



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Knoch, U;Rouhshad, A;Storch, N

Title:

Does the writing of undergraduate ESL students develop after one year of study in an English-medium university?

Date:

2014-01-01

Citation:

Knoch, U., Rouhshad, A. & Storch, N. (2014). Does the writing of undergraduate ESL students develop after one year of study in an English-medium university?. *Assessing Writing*, 21, pp.1-17. <https://doi.org/10.1016/j.asw.2014.01.001>.

Persistent Link:

<https://hdl.handle.net/11343/115245>

**Does the writing of undergraduate ESL students develop after one year of study
in an English-medium university?**

Abstract

English language skills are often listed by employers among key attributes needed for employment and there is a general dissatisfaction with English standards, especially writing skills, following graduation (e.g., Healy & Trounson, 2010; Rowbotham, 2011 in the Australian context). In the case of ESL students, research on whether English proficiency improves after studying at an English-medium university has to date been scarce, and has generally examined students' gains after a relatively short duration. The current study examined students' ESL writing proficiency following a year's study in an Australian university. The study used a test-retest design. A range of measures was used to assess writing, including global and discourse measures. All participants were also surveyed and a subset was interviewed. The study found that students' writing improved after a year of study but only in terms of fluency. There were no observed gains in accuracy, syntactic and lexical complexity. Global scores of writing also showed no change over time. Students stated in their questionnaires and interviews that they did not receive any feedback on their writing from their content lecturers. We discuss our findings in relation to the students' second language (L2) proficiency and the nature of their immersion experience.

Introduction

Reports on the number of international students studying in Australia show that in 2011 nearly a quarter of a million international students were enrolled in the higher education sector (International students in Australia, 2012). These figures attest to the fact that international students continue to form a substantial proportion of students in Australian universities. Most of these students are second language (L2) speakers of English. A high level of English language proficiency is considered very important in our increasingly globalised world, and is certainly an attribute sought by employers. Studying in an English medium university provides these learners with a rich immersion experience; an environment where the learners are exposed to plentiful L2 input and where they are asked to produce meaningful L2 output. Such an environment has been hypothesized by major theories of SLA to be conducive to L2 development (DeKeyser, 2007; Gass, 2003; Swain 1985, 1995).

Research on the effect of the experience of studying in an L2 medium university on the learners' L2 development is surprisingly scant. The bulk of research on the development of academic writing has focused largely on investigating the effect of a specific course of instruction (e.g., Shaw & Liu, 1998; Authors, 2009a), or following relatively short time periods of time at university such as one semester (e.g., Authors, 2008; Authors, 2007, 2009b). The present study set out to investigate the impact of studying in an L2-medium university on second language writing development of non-English-speaking background students after one year. The data forms part of a larger study examining the writing development of students during their three years of undergraduate study.

Literature review

Over the past decade there have been a growing number of studies on second language writing development. The majority of these studies involved a specific course of instruction. However, these courses varied in the type and length of instruction offered (e.g., a short period of study abroad or specifically-designed English for Academic Purposes semester length courses). Only a few studies focus on learners who have not taken a specific course of instruction. The studies also differ in how they measured writing development. Some employed band scores based on ratings by trained raters using band descriptors, others employed more fine-grained discourse-analytic measures (e.g., measures of fluency such as the number of words produced) and some made use of both types of measures. In the following we discuss these major strands of research in more detail as they illustrate the methodologies used as well as the major findings.

Writing development as a result of instruction

Studies that have examined second language writing development at university (e.g., Sasaki 2007, 2009, 2011; Shaw & Liu, 1998; Authors, 2009a) used a range of measures to gauge development, and the results seem to vary depending on the measures used. In a series of study-abroad studies, Sasaki (2007, 2009, 2011) compared the L2 writing development of study abroad and at-home groups. Using composition band scores, Sasaki showed that 1.5 to 11 months of study abroad (when

combined with general ESL and writing instruction) resulted in significant increases in the learners' L2 writing scores, whereas the writing of the at-home group counterparts did not show improvement over time.

