The impact of an EAP course on postgraduate writing

Neomy Storch*, Joanna Tapper

School of Languages & Linguistics, The University of Melbourne, 5th floor, Arts Centre, Melbourne, Victoria 3010, Australia

Abstract

This study assessed the impact of completing an English for Academic Purposes (EAP) course on the writing of postgraduate[1] learners. We begin this paper by describing a course offered for credit to postgraduate international students at a university in Australia, and then report on a large-scale study ($n = 69$) which investigated the improvement (if any) in the writing of students enrolled in this course. The students undertook the course at the beginning of their postgraduate studies at the university and concurrently with their degree programs. Two cohorts were identified: students with an IELTS score on entry to the University of 6.0 (or equivalent) and those with a score of 6.5. Data for the study were the in-class writing produced by the students at the beginning and at the end of semester, and a short questionnaire. The students' texts were analysed using a range of measures for fluency, linguistic accuracy and use of academic vocabulary. Texts were also analysed for overall structure, cohesion and coherence. The study found improvements for the entire group and for the two cohorts mainly in terms of accuracy, use of academic vocabulary, and structure. We discuss these findings in relation to key pedagogical features of the course.
© 2009 Elsevier Ltd. All rights reserved.

Keywords: Postgraduate EAP course; Assessing writing; Feedback

1. Introduction

In recent years there has been a dramatic increase in the number of international students in Australian universities, reflecting a trend observed in many other universities in English-speaking countries. For most of these international students, English is their second language, and in 2007, 44% were postgraduates (Studies in Australia, 2007). Given the large number of international students, there has been an increased awareness of the difficulties these students face when studying in their second language. In response to these perceived difficulties, Australian universities have established a range of ESL programs (see Melles, Millar, Morton, & Fegan, 2005). At the University of Melbourne these programs include individual tutorials, workshops and credit-bearing English for Academic Purposes (EAP) courses.

Many universities have also developed EAP courses specifically designed for postgraduate students (e.g. Cargill, Cadman, & McGowan, 2001; see also list in Melles et al., 2005), in recognition of the difference in tasks and skills required of postgraduate compared to undergraduate students. These courses, on the whole, aim to develop the

---

* Corresponding author. Tel.: +61 3 83445208.
  E-mail address: neomys@unimelb.edu.aua (N. Storch).
[1] The Australian term 'postgraduate student' corresponds to the US 'graduate student'. Both terms are used interchangeably in this article.

learners' writing skills, including development of more sophisticated and accurate levels of academic language use. A number of published papers provide a descriptive account of different courses and approaches adopted (e.g. Cargill et al., 2001; Casanave, 2003; Dudley-Evans, 1995; Frodesen, 1995; Silva, Reichelt, & Lax-Farr, 1994). However, to date there has been little research which has attempted to systematically assess the impact these programs and courses have had on the students' academic writing.

This study investigated the impact that a credit-bearing EAP course, developed specifically for international postgraduate students and titled Presenting Academic Discourse (PAD), had on the students' writing. The paper begins with a description of the EAP course and the rationale for the instructional approach adopted and then proceeds to outline the study, including the research questions that guided the study.

2. Presenting Academic Discourse (PAD)

This course aims to develop the advanced language and academic skills required for successful postgraduate study in English. The skills focused on include effective research skills, critical reading and writing, the ability to plan and confidently give an oral presentation on a research topic, and the ability to write a research proposal (which includes a review of the literature) fluently, accurately and using appropriate academic language.

The teaching approach adopted is best described as task-based (Skehan, 2003). The assessment tasks form the organising principle of the syllabus, and include:

- a summary of one text (500 words),
- a critical summary comparing two texts (1000 words),
- a proposal for a research project (2500–3000 words), and
- an oral seminar presentation based on the proposal (7–10 minutes long).

These tasks were chosen because they have been identified as authentic tasks required of postgraduate students, particularly in terms of the writing and critical engagement expected of graduate students (Belcher, 1995; Casanave & Hubbard, 1992; Hood, 2008). The approach is also student-centred. Students choose topics for their assignments which are directly related to their interests and study areas, and are encouraged to sustain the same topic throughout the assessment tasks so as to build cumulatively towards an extended understanding of an issue.

The course draws on the work of a number of scholars on composing processes (e.g. Bereiter & Scardamalia, 1987; Tribble, 1996), on genre approaches to the teaching of academic writing (e.g. Cheng, 2008; Hyland, 2004, 2007), and research showing the importance of feedback and revision to the composing processes, particularly for the development of graduate literacy (Leki, 2006). For each major assessment task students write a first draft for review by the instructor, after which the final version is submitted. The first draft is annotated by the students (see Storch & Tapper, 1997) which enables the students to articulate their concerns. Feedback is given mainly on structure and language. Since the EAP instructors are not discipline specialists, they give content feedback only if the claims lack clarity or are insufficiently developed. Only the final draft is graded. Common errors in our students' writing are often discussed in class (e.g. errors in use of articles in generic statements or inappropriate adverb/verb collocations), with examples taken from the students' writing and suggested reformulations elicited from the students or provided by the teacher.

The course is one semester long (12 weeks), with class time comprising two seminars (of 1 ½ hours each) per week. In the seminars, activities focus on text analysis and various language exercises. Model texts are analysed for structure and language use and then students are asked to analyse texts from their own disciplines (see Hood, 2005; Swales & Feak, 2000, 2004). For example, in teaching students how to write a critical review, students are presented with an authentic review and asked to analyse it for content (i.e. what aspects of the study are critiqued), how the review is structured, and the language used in the review (e.g. verbs of attribution, use of unreal conditional, phrases used to evaluate a study). Students are then asked to analyse and report to class on a review from their own discipline-specific area of study (see Casanave, 2003; Hirvella, 2004; Hyland, 2007).

A range of exercises is used to draw learners' attention to linguistic choices (lexical and grammatical) and form-meaning links. For example, students are asked to summarise a short text in groups; we then look at different versions produced in class, and subsequently at a model produced by the teacher. Grammatical and lexical choices are discussed with a focus on what is considered an appropriate formal, academic register, changing trends in academic writing, similarities and differences in discipline-specific writing and conventions (e.g. in the use of passive voice, first

person pronoun, see Chanock, 1997; Hinkel, 2002; Hyland, 1999a,b), and the undesirability of overly lengthy and complex sentences. Students also practice skills such as synthesizing information and editing individually and in pairs. The course reader contains materials on academic language use (e.g. different verbs of attribution, verb tense/aspect choice in the literature review sections) and model texts.

