Stokes number based <u>on a fluid time scale, which includes both the standard fluid time scale</u> …	Ishima et al. (1993a) shows that *the effect of large scale vortices* decreases because of increased relative velocity of particles and *smaller residence time within the vortex.* At this context, *particle concentration was assessed by the data of particle <u>passing through an LDA measurement volume.</u>* Modified Stokes number was described from increased particle size which is the effect on particle dispersion derived from increased relative velocity. This modified Stokes number was defined on the basis of <u>*a fluid time scale, which includes both the standard fluid time scale,*</u> and characteristic residence time…

Italics represent text-matches highlighted by the *Turnitin* report (that is, all of the text presented in the left-hand column). Underlined text draws attention to copied text remaining from the first report.

The left-hand column of Table 2 shows the text copied directly from one source that summarises the experiment performed in another study (that is, taken from the literature review of the key article). In the right-hand column we see that the student has now sought out the paper for himself, and has gleaned from it the points that are relevant to his own discussion. Unnecessary detail has now been eliminated, and the student has digested the material and integrated it into his own argument. The remaining matches are sometimes to disciplinary language confirmed through a concordancer search (e.g., "fluid time scale"), but other matches remain too close to the original source text to be acceptable. Clearly the student still has some distance to go before he is writing in a confident and appropriate academic voice, but we can at least see progress here.

Student 2. Too original = 0%

The second story is that of an EAL doctoral student whose first attempt at writing a research proposal resulted in a 0% matching report from *Turnitin*. While this might be possible, even desirable, in some disciplines, such a report in many Science, Technology, Engineering and Mathematics (STEM) disciplines requires further scrutiny. It is highly likely that the text is in fact "too original" in that it does not use the idiomatic phrases and terms one would expect in standard academic writing. As Pecorari (2009, p. 102) reminds us: "one of the reasons why lexical bundles are of practical interest is that they are part of fluent, native-like expression, and are therefore an important aspect of language learning." Without matching or recycling these phrases, the writing can become awkward, even inaccurate, as well as lacking the intertextuality that is required of scholarly writing (Pecorari & Shaw, 2012).

In this instance, an academic developer worked with the student and supervisor through a number of drafts of the document to identify expressions that did not meet the norms of the discipline. While the text displayed the components of voice outlined by the "Voice intensity rating scale" (Helms-Park & Stapleton, 2003; Zhao & Llosa, 2008), there was still a significant amount of work to be done in terms of vocabulary choices, word forms, verb tenses and phrasing. Examples of the kinds of errors are presented in Table 3 below, comparing the second iteration (which returned a text-match of 2%) and a later version (now up to 13% text-matching).

Table 3

Text-matches for Student 2

Extracts from draft 1 (2% match with *Turnitin*)	**Extracts from draft 2** (13% match with *Turnitin*)	**Reason for match in draft 2**
There are different types of T <u>effector</u> cells.	To date, *different types of T <u>helper</u> cells* have been described	Correct noun has now resulted in a match
The differentiation of the T cell <u>takes place</u> in the thymus	The *differentiation of the T cell <u>begins</u>* in the thy-mus.	Correct verb has now resulted in a match
And any defect in the Th2 cell type leads to atopic asthma and aller-gy. IL4 <u>is majorly respon-sible</u> for Th2 differentia-tion.	And any defect in theTh2 cell type leads to asthma and allergy[ref]. ... The cytokines TGF-β, IL6, IL-21 *and IL-23 <u>play a ma-jor role in</u> the differentia-tion* of naive T helper cells into Th17 cells.	Idiomatic term has now resulted in a match
CD40L is highly ex-pressed <u>in</u> these cells. CD40L <u>activates, prolif-erates and differentiates</u> B cells.	CD40L which is highly expressed <u>on</u> these cells help in the <u>activation, *proliferation and differen-tiation*</u> of B cells.	Nominalisation has now resulted in a match

Italics in the middle column represent text-matches highlighted by the *Turnitin* report. Underlined text draws attention to altered text from the first report leading to text-match.