Studies outside the study-abroad context, which employed band scores (e.g., Tsang & Wong, 2000; Xudong et al., 2010), and discourse measures (e.g., Shaw & Liu, 1998; Authors, 2009a; Tsang & Wong, 2000) to gauge development in writing, have shown mixed results following 10 to 14 weeks of instruction. For example, in Authors' (2009a) study, only the discourse measures of accuracy (measured via Errors/words, proportion of error free clauses and error free T-units) and lexical complexity (measured as a percentage of words which appeared on Coxhead's (2000) Academic Word List) showed improvement. However, measures of fluency and syntactic complexity showed no change following instruction. In Shaw and Liu's (1998) study, accuracy (measured via Errors/T-units), and complexity (measured via increased use of nominalization, reduction of clauses to prepositional or participial phrases, increased subordination or changes in the pattern of subordination) fell short of significance. However, the writing of undergraduate TESL students in Tsang and Wong's (2000) study showed a statistically significant increase in most of the discourse measures following 14 weeks of instruction. The increase was found in word count (a measure of fluency), average T-unit length and the number of T-units per sentence (measures of complexity), and the ratio of error-free T-units (a measure of accuracy). However, the ratio of error-free sentences and clauses, and the average clause length fell short of significance and the participants' essay scores did not

improve significantly. This study was conducted in Hong Kong and it is not clear what writing level the students had at the start of the course. The authors did mention that the opportunities for using English outside the classroom for the 11 participants were very limited.

Other discourse features of academic writing beyond accuracy, fluency and complexity have also been shown to improve following a course of instruction. For example, structural adequacy (i.e., introduction, conclusion and paragraphing) (Felix & Lawson, 1994), text structure and rhetorical quality (Authors, 2009a) improved following four to ten weeks of instruction. Similarly, use of informal expressions (e.g., contractions and colloquial language) significantly decreased (Shaw & Liu, 1998; Authors, 2009b). It was also shown, however, that the most frequent types of errors persist following 10 weeks of instruction (Authors, 2009a), and grammatical errors remain at the same high rate following four weeks of instruction (Felix & Lawson, 1994).

In conclusion, research on the impact of instruction on writing development of university L2 learners has not yielded conclusive findings. Part of the reason for this is that the studies differed in the length of the instruction, the instructional context and the measures used to gauge writing development. The periods of instruction in the studies reviewed were generally relatively short (no more than 14 weeks).

Writing development while studying in an L2-medium university (without a specific course of instruction)

Available research suggests that 12 weeks of studying in an L2-medium university results in a statistically significant increase in writing band scores (Authors, 2009b; Authors, 2008). However, discourse measures (i.e., fluency, accuracy, grammatical/syntactic complexity, and lexical complexity) have shown to be resistant to any significant change during the same period (Authors, 2007, 2009b; Authors, 2008).

Examining L2 writing development of undergraduate (after 3 years) and postgraduate students (after 1 to 1.5 years), O'Loughlin and Arkoudis (2009), using IELTS writing scores, showed that the average group score for writing increased by 0.206. The researchers reported that writing showed the least improvement compared to the other language skills investigated. Consistent with the results of other studies (e.g., Elder & O'Loughlin, 2003; Green, 2005), O'Loughlin and Arkoudis (2009) showed that students with lower scores showed the greatest improvement.

Research on factors that influence writing development

A number of researchers have also investigated factors which can explain improvement in L2 writing or lack thereof. Factors that have been reported in previous research to correlate to gains in writing proficiency include enrolment status (undergraduate versus graduate enrolment), accessing available language support, and integration into the host culture (see Elder & O'Loughlin, 2003; Green & Weir, 2003;

Authors, 2008). For example, Authors (2008), correlated responses on a questionnaire with students' discourse scores, and found that accuracy in writing was significantly influenced by participants' integration (operationalized as having Australian friends). Their study further showed that participants were fairly accurate in gauging their writing development: those who claimed not to have difficulties with their English scored significantly higher on all the analytic scores than those who admitted to having difficulties. Based on interviews conducted with a small number of students, the researchers suggested that one reason for the lack of improvement in writing was that many of the learners had little opportunity to produce extended writing or to receive feedback on their writing during the course of their study at the university. The researchers conclude that future research needs to take into consideration the nature of students' writing assessment tasks.

The study

One shortcoming of available research is that writing development was assessed over a relatively short period (i.e., 12 weeks). Perhaps, accuracy, fluency and complexity take a longer period to develop. Ortega (2003), for example, claims that complexity (measured via mean length of T-unit) improves substantially after 12 months of language instruction. Thus the present study used a longitudinal design (one year) and set out to investigate writing development using both rater-based scores and discourse measures. We also used a questionnaire to explore the possible impact of practice in L2 writing and other L2 skills on writing development. Finally, a subset of the

students was also interviewed. The participants in this study were students who were studying for their degree courses in an English-medium university.

The present study sets out to investigate the following research questions:

- (1) What aspects of L2 writing, if any, develop following one year of degree study in an L2-medium university?
- (2) Does proficiency level at entry influence writing development?
- (3) What variables (i.e., personal, social, and environmental) predict development in L2 writing?

Methodology

Participants completed two Diagnostic English Language Assessment (described in more detail below) essays one year apart (Time 1 and Time 2). They also completed a questionnaire at Time 2 and a subset of participants took part in semi-structured interviews. The methodology is described in more detail below.