Two separate streams of the course are available: a General and an Engineering stream. Enrolments in the course have increased steadily since its inception in 1999. In recent years enrolments have ranged from 60 − 90 students per semester, with the largest number of students enrolling in the Engineering stream. This is mainly because the Faculty of Engineering accepts students with an International English Language Testing System (IELTS) score of 6 (0.5 below the IELTS score required for entry to the University) on condition that they successfully complete PAD, concurrently with their degree program in the first semester; that is, the beginning of their study at the university.

3. Rationale for this study

Students' feedback comments on the course have always been very positive, and the growing number of students enrolling in the course provides further evidence of its success. Findings from a study conducted by the Faculty of Engineering (Leahy, 2004) seem to indicate that the course prepares students well for academic study. The study compared the progress of students enrolling as conditional entry (IELTS 6.0 + successful completion of PAD) with the progress of those entering unconditionally (IELTS 6.5). Using grade point averages, the study found no difference between the two cohorts in students' progress. Thus the available evidence suggests that the course is well received and successful in its aim of developing students' EAP. However, the EAP staff were also keen to investigate the direct impact, if any, the course has on the students' L2 writing development. We focused on writing because writing has been reported to be a major concern for postgraduate international students, their lecturers and supervisors (Cargill, 1996; Casanave & Hubbard, 1992). Writing also forms the major focus of PAD seminars and the assessment tasks.

As mentioned earlier, although there are a number of studies which have described different courses specifically designed to assist international postgraduate students with their writing, there is little available research on the impacts of these courses on students' writing. Studies which have investigated the effects of courses offered in English medium higher institutions on the development of learners' language skills have produced somewhat mixed results. Some studies report no gains (e.g. Read & Hays, 2003), whereas others report some gains in overall proficiency (e.g. Elder & O'Loughlin, 2003) or in writing proficiency (e.g. Green & Weir, 2003). It should be noted, however, that most of these studies have focused on test-preparation types of EAP courses, and gains were reported in terms of increases in average band scores. The problem with band scores, such as band scores on writing, whether global (i.e. overall writing proficiency) or analytic (e.g. scores on one criterion of writing such as 'lexical resources') is that they tend to collapse several features of writing under a single rubric (Purpura, 2004; Storch, in press). Thus improvements reported in terms of higher band scores do not necessarily reveal what aspects of writing led to the improvements.

The study by Shaw and Liu (1998) used a range of linguistic measures to compare international students' writing before and after they completed a full-time EAP course (prior to university study) in the United Kingdom. They reported that the learners' writing became more formal, employing language associated with written rather than spoken register, but showed no significant changes in terms of linguistic accuracy and complexity. Similarly, a small-scale study by Storch (forthcoming), which compared the writing of international students (undergraduates and postgraduates) at the beginning and end of semester, found that the students' writing showed improvement only on structure and register but not on linguistic accuracy or complexity. The students were studying in degree programs but did not participate in any EAP support programs. However, Polio, Fleck, and Leder (1998) found that the linguistic accuracy of students taking a preparatory EAP writing class improved after a semester (15 weeks) of study and attribute this improvement to the feedback learners received on their writing throughout the semester.

The present study investigated what aspects of academic writing improved, if any, at the completion of one semester of studying an EAP course that was specifically designed for postgraduate students. In this study we were particularly interested in tracing a possible relationship between features of the course and writing development. We focused on both linguistic and rhetorical aspects of writing; that is, we investigated whether over the semester, learners' writing became more fluent, accurate and displayed a greater command of academic vocabulary as well as whether it improved in terms of overall structure, cohesion and coherence. We decided not to investigate changes in linguistic complexity given the relative brevity of the study period (10 weeks) and because the focus of the course was on accuracy and academic language use, and not on writing more complex sentences. Ortega (2003) has shown that

linguistic complexity takes longer than 12 weeks to develop. As she also points out, more complex sentences do not necessarily equate with what is considered good writing. The duration of our study is thus a limitation, yet was necessary because the course being investigated is itself of brief duration (12 weeks). Another difficulty was to develop measurements of student progress in writing in an academic environment where there is no further direct assessment by the university of these students' proficiency in English.

Thus the research questions guiding this study were:

1. What aspects of the students' writing, if any, improved at the completion of the PAD course?
2. If improvements had occurred, could they be linked to what was taught in PAD?

4. Study design

The study used a test/re-test design, based on a writing task required of all students during class. During the first week of the semester, students wrote in class on the topic: 'Describe an important issue or concept or process that is relevant to your field of study.' A common topic was not set, since a basic principle of PAD is that students draw on their content knowledge when writing academic texts. Examples of topics chosen by the students included 'language development in lingua franca', 'preventing HIV/AIDs among injecting drug users in Indonesia', and 'information technology security'. The time allowed for the task was 30 minutes. This type of task is often used in our EAP classes as a very useful way for instructors to diagnose students' abilities in academic writing at the beginning of a semester. The task was returned to the students in the following week with minimal feedback comments. Photocopies of the students' written text were kept by the researchers.

In week 10 of the same semester, participants completed short questionnaires (administered by one of the researchers who was not involved in teaching the course). There were six questions altogether. Students were asked about previous tertiary study, current study program, work experience, and scores on English language tests. They were also asked about the writing they were doing concurrently in their study programs. These questions aimed to elicit information about the students' English language use and proficiency. The questionnaire also included one open-ended question asking students whether taking PAD had been useful for their academic writing. In the second seminar of week 10, all students again wrote in class on the week 1 topic. The use of the same topic is frequently used in pre-post test research designs (e.g. Elder & O'Loughlin, 2003; Shaw & Liu, 1998; Storch & Wigglesworth, 2007) to avoid the possible effect of different topics on the quality of writing. A list with students' names and week 1 topics was circulated to remind students of their topics.

4.1. Participants

Out of a total of 94 postgraduates taking the course, there were 69 participants in this study, of whom 16 were enrolled in the General stream and 53 in the Engineering stream. They came from a range of language and cultural backgrounds, with the majority (46) from PR China.

The majority of the participants (65 out of 69) were in the first semester of their postgraduate programs. The postgraduate programs included PhD, master with minor thesis, and master coursework-only degrees. The majority had entered with IELTS scores of 6 (or equivalent), as shown in Table 1.[2] For this research, the 18 participants with IELTS or equivalent scores of 6.5 or above were categorised as 'high proficiency'; and the 34 with IELTS of 6 were 'low'.[3]

5. Data analysis

The main source of data for this study was the written texts (scripts) completed in class in week 1 (time 1) and week 10 (time 2). These scripts were analysed for language use, and text structure and rhetorical quality. In analysing for

---

[2] Other proficiency tests such as TOEFL are acceptable for entry to the University, but IELTS is the main test used.
[3] The majority of the 17 participants who did not disclose their English scores probably had scores of IELTS 6, but we did not place them in the 'low' group as we did not have their scores.