Throughout the document a range of inappropriate word choices were identified and rectified, including wrong verb or noun, unidiomatic phrases, and use of verbs rather than nominalised forms. Wrong choices in vocabulary can sometimes be the result of translation issues, or attempts to reword information using synonyms, but unfortunately result in unacceptable expression. The same applies to awkward reconstructions of unidiomatic phrases that, in the worst cases, obscure meaning. In such situations, one can sympathise with students who claim that the imperative to express ideas in their own words seems to lead to further complications for the author (Abasi & Akbari, 2008). However, the task here is to develop an authorial voice that more closely resembles the requirements of the disciplinary community.

In this particular case the student was instructed in the language expectations of the discipline as well as some of the more general norms of academic writing, for example, the preference for nominalisation and high levels of lexical density. The outcome was a revised text that more closely matched the expectations of the academic discipline and recycled the language that demonstrated the writer's developing identity as a member of the target discourse community. Concordancing was used to verify the common usage of idiomatic phrases, which now result in the matches indicated in Table 3. Thus, the student's voice as a "legitimate peripheral" (Lave & Wenger, 1991) member of the disciplinary community is gradually coming into being.

Conclusion

When we say we are concerned with students developing their own voices in writing as a way to avoid plagiarism, this can be understood as their process of developing a greater sense of confidence and authority in the persona they create in the text (Hirvela & Belcher, 2001; Matsuda & Tardy, 2007; Ramanathan & Atkinson, 1999; Thompson, 2005). We would argue that this is just as much a matter of learning how closely they need to resemble the standard language expectations of their discipline, as it is a matter of finding new words and phrases to express their ideas. Thus, it is a process of learning to distinguish between the elements of writing that are available for recycling and those which are not - which items of clothing are available from the second-hand rack, and those which must remain in the wardrobes of their original owners.

References

Abasi, A. R., & Akbari, N. (2008). Are we encouraging patchwriting? Reconsidering the role of the pedagogical context in ESL student writers' transgressive intertextuality. *English for Specific Purposes*, *27*, 267–284.

Aitchison, C. (2009). Writing groups for doctoral education. *Studies in Higher Education*, *34*(8), 905–916.

Aitchison, C., Catterall, J. Ross, P., & Burgin, S. 2012. 'Tough love and tears': Learning doctoral writing in the sciences. *Higher Education Research & Development*, DOI:10.1080/07294360.2011.559195.

Baker, V. L., & Lattuca, L. R. (2010). Developmental networks and learning: Toward an interdisciplinary perspective on identity development during doctoral study. *Studies in Higher Education, 35*(7), 807–827.

Bitchener, J., & Basturkmen, H. (2006). Perceptions of the difficulties of postgraduate L2 thesis students writing the discussion section. *Journal of English for Academic Purposes, 5*, 4–18.

Boud, D., & Lee, A. (2005). "Peer learning" as pedagogic discourse for research education. *Studies in Higher Education, 30*(5), 501–516.

Catterall, J., Ross, P., Aitchison, C., & Burgin, S. (2011). Pedagogical approaches that facilitate writing in postgraduate research candidature in science & technology. *Journal of University Teaching & Learning Practice, 8*(2), 1–10.

Caffarella, R. S., & Barnett, B. G. (2000). Teaching doctoral students to become scholarly writers: The importance of giving and receiving critiques. *Studies in Higher Education, 25*(1), 39–52.

Cargill, M., & Adams, R. (2005). Learning discipline-specific research English for a world stage: A self-access concordancing tool? *Higher Education in a Changing World: Proceedings of the HERDSA International Conference*, Sydney, pp. 86–92.

Chandrasoma, R., Thompson, C., & Pennycook, A. (2004). Beyond plagiarism: Transgressive and nontransgressive intertextuality. *Journal of Language, Identity and Education, 3*(3), 171–193.

Chang, P., & Schleppegrell, M. (2011). Taking an effective authorial stance in academic writing: Making the linguistic resources explicit for L2 writers in the social sciences. *Journal of English for Academic Purposes, 10*, 140–151.

Chanock, K. (2008). When students reference plagiarised material – what can we learn (and what can we do) about their understanding of attribution? *International Journal for Educational Integrity, 4*(1), 3–16.