Participants

Data was collected from 101 non-English-speaking-background students in a test-re-test design study. All participants had satisfied the [name of University removed] English language entrance requirement in one of several different ways (e.g., having an IELTS overall score of 6.5, completing secondary education in an Australian school, etc).

The students came from a range of countries. The largest sub-group were students born in China (n = 46). Other frequently represented countries of origin were Malaysia (n = 16), Singapore (n = 6), Indonesia (n = 4), South Korea (n = 3), Vietnam

(n = 4), and Hong Kong (n = 3). The students were enrolled in a range of degree courses across the university, with the largest cohort (n = 43) enrolled in Business and Economics. Nineteen students were enrolled in a degree in the Faculty of Medicine, Dentistry and Health Sciences, 14 in Science, nine in Engineering degrees, and seven in Arts. Slightly more than half of the participants were female (n = 57). Just over a half of the participants (n = 53) had received their secondary education in Australia. Furthermore, it should be noted that of the 101 participants only 11 took English for Academic Purposes (EAP) as a credit-bearing subject in their first year at the University. The purpose of this course is to develop the academic English required for university study and one of its primary foci is on developing students' academic writing in English. It is important to note that the Australian context differs markedly from the US, where students from all background are required to take first year composition classes. This is not the case in Australia, where EAP and communication type courses are not compulsory. They are usually only taken by a small percentage of students (often international students) who voluntarily enrol or are required to enrol, for example, based on a post-entry language assessment.

Instruments:

The Diagnostic English Language Assessment (DELA)

All students at the [name of University removed] are required to meet the language proficiency requirement at entry. The minimum threshold proficiency level for international students is within the range of overall IELTS scores 6.5 to 7 (or equivalent) for most degree courses, depending on disciplines and programs (e.g., a score of 6.5 for most undergraduate degrees but a 7 for a Master of Education).

It is, however, believed that students who are at this minimum proficiency level can still benefit from English language support programs. Moreover, given that IELTS scores may be up to two years old at students' entry, IELTS scores may not provide an accurate account of students' proficiency level, and thus their possible language needs. Therefore, the University encourages all in-coming non-native-speaking

students to take the Diagnostic English Language Assessment (DELA) at entry. While the assessment is not mandatory for most students, students entering the University with English-language scores below certain levels (e.g., below IELTS 7) are required to take the DELA.

DELA is a professionally validated test designed by the [name of Centre, and name of university removed]. DELA consists of Listening, Reading, and Writing subtests. Each subtest is assessed on a 6-point scale, with 6 representing advanced proficiency level. The writing subtest requires students to write an argumentative essay of approximately 300 words on a given topic. The exact topic used in this study cannot be disclosed as it is still an operational task. Brief statements for and against (one or two sentences in length) are provided to help students with possible ideas for their writing. Students are given 30 minutes to complete the task. The writing subtest is assessed based on the three criteria of Organisation, Content, and Form and essays receive a band score (based on a 6-band scale) for each of the criteria. Then, the band scores are averaged to compute the overall writing score.

Based on students' performance in DELA, the University provides recommendations to students regarding the degree of language support (i.e., credit-bearing ESL or workshops) they may need. Participants who score on average 3.3 or below are categorized as *at-risk* and thus are required to take a credit-bearing ESL subject. Those who score from 3.4 to 4 are considered *borderline* in the sense that they may still need support and are recommended to attend university workshops to enhance their proficiency. Those participants whose DELA score exceeds 4 are coded as *proficient* learners. However, it is important to note at this point that an average score of, for example, 4.5 on DELA only equates approximately to an IELTS score of 7 and therefore learners being placed in this 'proficient' group are certainly still able to improve and develop their language ability further. Students in this highest group are also advised that they could benefit from further courses and directed self-study.

Questionnaire

After one year of degree study, participants were asked to complete a questionnaire (see Appendix), which included questions about participants' practice in L2 writing in social and educational contexts, and practice in L2 speaking, listening, and reading. For example, to investigate their writing in the educational context, students were asked about the nature of their written assignments during their study and lecturers' feedback on their written assignments. Variables in the questionnaire were based on our review of the literature and on a questionnaire used by Authors (2008).

Interview questions

As part of a larger study, a subset of students was also invited to participate in interviews. The interview questions enquired about participants' opinions about their writing development, and availability of lecturers' feedback comments on their writing. Moreover, participants were asked whether they had assumed that their writing would improve prior to commencing their degree.