Table 1
Participants' IELTS scores and areas of study.

IELTS (or equivalent) score	Number	Area of study
High: IELTS 6.5 or above	18	PhD (5) in Applied Linguistics, Botany, Economics, Medicine
		Master of Applied Linguistics (4); Public Health (4); Engineering/IT (5)
Low: IELTS 6	34	Master of Telecommunications (7); Engineering Project Management (6); IT (6); other Engineering/IT >degrees (13);
		Public Health (1); Population Health (1)
Unknown	17	Master of IT (6); Telecommunications (5); other Engineering/IT (6)

language use, we utilised measures of writing fluency, accuracy, and academic vocabulary. To measure for fluency and accuracy, all scripts were coded for length (in words), T-units, clauses, and errors. A T-unit consists of one independent clause and any dependent clauses attached to it (Hunt, 1966). The following is an example of a T-unit from our data (a participant's script at time 1). This T-unit is made up of a dependent clause and an independent clause, separated by a slash.

Because there are too many people looking for water/these wells becomes (sic) dry after one or two months.

5.1. Fluency

Fluency was measured in terms of the number of words and words per T-unit. Although Ortega (2003) argues that words per T-unit is a measure of complexity rather than fluency, others, such as Wolf-Quintero, Inagaki, and Kim (1998), based on an extensive review of studies, categorise it as a measure of writing fluency; reflecting the length of production units, rather than the structural complexity of the units. The word count tool of the software program Office Word was used. In counting words, we excluded titles students gave to their texts but included incomplete sentences at the end of unfinished texts.

5.2. Accuracy

In order to assess accuracy, all scripts were coded for errors. We experienced similar difficulties in identifying, classifying and counting errors as those described by Mayor, Hewings, North, Swann, and Coffin (2007) in their research on features of high- and low-scoring IELTS scripts. When the writers' meanings were unclear, we used the preceding and following parts of the texts to identify and classify errors; therefore, unfinished sentences at the end of texts were not coded for errors. We used error categories based on Bardovi-Harlig and Bofman (1989) (see Appendix A for categories and examples). The categories included errors in syntax (e.g. errors in word order, missing elements), morphology (e.g. verb tense, subject-verb agreement), grammar (e.g. use of articles, prepositions), and in lexis (word choice). After developing the coding system, we each analysed errors in the same eight scripts in order to reach agreement on coding and counting. At that stage because of orthography difficulties, we decided to exclude spelling errors, and to count punctuation as an error only when meaning was affected.

A range of accuracy scores was calculated: a ratio of error free T units per total T-units (EFT/T), a ratio of error free clauses per total clauses (EFC/C), and the total number of errors per total number of words (E/W). The last measure (E/W) addresses the concern that some researchers (e.g. Bardovi-Harlig & Bofman, 1989) have raised about using only ratio scores to assess learners' accuracy: that ratio scores do not distinguish between T-units or clauses that contain multiple errors from those which contain only a single error.

5.3. Use of academic vocabulary

Use of vocabulary, in terms of both range and appropriacy, is an important aspect of academic writing. However, the analysis of vocabulary used in writing by non-native speakers of English is a challenge to researchers (see Duran, Malvern, Richards, & Chipere, 2004; McCarthy & Jarvis, 2007). Mayor et al. (2007, p. 254) report that it was difficult to find a "revealing approach to the range and delicacy" of vocabulary in the IELTS writing tasks that they analysed,

or to find close links between lexical features and high and low scores. Hinkel's (2003) study of the academic writing of nonnative English speaking university students identified the occurrence of certain categories of nouns and verbs. Most studies investigating lexical complexity use students' writing on the same given topics and scripts with at least 300 words. In contrast, our data were purposely gathered from students who chose their own topics, according to their varied fields of study and interest, and in some cases were less than 300 words long.

We therefore decided to investigate vocabulary use in the scripts by using the Academic Word List (AWL) developed by Coxhead (2000). Coxhead's AWL derives from work by Nation and others on the nature of lexis found in academic texts. An academic text contains four categories of word families (Nation, 2001). The largest category is *high frequency words* that comprise about 80% of the running words in a text ('running' words, or 'tokens', include repeat uses of the same words, as opposed to 'types', which are the individual words). According to Nation (2001), for learners of English, these high frequency words are among the first 2000 words they are likely to encounter at university. The *technical* category includes words closely related to specific study areas (about 5% of academic texts). A third category of *low frequency words* (about 5% of an academic text) includes thousands of rare words, and technical words from areas other than a learner's own technical area. The final category consists of *academic words* that are found across disciplines (but are not among the high frequency words) and comprises 9 to 10% of an academic text. Hyland and Tse's study (2007) re-examined the frequency of Coxhead's list in an expanded corpus which included student texts (predominantly theses and dissertations). They found that the AWL covered 10.6% of their entire corpus.

Academic words are said to pose a challenge for nonnative English speaking students because of their lack of salience and relative low frequency. Unlike technical or discipline specific words, academic words (e.g. analyse, inherent) play a supportive rather than a central role in construing the meaning of a text (Coxhead, 2000). However, they are very important; for example, in describing and evaluating empirical studies (Nation, 2001). Furthermore, their low frequency means that they are unlikely to be learnt incidentally (Worthington & Nation, 1996).

It was impractical to analyse for technical, low frequency and high frequency words in scripts that were written across disciplines on a variety of topics, but the AWL seemed a promising way to analyse academic vocabulary development. The AWL of 570 word families was derived from a corpus of academic texts drawn from four 'sub-corpora' from arts, commerce, law and science (see Coxhead, 2000 for details). Coxhead reports that the AWL accounts for 12% of the commerce subcorpus, 9.4% for law, 9.3% for arts and 9.1% of the science subcorpus. Given that our students came mainly from the field of engineering (see Table 1), and that the AWL words achieved the lowest coverage in the science subcorpus, we reasoned that a frequency of 9% of tokens from the AWL could be set as an ideal benchmark for these postgraduate students to attain, regardless of their discipline. At the same time, we acknowledge that this figure of 9% is a fairly rough guide, given the variations that are found in AWL's coverage across and within disciplines (Hyland & Tse, 2007).

Each student script was checked for the presence of words on the AWL. The numbers of AWL types and tokens were counted for each student at times 1 and 2, and the number of tokens was also calculated as a percentage of the words written at times 1 and 2. In order to avoid distortions due to occurrences of some high-frequency words, we decided to use a cut-off point when counting tokens. To establish this cut-off, we looked at the word *technology*, which was the item from the AWL that occurred most frequently in participants' writing. The range of times it occurred was 0–9; so we chose 5 as the cut-off for counting repetitions. Thus, we counted as 5 tokens any AWL item that occurred more than five times in a text.