Cheng, W. (2008). Concgramming: A corpus-driven approach to learning the phraseology of discipline-specific texts. *CORELL: Computer Resources for Language Learning*, Vol. 1.

Cho, S. (2004). Challenges of entering discourse communities through publishing in English: Perspectives of nonnative-speaking doctoral students in the United States of America. *Journal of Language, Identity and Education, 3*(1), 47–72.

Conroy, M. (2010). Internet tools for language learning: University students taking control of their writing. *Australasian Journal of Educational Technology, 26*(6), 861–882.

Cotterall, S., & Cohen, R. (2003). Scaffolding for second language writers: Producing an academic essay. *ELT Journal, 57*(2), 158–166.

Davis, M., & Carroll, J. (2009). Formative feedback within plagiarism education: Is there a role for text-matching software? *International Journal for Educational Integrity, 5*(2), 58–70.

DiPardo, A., Storms, B. A., & Selland, M. (2011). Seeing voices: Assessing writerly stance in the NWP Analytic Writing Continuum. *Assessing Writing, 16*, 170–188.

Ede, L. (1992). *Work in progress: A guide to writing and revising* (2nd ed.). New York: St. Martin's Press.

Eira, C. (2005). Obligatory intertextuality and proscribed plagiarism. *2nd Asia-Pacific Educational Integrity Conference.* Newcastle, Australia, 2–3 December.

Flowerdew, J. (2000). Discourse community, legitimate peripheral participation, and the nonnative-English-speaking scholar. *TESOL Quarterly, 34*(1), 127–150.

Gale, X. L. (1994). Conversing across cultural boundaries: Rewriting "self". *Journal of Advanced Composition, 14*, 455–462.

Grobman, L. (2009). The student scholar: (Re)negotiating authorship and authority. *College Composition and Communication, 61*(1), 175–196.

Guerin, C., & Green, I. (2012). Voice as a threshold concept in doctoral writing. In *Narratives of Transition*. Conference Proceedings of 10th Quality in Postgraduate Research Conference, Adelaide, South Australia. http://qpr.edu.au.

Harwood, N. (2005). "Nowhere has anyone attempted . . . In this article I aim to do just that": A corpus-based study of self-promotional I and we in academic writing across four disciplines. *Journal of Pragmatics, 37*, 1207–1231.

Helms-Park, R., & Stapleton, P. (2003). Questioning the importance of individualized voice in undergraduate L2 argumentative writing: An empirical study with pedagogical implications. *Journal of Second Language Writing, 12*, 245–265.

Hirvela, A., & Belcher, D. (2001). Coming back to voice: The multiple voices and

identities of mature multilingual writers. *Journal of Second Language Writing, 10*, 83–106.

Howard, R. M. (1995). Plagiarisms, authorships, and the academic death penalty. *College English, 57*(7), 788–806.

Ivanic, R., & Camps, D. (2001). I am how I sound: Voice as self-representation in L2 writing. *Journal of Second Language Writing, 10*, 3–33.

Kamler, B., & Thomson, P. (2006). *Helping doctoral students write: Pedagogies for supervision*. London & New York: Routledge.

Kamler, B., & Thomson, P. (2008). The failure of dissertation advice books: Toward alternative pedagogies for doctoral writing. *Educational Researcher, 37*, 507–514.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Lee, A., & Boud, D. (2003). Writing groups, change and academic identity: Research development as local practice. *Studies in Higher Education, 28* (2), 187–200.

Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes, 25*, 56–75.

Matsuda, P. K., & Tardy, C. M. (2007). Voice in academic writing: The rhetorical construction of author identity in blind manuscript review. *English for Specific Purposes, 26*, 235–249.

McCarthy, G., & Rogerson, A. (2009). Links are not enough: Using originality reports to improve academic standards, compliance and learning outcomes among postgraduate students. *International Journal for Educational Integrity 5*(2),47–57.

McCulloch, S. (2012). Citations in search of a purpose: Source use and authorial voice in L2 student writing. *International Journal for Educational Integrity, 8*(1), 55–69.