Data collection procedures

In early 2009 (Time 1), a large number of non-English-speaking-background students took the DELA as part of the normal administration. For the writing subtest, they were asked to write on the topic of corporate advertising in schools. As part of this test administration (conducted during the orientation week of the first semester), test takers also completed a background questionnaire.

One year later (Time 2), all undergraduates who had taken the DELA in 2009 were invited to take part in this study. One hundred and one students agreed to participate and formed the participants in the current study. At Time 2, participants completed the writing component of the DELA. They were given the same essay prompt as the length of time between the testing session and the low-stakes nature of the assessment made it very unlikely that the students would remember the essay topic in detail from

the first administration. We decided to use the same essay topic to control for topic effect. After completing their essay, participants were asked to complete the questionnaire.

As part of a larger study investigating the writing development of students after three years of undergraduate study, we also interviewed a subset of students. We will only report on questions relevant to the first year of study in this paper. The semi-structured interviews were conducted in a private room by two of the researchers.

Data Coding and analysis

To answer research question 1, we compared the writing scripts produced at Time 1 and Time 2 using both the DELA writing scores and a series of discourse-analytic measures. The measures we employed were a mixture of automatically coded features and measures that required hand-coding. The measures were grouped into the following categories: accuracy, fluency, grammatical and lexical complexity. Table 1 below lists the measures used in this study.

[insert Table 1 around here]

Accuracy was measured via the percentage of error-free clauses, and the percentage of error-free T-units. The coding for T-unit and clause boundaries were based on the coding scheme elaborated by Cummings et al. (2006). The coding of error-free clauses and error-free T-units followed the coding scheme used by Authors (2012). Intercoder reliability checks were established for a subset of the data (20%) using both a Spearman rho coefficient to compare the number of units in each essay and a percentage agreement index to compare exact agreement of the location of units. The intercoder reliability indices for T-units, clauses, and error-free T-units and clauses can be found in Table 2 below.

[insert Table 2 around here]

Fluency was measured by counting the number of words in each essay. This was calculated automatically using a Perl script program designed for that purpose by a qualified programmer. In this program contractions and numbers are counted as one word.

Grammatical complexity was measured via the number of words per T-units (W/T-unit), the number of words per clauses (W/Clauses), and the number of clauses per T-unit (C/T-unit). We note that the number of words per T-unit is a contentious measure of grammatical complexity. Not all scholars agree that longer T-units necessarily indicate more complex writing. For example, a T-unit could be shorter because the writer uses reduced clauses, a more complex type of clause structure. However, the mean length of T-units is a frequently used measure of complexity (see review of studies in Ortega, 2003) and thus it enabled us to compare our findings with those of other studies.

Written scripts were also coded for different measures of lexical complexity. These measures included percentage of words from the Academic Word List (AWL), lexical sophistication, and D-value. The Academic Word List (AWL), developed by Coxhead (2000), consists of 570 word families. AWL is comprised of 10% of words in a variety of academic texts. Computation of the percentage of AWL was carried out using the program Vocab Profile (Cobb, 2002). The same program was also employed for measuring lexical sophistication. Lexical sophistication was measured by dividing the number of sophisticated words by the number of content words. Sophisticated words were defined as words that either appear on the Academic Word List or are Off-List words (i.e., words that do not appear on any of the lists but are generally less common, sophisticated words). The words appearing on the off-list were manually scanned to ensure that they were in fact sophisticated words rather than informal, misspelled or non-words. D-value is a measure of lexical richness which is derived by computing a set of type/token ratios for each text (Malvern, Richards, Chipere &

Duran, 2004). As a single measure of type/token ratio is text-length dependent, Malvern et al. proposed computing a set of type/token ratios for each text and called it d-value. D-value calculation was done automatically using the program d-tools from the website: www.lognostics.co.uk.

All writing scripts were also rated by an experienced DELA rater with 30 years of rating experience and a Masters degree in Applied Linguistics. The rating was based on the 6-band DELA rating scale. To investigate the inter-rater reliability for marking the scripts, a subset of 20 writing samples (20%) were double-rated. The resulting Cronbach's alpha was 0.853.

We then conducted a statistical analysis using paired sample t-tests to answer research question 1. Because we employed a number of measures in our study, we adjusted the alpha value using a Bonferroni correction. The alpha value was therefore lowered to .004. Partial eta squared effect sizes are reported for variables identified as significant. We also conducted a further sub-analysis to compare the writing development of students who completed their secondary education in Australia with those who arrived in Australia to start their degree programs.

To answer research question 2, we grouped the participants into three groups based on their average writing score at Time 1: at risk, borderline, proficient. We chose these three groupings as they are used to group students into support options following the DELA. As mentioned above, it is important to remember that the label 'proficient' (approximately equivalent to an IELTS level of 7) includes students who may still gain from improving their English proficiency.