5.4. Text structure and rhetorical quality

In addition to analysing the scripts by using a range of linguistic measures, we analysed participants' writing for overall structure, coherence and cohesion, and content. Given the brevity of the writing and varied topics, the criteria for evaluating these aspects of writing had to be simple and flexible enough to enable us to trace changes from time 1 to 2. We consulted the IELTS Task 2 writing band descriptors (no date) and the PAD writing assessment scheme, but both proved inappropriate. Task 2 in IELTS is an argumentative essay, and the PAD assessment scheme was designed to assess longer, research-based texts. Thus we developed a guide to analyse text structure and rhetorical quality, shown in Appendix B. Using this guide, each researcher assessed the scripts of five students (i.e. 10 scripts altogether) with time 1 and time 2 identification removed; inter-rater reliability was 100%.

The comments on the guide were then used to decide if the script at time 2 showed overall improvement compared with time 1. A brief summary was derived from the comments to provide a rationale for the decision made. A sample of one student's scripts is given in Appendix C, together with the corresponding quantitative and qualitative analyses.

6. Results

The scripts ranged over many topics and varied in length, accuracy, and rhetorical quality. The results of our analyses are presented for the entire data set ($n = 69$); for the two proficiency cohorts: high proficiency (IELTS or equivalent test score of $\geq.5$) ($n = 18$) compared to low proficiency cohort (IELTS or equivalent score of 6) ($n = 34$); and across times 1 and 2. The 17 participants whose English scores were not reported are omitted from high/low comparisons, but are included in the 'entire set' figures for general comparisons of times 1 and 2.

6.1. Language use: fluency

Table 2 shows that the scripts produced at time 2 were on average slightly shorter for the entire set, but the difference was not statistically significant (t (68) = 1.786, $p = .079$). The average number of words per T-unit remained fairly stable over time for the entire set. Given the similarity of scores, no statistical tests of significance were performed. The average W/T for the high group dropped, while it rose for the low group. However, Ortega's (2003) meta-analysis shows that mean differences of less than four words on W/T are not considered statistically significant for such cohort sizes.

6.2. Language use: accuracy

Errors were categorised and counted for each student's scripts at times 1 and 2. Accuracy scores and distribution of errors are summarised in Table 3, which shows an improvement in accuracy scores, reflected in all measures of accuracy for the entire set. Although the difference between time 1 and 2 on the measure EFT/T was not statistically significant, it was statistically significant for the difference in E/W (t(68) = 2.505, $p = .015$), and very close to significance for EFC/C (t (68) = -1.882, $p = .064$).

A two-way ANOVA analysis revealed that there was an effect for time and for groups but no interaction effect. That is, although both groups improved over time, there was no statistically significant difference in terms of gains between the two groups. Tables 4–6 show the results of the two-way ANOVA analyses for all three measures of accuracy.

The most frequent types of errors at both time 1 and 2 were grammatical and derivational (e.g. use of articles, plural −s, verb tense/form, subject/verb agreement, word form), and lexical (incorrect word choice). We found that for some individuals, the number of errors at times 1 and 2 was similar, while for others, errors had increased or decreased, and the categories of errors might change also.

Table 7 shows all types of errors found in the scripts at times 1 and 2. Overall, the raw number of errors in 17 error categories decreased at time 2. The top four types of errors persisted at time 2, with articles being the most frequent source of errors (see Table 7). In her large-scale study, Ferris (2006) also found that despite improvements in overall accuracy following corrective feedback, errors in articles persisted.

Table 2
Results for fluency for the entire set, high and low groups.

Fluency	Time 1		Time 2	
	M	SD	M	SD
Total words				
Entire set	216.30	73.40	200.90	58.77
High proficiency	225.28	49.13	221.11	57.25
Low proficiency	197.82	71.50	194.47	60.45
Average W/T				
Entire set	14.91	3.28	14.84	3.25
High proficiency	16.93	2.71	14.60	1.59
Low proficiency	14.98	3.40	15.91	4.23

Table 3
Results for accuracy for the entire set, high and low groups.

Accuracy	Time 1		Time 2	
	M	SD	M	SD
Errors/Words				
Entire set	.09	.04	.07	.04
High proficiency	.06	.03	.04	.03
Low proficiency	.10	.04	.08	.03
EFT/T				
Entire set	.35	.21	.39	.21
High proficiency	.48	.23	.58	.23
Low proficiency	.29	.19	.33	.17
EFC/C				
Entire set	.46	.20	.50	.18
High proficiency	.57	.21	.65	.17
Low proficiency	.40	.18	.45	.15

6.3. Academic vocabulary use

Table 8 shows the total and average number of occurrences of items on the AWL in students' texts at times 1 and 2 for all participants, and for the high and low groups.

These figures show that by time 2, there was an overall increase in the use of AWL tokens. The occurrence of types also increased, but given the difficulties in type/token analyses (see Duran et al., 2004) we did not perform this analysis. Use of tokens, when expressed as a percentage of the average number of words, showed a statistically significant improvement at time 2 for the entire set, (t(68) = −4.202, $p < .001$), and for the low proficiency group (t(33) = −2.390, $p < .001$). It was close to statistical significance for the high proficiency group (t(17) = −1.592, $p = .068$).

Since word choice was found to be the third-most frequent type of error in the analysis of students' accuracy (see Table 7), we checked how items from the AWL were used. We found that the AWL words were overwhelmingly used appropriately. At time 1, the 1,102 tokens were used appropriately, and at time 2, of 1261 tokens only six were used inappropriately (by 2 high and 3 low students). The following is an example taken from a high student's writing showing the inappropriate use of the word 'access':

Third, Malaysians are exposed to various *access* to English.

As with the occurrence of errors, for individuals the use of academic words varied. Some students used many tokens of a particular type; for example: 'project, code, analyse' and this seemed to relate to their topics. As Nation (2001) notes, a few items on the AWL are closely related to specific topics. Frequent use of a type could also indicate limited use of synonyms or pronouns. Interestingly, several students used different items from the AWL at time 1 and time 2, even though the topic remained constant. The range in the number of types occurring at times 1 and 2 also varied greatly among individuals, but did not seem to be closely related to English entry level.

Table 4
Two-way ANOVA results for E/W.

Source of variance	df	MS	F	sig
Between subjects				
Group	1	.042	20.48	.000*
Within subjects				
Time	1	.008	9.98	.003*
Time × Group	1	.001	.875	.354

* note: significant at $p > .05$.

Table 5
Two-way ANOVA results for EFT/T.