McGowan, U. (2008a). International students: A conceptual framework for dealing with unintentional plagiarism. In T. S. Roberts (Ed.), *Student plagiarism in an online world: Problems and solutions* (pp. 92–107). Hershey, New York: Information Science Reference.

McGowan, U. (2008b). *Avoiding plagiarism: Achieving academic writing.* Retrieved January 12, 2011, from http://www.adelaide.edu.au/clpd/plagiarism/.

Palfreyman, D., & Smith, R. C. (Eds.). (2003). *Learner autonomy across cultures: Language education perspectives*. Basingstoke & New York: Palgrave Macmillan.

Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second language writing. *Journal of Second Language Writing, 12*(4), 317–345.

Pecorari, D. (2009). Formulaic language in Biology: A topic-specific investigation. In M. Charles, S. Hunston & D. Pecorari (Eds.), *Academic writing: At the interface of corpus and discourse* (pp. 91–104). London: Continuum International Publishing.

Pecorari, D., & Shaw, P. (2012). Types of student intertextuality and faculty attitudes. *Journal of Second Language Writing*, *21*, 149–164.

Pennycook, A. (1996). Borrowing others' words: Text, ownership, memory, and plagiarism. *TESOL Quarterly, 30*(2), 201–230.

Petersen, E. B. (2007). Negotiating academicity: Postgraduate research supervision as category boundary work. *Studies in Higher Education, 32*(4), 475–487.

Picard, M., & Guerin, C. (2011). "Be original, but not too original": Developing academic voice through innovative use of text-matching and concordancing software. In C. Nygaard, N. Courtney & C. Holtham (Eds.), *Beyond transmission: Innovations in university teaching* (pp. 221–234). Faringdon, UK: Libri.

Ramanathan, V., & Atkinson, D. (1999). Individualism, academic writing, and ESL writers. *Journal of Second Language Writing, 8*(1), 45–75.

Rolfe, V. (2011). Can *Turnitin* be used to provide instant formative feedback? *British Journal of Educational Technology*, *42*(4), 701–710.

Stolley, K., & Brizee, A. (2010). Avoiding plagiarism: Overview and contradictions. Purdue Online Writing Lab. Retrieved January 12, 2011, from http://owl.english.purdue.edu/owl/resource/589/01/.

Taş, E. E. I. (2010). "*In this paper I will discuss…*": Current trends in academic writing. *Procedia Social and Behavioral Sciences, 3*, 121–126.

Thompson, C. (2005). "Authority is everything": A study of the politics of textual ownership and knowledge in the formation of student writer identities. *International Journal for Educational Integrity, 1*(1), np.

Viete, R., & Ha, P. L. (2007). The growth of voice: Expanding possibilities for representing self in research writing. *English Teaching: Practice and Critique, 6*(2), 39–57.

Wisker, G. (2005). *The good supervisor: Supervising postgraduate and undergraduate research for doctoral theses and dissertations*. Basingstoke and NY: Palgrave Macmillan.

Yoon, C. (2011). Concordancing in L2 writing class: An overview of research and issues. *Journal of English for Academic Purposes, 10*, 130–139.

Zhao, C. G., & Llosa, L. (2008). Voice in high-stakes L1 academic writing assessment: Implications for L2 writing instruction. *Assessing Writing, 13*, 153–170.

About the Authors

Dr Cally Guerin is an academic developer in the Researcher Education and Development Unit in the School of Education, University of Adelaide, where she runs workshops for doctoral students and their supervisors. Her research interests are in the areas of doctoral education, thesis writing groups, research supervision, and the multicultural university. She is a regular contributor to the doctoral writing blog site at http://doctoralwriting.wordpress.com/ .

Dr Michelle Picard is a senior lecturer at the School of Education at the University of Adelaide. She is Director of the Researcher Education and Development Unit and coordinates the MEd program. Her research focus is the development of high level academic language and research skills in postgraduate students. Michelle publishes in the fields of Academic Language and Learning and Higher Education. She has lectured in Australia, South Africa, the United Arab Emirates, and the Sultanate of Oman.