Following the grouping of the students into the three proficiency levels, we compared their change from Time 1 to Time 2 on each of the discourse measures and the band scores separately. This was of interest, to see if any of these groups differed in their change after one year. To arrive at the results, we conducted a multi-step statistical

analysis. We first conducted a mixed method ANOVA analysis with time (with two levels) as the within-person factor and group (with three levels) as the between-person factor. While this analysis was interesting, it only provided us with overall alpha values for proficiency group, time, and the interaction effect. As a follow-up to results found to be significant in the ANOVAs, comparisons between means were made between relevant pairs of means based on the least significant difference (LSD). For space reasons, we will only present a summary of the results of this analysis.

To answer research question 3, we entered all responses into SPSS. The questionnaire questions were grouped into themes, depending on the focus of the question. We grouped the questions into the following thematic topics:

- practice in L2 writing in social and educational contexts
- practice in speaking, listening, and reading

We then grouped our learners into categories based on the discourse analysis:

1. students who improved in accuracy from Time 1 to Time 2 and students who did not improve
2. students whose fluency improved from Time 1 to Time 2 and students who did not improve in fluency

We chose to focus on these two distinctions for the following reasons. Firstly, students could be easily grouped based on accuracy and fluency. We were not able to categorize students easily on the complexity measures as each of them conceptualizes complexity differently and therefore students who improved on one measure might not have improved on another. Secondly, we noticed from our analysis that while we expected an increase in accuracy for the larger cohort, the largest increase was found in fluency. We therefore decided to include both aspects into our analysis for research question 3.

To arrive at our results for research question 3, we conducted a series of chi square analyses on each of our relevant questionnaire questions and the grouping of students described above. In our results section, we will only comment on the most salient, interesting results from this analysis.

As mentioned earlier, three interview questions elicited opinions of a subset of participants on availability of lecturers' comments on their writing, their assumption about writing improvement prior to starting their degree, and their overall impression of how their writing changed during their study. We also drew on our interviews, where relevant, to answer question 3.

Results

To answer the first research question, DELA scores and discourse measure scores are compared between Time 1 and Time 2. Descriptive and inferential statistics are presented for all participants, without regard to their proficiency.

Analytic essay scores

Table 3 presents the descriptive statistics and Paired Sample t-test results of DELA rating scores for all participants. As the table shows, for analytic and overall scores there were minute differences in mean scores between Time 1 and Time 2 and thus the differences fell short of significance.

[insert Table 3 about here]

Discourse measures

Table 4 presents the descriptive and inferential statistics for the discourse measures. As the table shows, only fluency improved significantly from Time 1 to Time 2 ($p < .004$) with a large effect size. All other discourse measures, however, fell short of significance. That is, they did not change from Time 1 to Time 2 (a year later).

[insert Table 4 about here]

We also conducted a further analysis, comparing the writing development for students who completed their secondary education in Australia and those who completed school overseas. The findings for those two sub-groups mirrored those of the whole group in that the only significant result was found in an increase in fluency.

To answer the second research question we examined whether proficiency level of the participants, as an intervening variable, would influence the change(s) in their discourse measures over time. To do this, we divided participants into three different proficiency groups (at risk, borderline and proficient) based on their average writing scores at Time 1. As noted earlier, our decision on cut-off points was motivated by the type of language support generated by the DELA scores. In the current data, there were 14 at-risk students, 36 borderline students, and 51 proficient students. Table 5 to Table 8 show the means of the discourse measures at Time 1 and Time 2 and the differences between them for the three proficiency levels.

[insert Tables 5-8 about here]

As the tables show, the differences in the mean scores of discourse measures were minute for all measures for the three proficiency levels. The mixed method ANOVAs with follow-up comparisons using the LSD did not result in any significant differences between the three proficiency groups.

Thus far, the results of the analysis have revealed that participants' DELA scores and discourse measures, except for fluency, have not changed significantly from Time 1 to Time 2. Further, proficiency level was not an intervening variable, as it did not influence the improvement/lack of improvement in discourse measures over one year of degree study.

Questionnaire analysis

The third research question asked whether participants' L2 writing practice and practice in other L2 skills (speaking, listening and reading) can predict changes in students' writing accuracy and fluency over one year. As mentioned in the methodology section, we decided to only focus on accuracy and fluency in this section, as the results for complexity were more varied and less straightforward to interpret. To answer this research question, participants were coded as either "improved" or "not improved" based on their changes in percentages of EFC (accuracy) and word count from Time 1 to Time 2 (fluency).