Source of variance	df	MS	F	sig
Between subjects				
Group	1	.1089	18.01	.000*
Within subjects				
Time	1	.108	5.55	.022*
Time × Group	1	.016	.83	.366

* note: significant at $p > .05$.

In terms of AWL development, 44 individuals increased their use of academic words, as measured by the occurrence of AWL tokens as a percentage of words written at times 1 and 2 (although they did not all reach the 9% level). In addition, while at time 1, there were 23 students already using 9% or more of academic words (including 10 from the high and 10 from the low proficiency groups), at time 2, 34 students used 9% or more AWL words (with the increase coming from the low and the unknown score groups, as the high group remained at 10).

6.4. Text structure and rhetorical quality

Holistic quality changes from time 1 to time 2 are shown in Table 9.

Most of the scripts at times 1 and 2 began with some sort of introduction, although there were qualitative differences between the introductions at times 1 and 2. At time 1, students tended to include extensive background information about their topics, and a vague reference to the main point (thesis statement). Few introductions contained a statement telling the reader how the text would evolve (an advanced organiser). At time 2, there was a tendency to cut down on the background information, and to include a clearer thesis statement (and sometimes an advanced organiser). Changes in the organisation and structure of paragraphs, links between paragraphs and conclusions also occurred at time 2, with the majority showing improvement.

Improvements from time 1 to time 2 were mainly in structure (33 students) and content (23), followed by cohesion (7). Some students improved in more than one area; some who improved were already competent, while others started at a lower rhetorical quality level and also improved. Among those whose work showed 'no change', some were already strong writers, some improved in one area but did more poorly in another area, and others remained poor writers at both times. Those whose writing was worse at time 2 (7 students) tended to by trying out different content. The problems that these seven students had were mostly with coherence (3 students), content (2 students), and structure (2 students). Examples of the researchers' summary comments are:

More specific content, poor cohesion, unfinished.

Content less well developed, structure the same.

Too much information in the introduction, different content focus, no conclusion.

Table 6
Two-way ANOVA results for EFC/C.

Source of variance	df	MS	F	sig
Between subjects				
Group	1	.820	19.685	.000*
Within subjects				
Time	1	.092	4.841	.032*
Time × Group	1	.004	.188	.666

* note: significant at $p > .05$.

Table 7
Types and number of errors by all students, times 1 and 2.

Type of error	Time 1(total errors = 1260)		Time 2 (total errors = 996)	
	Number of errors	Error type as % of total time 1 errors	Number of errors	Error type as % of total time 2 errors
Articles	218	17.3	209	20.9
Plural -s	222	17.6	173	17.3
Word choice	188	14.9	150	15.0
Verb tense/form	164	13.0	123	12.3
Word form	118	9.3	91	9.1
Subject/verb agreement	73	5.7	65	6.5
Prepositions missing or redundant	69	5.4	50	5.0
Word order	24	1.9	22	2.2
Absence of major constituent	16	1.2	20	2.0
Absence of minor constituent	19	1.5	12	1.2
Linking of ideas	27	2.1	24	2.4
Noun, pronoun agreement with verb	12	0.9	8	0.8
Possessive	11	0.8	9	0.9
Passive form (missing or incorrect)	18	1.4	4	0.4
Determiner	62	4.9	19	1.9
Collocation	5	0.3	5	0.5
Capitalisation	4	0.3	2	0.2
Punctuation	10	0.7	10	1.0

6.5. Qualitative data: questionnaires

In analysing questionnaire responses, we focused on students' perceptions of whether PAD had helped their academic writing. Responses to the question "Has taking PAD helped you with other assignments or research writing?" were tabulated on a data display matrix under 13 categories that were developed from the students' responses. An example of one student's response and how it was categorised follows:

> Yes, it improved my knowledge in academic writing by providing relevant assignments and activities. Though I usually write reports, project proposals, and modals in my previous employment, this course provided me with more skills in improving my writing and be more confident in my writing. (High proficiency, Public Health).

This response was coded in the categories:

- Improved my understanding of academic writing
- Given me more confidence about writing
- Improved my writing

Fifty-two participants responded, some with more than one comment. The top five categories of responses were:

Improved my understanding of academic writing	25
Helped me to structure ideas and texts	13
Improved my grammar (specific areas mentioned included articles, prepositions, subject-verb agreement)	11
Helped me to evaluate others' work and make my own judgments	9
Improved my vocabulary	7

Examples of students' write-in comments included:

- I have a clear view about 'what is academic writing' and know the requirements and differences between formal and informal writings. (High proficiency, M. Public Health)

- Yes. It helps me to see the structure of writing a research paper. It also helps me to criticise research papers. Both of the skills I've never exposed to before. (High proficiency, M. Applied Linguistics)
- PAD introduces the way to structure a report and assignment. (Low proficiency, M. Telecommunication Engineering)
- Helps me with the structure of academic writing and referencing, which is very important. (Low proficiency, M. Telecommunication Engineering).
- I can organise my report better. I learn how to paraphrasing the text and how to present them. (Low proficiency, M. Engineering Management)

It is notable that these write-in comments reflect items from the PAD syllabus that had been discussed in class and stressed as feedback criteria on assignments.

7. Discussion

The vital question for us is the extent to which we can interpret these results as showing connections between the development of the students' academic writing over one semester and their study of the course Presenting Academic Discourse. As mentioned before, the main focus of the PAD syllabus was on structure, accuracy, and academic vocabulary presented to students through seminar tasks, feedback, and course materials. The qualitative data derived from the write-in comments on the questionnaire indicate that students were aware of their increased understanding of what is expected in academic writing for this environment. They also mentioned such areas of writing as structure, grammar, referencing and critique of source texts—all of which were prominent in the PAD syllabus.

Our analysis of the students' writing in terms of text structure and rhetorical quality shows that most students' writing improved over the course of the semester, particularly in terms of structure. This improvement in structure may be attributable to the students' exposure to academic texts in their concurrent studies. The study by Storch (forthcoming) showed that even students not enrolled in EAP courses improve the structure of their writing after a semester of degree study. Nevertheless, the students in this study, as is evident in their most frequent write-in comments in the questionnaire, felt that it was PAD that improved their understanding of academic writing and structuring their texts.

When we turn to the more quantitative findings, we find that the analysis of language use over time showed mixed results. Measures of fluency (in terms of total number of words and words per T-unit) showed no change over time.

Table 8
Results for words on Academic Word List appearing in students' scripts.