We will report questionnaire results in two main areas: (1) L2 writing practice, and (2) practice in the other three language skills. The questionnaire elicited data about L2 writing practice in both social contexts (i.e., email, text messaging, and chat), and educational contexts. Questions relating to L2 writing practice in educational contexts asked whether in their first year lecturers commented on their English ability, whether they received English writing support, and the nature of assessment tasks (e.g., group vs. individual).

In their answers, 77 participants (76%) mentioned that they did not receive comments on their written assessments from their lecturers. Further, 64 participants (63%) mentioned they did not receive L2 writing support in their first year. Most of those who received support, sought it from the Academic Skill Unit (which offers free workshops on language skills and provides free language tutorials for students) or their friends. Moreover, 84 participants (83%) mentioned that assignments in their first year were mostly individual. The results of the Chi-square showed that there was no relationship between L2 practice writing and improvement or lack of improvement in writing accuracy and fluency.

The questionnaire also elicited data about participants' practice in L2 speaking, listening and reading. These questions asked about participants' medium of

communication at home, and frequency of conversations in English. Sixty one participants (60%) mentioned they always used a language other than English for communication at home. However, seventy two participants (71%) mentioned that they had conversations in English very often (more than once a day) or often (at least daily) and 53 (52%) noted that they communicated in English with most/half of their friends. Furthermore, 56 participants (55%) noted that they read books for pleasure always /often in English and 72 participants (71%) mentioned that they always or often watched TV in English. The results of the Chi-square showed that there was no relationship between improvement in accuracy and practice in these three language skills. As far as fluency is concerned, one significant relationship was found for daily communication in English. The results showed that participants who mainly communicated in English for daily communication were more likely to write longer essays at Time 2.

Few of the questions in the questionnaire were open-ended. One of the questions asked the participants to write about their difficulty in writing. The 37 students (36%) who responded to the open-ended question mentioned a wide range of difficulties from grammar and lexical difficulties to writing under time pressure.

As mentioned above, the majority of the participants mentioned that they did not receive feedback from lecturers on their assignments. In the open-ended section of the questionnaire, twenty participants (19%) noted that they received only general feedback from their lecturers. That is, they were told that they had problems with English grammar, but were not given any more specific pointers on what to improve. This was also supported by our analysis of the interviews. Only two interviewees mentioned they received feedback on their language ability. The rest of the participants either did not receive feedback or mentioned that they received generic comments such as “*you need to improve your writing*”. Further evidence of this comes from an extract from the interview with one participant:

I think feedback ... every write assignment we only get feedback on how the assignments is good or bad but the ge- but the lecturer don't te- mm tell you ... this word is not right or this sentence is not right or your mmm or you like ... structure is not right. They don't talk about that. They only talk about you should write this, you should write that, you don't write this, you don't write that."... no matter how you write, you get marks.

When asked whether, before starting their degree, they assumed their writing would improve after studying in an L2 medium university, all but one interviewee claimed that they had assumed that their L2 writing would improve. Our findings show no evidence that this in fact occurred. Interestingly, nearly half of the participants (43%) felt that their English writing improved during the first year. Participants were also asked to explain the reason for their perceived improvement. Fourteen of the participants (13%) attributed their improvement to exposure and practice in using English. Six of the participants (5%) also attributed it to learning about the requirements of the academic essays and the assessment criteria.

Discussion

The current study investigated the second language writing development of non-English-speaking students after one year of degree study in a second language environment. The results of the study showed that analytic and overall DELA writing scores did not change significantly after one year of degree study. This is inconsistent with the findings of Authors (2008), and Authors (2009b) which documented significant improvement in both analytic and overall DELA scores after 12 weeks of degree study in the same institution. This inconsistency could perhaps be explained in terms of the participants' concurrent EAP program and proficiency level. In Authors (2008) one-third of the participants (13 out of 39) were enrolled in a credit-bearing EAP in their first semester of study. However, in the current study only 11 (10.89%)

of the participants were enrolled in a credit-bearing EAP in their first semester. Participation in EAP programs, the instruction, practice in academic writing, and feedback students receive on their writing in such programs, may explain why students in Authors' (2008) study showed improvement. Another reason that may explain why the findings of this study were inconsistent with those of previous research is L2 proficiency. In Authors' (2009b) study, the students were assessed on DELA as needing additional support. This suggests they had relatively low proficiency level. In the current study 60 (59.40%) of participants' overall DELA score at Time 1 exceeded 4 (out of 6); that is, they had a higher proficiency level (compared to participants in Authors' [2007] study). Previous research (Elder & O'Loughlin, 2003; Green, 2005; O'Loughlin and Arkoudis, 2009) suggests that it is students with a lower proficiency level that are more likely to show improvement over time.

The only significant improvement participants in the current study showed was in their *fluency* (measured via text length). That is, they could write longer texts in the time allowed. This is inconsistent with the results of previous studies (e.g., Authors, 2007; Authors, 2008; Authors, 2009b) in which fluency (measured via text length) did not improve after immersion in an L2 medium environment. Perhaps, this could be explained by the longitudinal nature of the current study. Unlike the previous studies that measured change after 12 weeks, the current study investigated changes after one year. Therefore, this study suggests that fluency, like complexity (Ortega, 2003), may require more than one semester to develop. This is clearly an area that needs to be further investigated.

One factor that could explain the improved fluency is extensive exposure to the target language. In the questionnaires, most of the participants noted that they had opportunities to practice their L2 in both academic and non-academic contexts. For example, 86 (85%) of the participants noted in the questionnaire that they had a fair amount/a lot of writing during their first year at university. Further, 84 (83%) of

participants mentioned that assignments were mostly individual (as opposed to group assignments).

Extensive exposure to the target language could also explain improvement in *lexical complexity* as measured by d-value, although this improvement was only approaching statistical significance. D-value is a measure of type/token ratio which takes into consideration text length. In other words, after one year of degree study participants produced a greater range of vocabulary. This improvement in the range of vocabulary could be due to the learners' exposure to second language and living in an input-rich environment. However, lexical sophistication, and percentage of Academic Word List did not change after one year. This may be attributable to the nature of the task used in the test. Perhaps our participants' lexical sophistication and percentage of AWL could have shown improvement had they been asked to write about a discipline-related topic (see Authors, 2009a).

We found, however, that greater exposure to the target language did not result in improved grammatical complexity or accuracy. Findings regarding lack of improvement in grammatical complexity are consistent with the findings of previous studies (e.g., Authors, 2007, 2009b). Studies which have shown evidence of improvement in complexity tend to be with students who are undertaking concurrent language classes (e.g., Authors, 2011; Tsang & Wong, 2000). Thus our findings confirm that grammatical complexity is unlikely to improve in the absence of language instruction. Ortega (2003) speculated that it takes one year of language instruction for grammatical complexity to improve. Our study suggests that grammatical complexity may in fact take longer than one year to show improvement. We acknowledge that our claim regarding lack of grammatical complexity is limited to the three measures (W/T, C/T and W/C) employed in the current study, and a different picture may emerge, if one employs a different measure of grammatical complexity (e.g., see Lu, 2010).

A number of reasons may explain the lack of improvement in grammatical accuracy found in this study. It is possible that participants produced more fluent essays at the expense of accuracy. Another reason to explain this finding is related to pedagogy. Many of our participants commented on the dearth of corrective feedback (CF) on their written assignments in the questionnaire and the interviews. Students also mentioned that they felt there was no need for further work on their English if no subject lecturers commented on their expression. A large number of studies on written CF (e.g., Authors, 2009c; Bitchener, Young & Cameron, 2005) show the benefit of CF for improving accuracy. Generic feedback on writing, which does not draw learners' attention to language use, may be a disservice to the students. It may provide them with the impression that their writing does not need any further improvement.

This finding has major implications for the assessment practices of university lecturers who are usually not trained in identifying issues specific to ESL students but who anecdotally complain about the quality of writing submitted as part of assignments or exams. This study shows that students appreciate feedback on their writing and if this is lacking, are then not aware of any deficiencies in the quality of their texts. Therefore a closer focus on the assessment practices of university content lecturers and professional development in this area could prove beneficial in raising awareness amongst staff and students, and ultimately improving writing standards across university courses. This would also result in the focus of writing support not only being situated in language support units, but being addressed more closely at all levels of study. However, it is not clear whether content lecturers can be trained to provide corrective feedback or whether they would be prepared to add further complexities to their marking and general workload. One solution could be to ask lecturers to 'screen' content essays and refer any students in need of language support to the appropriate support options available on campus. This study has shown that students perceive the lack of corrective feedback as a sign that they do not need to improve. This situation clearly needs to be addressed.

Like most research, the study has several limitations. It is possible that the task used in this study, a generic argumentative essay task of the type used in IELTS and TOEFL assessments, was not able to capture improvement in the writing of students. Tasks which are more discipline-specific, such as writing scientific reports and on topics related to the students' study discipline, may have captured such improvements, particularly improvement in lexical sophistication. Similarly, it is possible that the measures used in this study were too blunt to detect any developments. For example, improvement in content was not examined while it is possible that students improved in their ability to critically evaluate a topic. Another important factor is students' task motivation. It is possible that students completed the writing task at Time 2 as part of a research project and therefore may have felt less pressure to perform to the best of their ability. At Time 1, writing was used to diagnose L2 writing ability and requisite support options, and may have therefore been perceived as a higher stakes performance. Students' perceptions of writing tasks and conditions merit further consideration in future research on writing development.