	Time 1	Time 2
Entire set (*n* = 69)		
Total No. words	14925	13862
Average No. of words per script	216.30	200.90
No. AWL token	1102	1261
Average No. of AWL tokens	15.97	18.28
Tokens as % of words	7.38%	9.25%
High (*n* = 18)		
Total No. words	4055	3980
Average No. of words per script	225.28	221.11
No. AWL tokens	334	371
Average No. of AWL tokens	18.56	20.61
Tokens as % of words	8.05%	9.35%
Low (*n* = 34)		
Total No. words	6726	6612
Average No. of words	197.82	194.47
No. AWL token	511	602
Average No. of AWL tokens	15.03	17.71
Tokens as % of words	7.85%	9.34%

Table 9
Changes in holistic quality from time 1 to time 2.

Time 2	Improved	No change	Worse
Entire set	40	22	7
High	14	1	3
Low	22	9	3

This is perhaps not surprising given that the participants were given a relatively short time to produce the writing (30 minutes). Some scripts were shorter at time 2 due to the use of reduced clauses (see also Mayor et al., 2007), fewer repetitions and more succinct development of content. Future research needs to consider clause structures and their possible effect on the number of words produced.

Use of academic vocabulary as measured by the AWL showed statistically significant improvements over time, with the number of students reaching the 9% threshold rising from 23 to 34. This improvement may be partially attributable to the students' exposure to academic texts in their concurrent studies. This is particularly so in the case of postgraduate students who are required to read extensively in their own disciplines. However, it could also be attributable to PAD, given that in PAD there was a strong focus on academic vocabulary in the seminars, materials provided and feedback on students' writing. For example, students' attention was drawn in seminars to appropriate expressions (some of which appear on the AWL) when referring to work by other authors, when drawing comparisons, and when commenting critically on research methods, salient information and results. At the same time, we encouraged students to notice how these expressions were used in their disciplines. Such explicit teaching of vocabulary is in line with Hinkel's (2002, 2003) recommendations. Hinkel argues that mere exposure to academic texts is insufficient to increase ESL learners' lexical repertoire.

Unlike previous studies which reported no improvement in accuracy over time despite immersion in the target language (e.g. Shaw & Liu, 1998; Storch, forthcoming), this study found that participants' grammatical accuracy improved over time, for the entire cohort and for both high (IELTS 6.5 and above) and low (IELTS 6). Our results concur with those of Polio et al. (1998). What may explain the similar results is the provision of targeted explicit teaching and systematic feedback on writing. Although university students are generally required to produce lengthy assignments for assessment, language use and accuracy are often allocated only a small proportion of the overall grade, and students receive little if any feedback comments on language matters from lecturers (Storch & Tapper, 2000). Ferris (2003) argues that for language learners, feedback on writing may be the single most important element that affects their successful development as writers, a view echoed by Leki (2006) writing about the development of graduate literacy. A further strength of PAD is that the feedback is provided on writing which is relevant to the learners, unlike what happens in many EAP type writing assignments (see Leki & Carson, 1997).

We acknowledge that our results need to be interpreted with caution. One of the limitations of this study, as is the case with many studies using a pre-post test design, is that assessment of students' writing development was based on one text produced at the beginning and one at the end of the semester. Clearly a single text may not be representative of the students' writing ability. Future studies should collect a number of texts produced by students over the semester. Another limitation of this study was the measures used. Assessing texts for rhetorical quality was at times difficult when the texts were very short. At the same time, quantitative measures, such as measures of fluency which are based on word counts may fail to capture improvements achieved in shorter texts. Reduction in words may be due to more direct writing (fewer repetitions or digressions) or more linguistically complex sentences with reduced clauses. As other researchers have noted (e.g. Hyland & Tse, 2007), some items on the AWL seem to be used more frequently by some subdisciplines than others. Nevertheless, using the AWL to trace vocabulary development is useful when dealing with student writing in different disciplines and on a range of topics. Another limitation of the study was its duration. Some aspects of writing may need longer than 10–12 weeks to show improvement.

To date, there have been relatively few studies which have attempted to measure the impact of EAP instruction on students' language development. In these times when EAP provisions are under the spotlight and subjected to calls for accountability, this study shows that EAP type courses can have a positive impact on students' writing and that this impact is measurable in quantitative terms.

Appendix A. The 18 error categories

(developed from Bardovi-Harlig & Bofman, 1989)
Syntactical

1. Word order
2. Absence of major constituent, such as subject, verb, object
3. Absence of minor constituent (e.g. 'Enterprises may not be professional [enough] to master the coordination of … ')
4. Errors in linking ideas (missing, redundant, or incorrect)

Morphological: nominal

5. Plural
6. Agreement (noun or pronoun with verb)
7. Possessive

Morphological: verbal

8. Tense and verb form. Errors of tense, aspect, mood and form for the same verb were counted as one error.
   e.g. 'Then we cannot argue that this project manager is fail [has failed]'.
   But: 'It is sometimes necessary cross [to cross] geographical differences' was coded as 2 errors of verb form and word choice (possible meanings were 'to overcome' or 'to deal with').
9. Agreement of verb with subject. Agreement errors involving both subject and verb in the same phrase were counted as one error.
   e.g. 'Every details [detail] need [needs] to be considered'.
10. Passive form (missing or incorrect)
11. **Derivational (word form)**
    e.g. 'very technologic [technological] parameters'

Grammatical

12. Determiners (e.g. this, that, it, those). Missing, redundant or incorrect.
    e.g. ' when building cantilever bridges. Those [These] bridges … '
    e.g. 'Looking at its [this] background and current situation, …'
13. Articles. Errors of article and noun plurals were counted as one error.
    e.g. When the context shows that ' the problem' should be 'problems', one error was counted.
14. Prepositions (missing or redundant)

Lexical

15. Word choice. (Register errors such as 'lots of' were not included).
    e.g. 'Many countries still out of [lack] responsibility'.
    e.g. 'especially in developing countries, such as my hometown [home] — China'.
    Prepositions were coded as word choice if the choice was incorrect.
    e.g. 'The glaciers in [at] the two poles of the earth'.
16. Collocation. Erroneous expressions and phrasal verbs were counted as one error.
    e.g. the key of the [to] success
    e.g. I am interested to conduct [in conducting]…
    If meaning was so obscure that reformulation was impossible, a phrase or clause was counted as one collocation error.
    e.g. 'The definition should ''with which'' or ''follow with'' conclude the rights, the duties …' was one error.

Mechanics
(Spelling omitted)

17. Capitalisation
18. Punctuation (if meaning was affected)

A repeated error was counted each time it occurred. Errors were counted according to the minimal number of corrections required to make a phrase or clause error-free, while maintaining the apparent meaning indicated by the context. For example, when taking context into account, a minimum reformulation of the following sentence yields 5 errors.

Original text: Communication is a critical field for a successful project manager, how need to communicate his customers.

Minimal reformulation: Communication is a critical field for a successful project manager, who (1) needs (2) to know (3) how (4) to communicate with (5) his customers.

(The phrase was **not** reformulated as ''who needs to communicate with his customers'', because from the context it was clear that the student wanted to convey the importance of a project manager knowing *how* to communicate.)

Appendix B. Guide to analysis of text structure and rhetorical quality

	Comments	
	Text 1	Text 2
Structure		
Introduction:		
• is present and well developed (contains background information and an advanced organiser)		
• is present but development is limited		
• is absent		
Body paragraphs:		
• separate and well structured		
• no separate paragraphs		
Conclusion:		
• present as a separate paragraph and appropriate		
• combined with body		
• absent		
Cohesion and coherence		
• main point/topic is clear in each paragraph and clear and appropriate links are present between and within paragraphs.		
• some paragraphs lack clear cohesion/coherence.		
• entire text difficult to follow		
Content quality		
• relevant and well supported		
• main points not well developed/repetition		
• irrelevant and/or not well developed		

Appendix C. Sample of a student's scripts (low group)

Time 1

The aim of this writing is to critically examine the reason why Tonga should consider HIV/AIDS in their agenda at the policy level.

First Tonga is a very small community of just more than 100,000 people spread over some 176 island. Even though, the report cases of HIV/AIDS is very low, but it has all the necessary ingredients of developing a serious epidemic the future. The prevalence of HIV/AIDS in Tonga was reported in 2001 to be 14 cases but there are a lot of questions behind in this issue.

Second, there has been an unrest recently happen in this region. The royal family has the largest authority in the country and the political power is weak. As a result, it leads to weak in HIV/AIDS surveillance system. It was evidence from previous years that the government has done very little about it. Sex workers, exploitation of children, men sex

with men, practice of unsafe sex and so forth has been increasing to a substantial amount leaving Tonga just like Papua New Guinea (PNG).

Last but not least, look at PNG. If we compare Tonga to PNG, more or less all the pacific islanders has not exception. Having saying this, in 2003 the reported cases for HIV living with AIDS in PNG was more than 700. In September 2004, the WHO reported to be more than 10,00 people and surprisingly in December of the same year (2004) the reported case was more 11,000 people. Therefore based on this reported and elsewhere literature it is believed that Tonga should not ignore the fact that HIV/AIDS is really exist in this nation. There is only one explanation why incidence of HIV/AIDS reported cases increase in PNG, is because they ignored the fact that HIV/AIDS really exist.

To this end, in my conclusion, it is of the essence that Tonga should consider HIV/AIDS seriously and accept the facts that it si exist in this nation. To achieve this goal, the political leaders, legal profession, public health officials and all Non Government Organisations should hand in hand address in their agenda meeting that HIV/AIDS has not exception to this devastating disease pandemic.

Time 2

Tonga is a small community of just more than 100,000 population. With a population of more than 100,000 one could be believed that HIV/AIDS prevalence and incidence is low. To support my point, with an estimated 40 million persons living with HIV/AIDS around the world, the epidemic claimed the lives of 14 cases in Tonga. Although, this is in fact true Tonga should not neglected HIV/AIDS until vaccine to cure the virus is discovered. The aim of this paper is to argue that Tonga should consider HIV/AIDS as a priority.

Three reasons that Tonga should consider HIV/AIDS seriously: the HIV surveillance system is weak; Tonga has all the necessary ingredients to develop any outbreak in the future; and the government has done very little to prevent the spread of the virus because there is no Community-Based Education (CBE) centre. CBE place insights into the importance of prevention and control of HIV/AIDS in Tonga. There is no law to prevent high risk group form transmit virus to other, so if Tonga continue to ignore the existence of HIV/AIDS, the people will suffer. In my view, Tonga should consider this issue seriously. HIV/AIDS will come to stay until something is done about it.

Analysis of sample student scripts
Quantitative analysis

	Time 1	Time 2
Fluency:		
Words	358	201
T-units	20	14
Clauses	33	21
Dependent clauses	13	7
Words/T-units	17.9	14.36
Accuracy		
No. of errors	32	15
EFT/T	0.30	0.36
EFC/C	0.42	0.52
Academic vocabulary use		
Tokens on AWL	18	9
Tokens as % of words	5%	4.45%

Qualitative analysis

	Text 1	Text 2
Structure		
Introduction:		
• is present and well developed (contains background information and an advanced organiser)		√ (topic is clear)
• is present but development is limited	√	
• is absent		

(*continued*)

	Text 1	Text 2
Body paragraphs:		
• separate and well structured	Separate paragraphs but not all clearly linked	Only one, well structured parag.
• no separate paragraphs		
Conclusion:		
• present as a separate paragraph and appropriate	√	
• combined with body		√
• absent		
Cohesion and coherence		
• main point/topic is clear in each paragraph and clear and appropriate links are present between and within paragraphs.	First paragraph lacks cohesion	More explicit and relevant connections between ideas (3 reasons)
• some paragraphs lack clear cohesion/coherence.	Enumerators used as links, but are not appropriate	
• entire text difficult to follow		
Content quality		
• relevant and well supported	Some points well supported	More succinct in quality
• main points not well developed/repetition		
• irrelevant and/or not well developed		

Summary:

The writing shows improvement at time 2. The introduction and topic sentences are better at time 2. Cohesion is more explicit. Content quality similar to time 1, but more is coherent and succinct.

References

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition, 11*, 17–34.

Belcher, D. (1995). Writing critically across the curriculum. In D. Belcher, & G. Braine (Eds.), *Academic writing in a second language* (pp. 135–154). Norwood, NJ: Ablex.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum.

Cargill, M. (1996). An integrated bridging program for international postgraduate students. *Higher Education Research and Development, 15*, 177–188.

Cargill, M., Cadman, K., & McGowan, U. (2001). Postgraduate writing: using intersecting genres in a collaborative content-based program. In I. Leki (Ed.), *Academic writing programs* (pp. 85–96). Alexandria, VA: TESOL.

Casanave, C. P. (2003). Multiple uses of applied linguistics literature in a multi-disciplinary graduate EAP class. *ELT Journal, 57*, 43–50.

Casanave, L. P., & Hubbard, P. (1992). The writing assignment and writing problems of doctoral students: Faculty perception, pedagogical issues, and needed research. *English for Specific Purposes, 11*, 33–49.

Chanock, K. (1997). 'Never say I'? The writers' voice in essays in the humanities. In Z. Golebiowski (Ed.), *Procedures. First national conference on tertiary literacy: Research and practice* (pp. 54–63). Melbourne: Victoria University of Technology.

Cheng, A. (2008). Analysing genre exemplars in preparation for writing: the case of an L2 graduate student in an ESP genre-based instructional framework of academic literacy. *Applied Linguistics, 29*, 50–72.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213–238.

Dudley-Evans, T. (1995). Common-core and specific approaches to the teaching of academic writing. In D. Belcher, & G. Braine (Eds.), *Academic writing in a second language* (pp. 293–312). Norwood, NJ: Ablex.

Duran, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics, 25*, 220–242.

Elder, C., & O'Loughlin, K. (2003). *Investigating the relationship between intensive EAP training and band score gains on IELTS*. In: *IELTS Research Reports, 4*. Canberra: IELTS Australia. pp. 207–254.

Ferris, D. (2003). *Response to student writing. Implications for second language students*. Mahwah, New Jersey: Lawrence Erlbaum.

Ferris, D. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing* (pp. 81–104). Cambridge: Cambridge University Press.

Frodesen, J. (1995). Negotiating the syllabus: a learner-centred interactive approach to ESL graduate writing course design. In D. Belcher, & G. Braine (Eds.), *Academic writing in a second language* (pp. 331–350). Norwood, NJ: Ablex.

Green, A., & Weir, C. (2003). *Monitoring score gain on the IELTS Academic Writing module in EAP programmes of varying duration. Phase 2 report*. Cambridge: UCLES.

Hinkel, E. (2002). *Second language writers' text. Linguistic and rhetorical features*. Mahwah, New Jersey: Lawrence Erlbaum.

Hinkel, E. (2003). Simplicity without elegance: features of sentences in L1 and L2 academic texts. *TESOL Quarterly, 37*, 275–301.

Hirvella, A. (2004). *Connecting reading and writing instruction*. Ann Arbor: University of Michigan Press.

Hood, S. (2005). The co-patterning of attitude and field in academic writing: What gets evaluated how? *Australian Review of Applied Linguistics, Series S, 19*, 23–40.

Hood, S. (2008). Summary writing in academic contexts: Implicating meaning in processes of change. *Linguistics and Education, 19*, 351–365.

Hunt, K. (1966). Recent measures in syntactic development. *Elementary English, 43*. 732–739.

Hyland, K. (1999a). Disciplinary discourses: writer stance in research articles. In C. N. Candlin, & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 99–121). Essex, UK: Addison Wesley Longman.

Hyland, K. (1999b). Academic attribution: citation and the construction of disciplinary knowledge. *Applied Linguistics, 20*, 341–367.

Hyland, K. (2004). *Genre and second language writing*. Ann Arbor: University of Michigan Press.

Hyland, K. (2007). Genre pedagogy: language, literacy and L2 writing instruction. *Journal of Second Language Writing, 16*, 148–164.

Hyland, K., & Tse, P. (2007). Is there an ''academic vocabulary''? *TESOL Quarterly, 41*, 235–253.

IELTS Task 2 Writing band descriptors (public version). Retrieved 28 November, 2008 from: http://www.ielts.org/pdf/UOBDS_writingT2.pdf.

Leahy, F. (2004). Proposal for alteration to selection requirements for Master of Information Technology. Faculty of Engineering Report produced for SEPSC 5/07, 20 August 2007.

Leki, I. (2006). ''You cannot ignore'': L2 graduate students' response to discipline-based written feedback. In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing* (pp. 266–286). Cambridge: Cambridge University Press.

Leki, I., & Carson, J. (1997). ''Completely different worlds'': EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly, 31*, 39–69.

Mayor, B., Hewings, A., North, S., Swann, J., & Coffin, C. (2007). A linguistic analysis of Chinese and Greek L1 scripts for IELTS Academic Writing Task 2. In L. Taylor, & P. Falvey (Eds.), *IELTS collected papers: Research in speaking and writing assessment. Studies in Language Testing, 19* (pp. 250–315). Cambridge: Cambridge University Press.

McCarthy, P. M., & Jarvis, S. (2007). Vocd: A theoretical and empirical evaluation. *Language Testing, 24*, 459–488.

Melles, G., Millar, G., Morton, J., & Fegan, S. (2005). Credit-based discipline specific English for academic purposes programmes in higher education. Revitalizing the profession. *Arts and Humanities in Higher Education, 4*, 283–303.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York: Newbury House.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*, 492–518.

Polio, C., Fleck, C., & Leder, N. (1998). ''If I only had more time:'' ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing, 7*, 43–68.

Purpura, J. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.

Read, J., & Hays, B. (2003). *The impact of the IELTS test on preparation for academic study in New Zealand*. In: *IELTS Research Reports, 5*. Canberra: IELTS Australia. pp. 237–262.

Shaw, P., & Liu, E. T.-K. (1998). What develops in the development of second language writing. *Applied Linguistics, 19*, 225–254.

Silva, T., Reichelt, M., & Lax-Farr, J. (1994). Writing instruction for ESL graduate students: examining issues and raising questions. *ELT Journal, 48*, 197–204.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*, 1–14.

Storch, N. The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing*, in press

Storch, N., & Tapper, J. (1997). Student annotations and writing conferences: What NESB and NS students say about their own writing. *Journal of Second Language Writing, 6*, 245–264.

Storch, N., & Tapper, J. (2000). Discipline specific academic writing: what content teachers comment on. *Higher Education Research & Development, 19*, 337–356.

Storch, N., & Wigglesworth, G. (2007). *Students' engagement with feedback on writing. Paper presented at the Social and Cognitive Aspects of Second Language Learning and Teaching conference*. New Zealand: University of Auckland.

Studies in Australia. (2007). International students already in Australia. Retrieved 12 January 2009, from: http://www.studiesinaustralia.com.

Swales, J. M., & Feak, C. B. (2000). *English in today's research world: A writing guide*. Ann Arbor: University of Michigan Press.

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students. Essential tasks and skills* (2nd ed). Ann Arbor: University of Michigan Press.

Tribble, C. (1996). *Writing*. Oxford: Oxford University Press.

Wolf-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. Honolulu, Hawaii: University of Hawaii at Manoa.

Worthington, D., & Nation, I. S. P. (1996). Using texts to sequence the introduction of new vocabulary in an EAP course. *RELC Journal, 27*, 1–11.

Neomy Storch is a senior lecturer in English for Academic Purposes and Applied Linguistics in the School of Languages and Linguistics, the University of Melbourne. Her research interests have focused on issues related to ESL pedagogy, and in particular L2 writing instruction, feedback and the nature of peer interaction.

Joanna Tapper is an Honorary Fellow in the School of Languages and Linguistics, the University of Melbourne. Previously, she taught English for Academic Purposes, business and technical writing, and applied linguistics. Her research interests have included analysis of spoken interactions, student experiences with group-work and the teaching of academic writing.