Conclusion

The study shows that L2 writing does not develop substantially after one year of degree study in an L2-medium university. While an improvement in fluency was recorded, the same was not the case for measures of accuracy or complexity. Students commented on the lack of feedback on their writing by their university lecturers who tended to focus only on the content of their written assignments, and this may explain the lack of improvement in the students' writing. The study has clear implications for pedagogy. Institutions could focus on 'training' content lecturers to provide some corrective feedback or to identify students needing support and guide these to the appropriate language development courses on campus.

While this study has focussed on writing development in the first year of degree study, future research should focus on writing development beyond the first few semesters. It is feasible that the writing demands on students increase as they progress

through their degrees with possible improvement later in the student lifecycle. Given the importance of L2 proficiency, including L2 writing, for employment and for study, further research is clearly necessary in this area.

References

[all references by authors and colleagues have been excluded for blind review]

Authors. (2007).

Authors. (2008).

Authors. (2009a).

Authors. (2009b).

Authors. (2009c).

Authors. (2011).

Authors. (2012).

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191-205.

Cobb, T. (2002). Web Vocabprofile retrieved from <http://www.lex tutor.ca/vp/>

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

Dekeyser, R. M. (2007). Introduction: Situating the concept of practice. In R. M. Dekeyser, (Ed.), *Practice in a second language* (pp. 208-226). New York: Cambridge University Press.

Elder, C., & O' Loughlin, K. (2003). Investigating the relationship between intensive EAP training and band score gains on IELTS. *IELTS research reports*, 4, 207-254.

Felix, U., & Lawson, M. (1994). Evaluation of an integrated bridging course on academic writing for overseas postgraduate students. *Higher Education Research & Development*, 13(1), 59-69.

Gass, S. (2003). *Input and interaction*. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (224-255). Oxford: Blackwell.

Green, A. (2005). EAP study recommendations and score gains on the IELTS Academic Writing test. *Assessing Writing*, 10, 44-60.

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

Healy, G., & Trounson, A. (2010, 10 February). Universities told to boost English programs. *The Australian*. Retrieved from <http://www.theaustralian.com.au/higher-education/universities-told-to-boost-english-programs/story-e6frgcjx-1225828473695>

International students in Australia (2012, August). Retrieved from <http://www.studiesinaustralia.com/studying-in-australia/why-study-in-australia/international-students-in-australia>.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496. doi: 10.1075/ijcl.15.4.02lu.

Malvern, D., Richards, B., Chipere, N., & Duran, P. (2004). *Lexical diversity and language development: quantification and assessment*. Basingstoke: Palgrave Macmillan.

O'Loughlin, K., & Arkoudis, S. (2009). Investigating IELTS exit score gains in higher education. *IELTS Research Reports*, 10, 95-180.

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.

Rowbotham, J. (2011, 12 January). Employer call for English fluency. *The Australian*. Retrieved from <http://www.theaustralian.com.au/higher-education/employer-call-for-english-fluency/story-e6frgcjx-1225985875346>

Sasaki, M. (2007). Effects of study-abroad experiences on EFL writers: A multiple data analysis. *The Modern Language Journal*, 91, 602-620.

Sasaki, M. (2009). Changes in EFL students' writing over 3.5 years: A socio-cognitive account. In R. M. Manchon (Ed.), *Learning, teaching and researching writing in foreign language contexts* (pp. 49-76). Clevedon, England: Multilingual Matters.

Sasaki, M. (2011). Effects of varying lengths of study-abroad experiences on Japanese EFL students' L2 writing ability and motivation: A longitudinal study. *TESOL Quarterly* 45(1), 81-105. doi: 10.5054/tq.2011.240861.

Shaw, P., & Liu, E. (1998). What develops in the development of second language writing?. *Applied Linguistics*, 19(2), 225-254.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In M. S. Gass & C. G. Madden.

U.S.A, Newbury House Publishers. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 235-256). Rowley, MA: Newbury House.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *principle & practice in Applied linguistics* (pp. 125-144). Oxford: Oxford University Press.

Tsang, W. & Wong, M. (2000). Giving grammar the place it deserves in process writing. *Prospect*, 15(1), 34-45.

Xudong, D., Cheng, L., Varaprasad, C., & Leng, L. (2010). Academic writing development of ESL/EFL graduate students in NUS. *Reflections on English language teaching*, 9(2), 119-138.

Appendix: Questionnaire

1. Did you study at an English-speaking university before you came to the university?
 - a. Yes
 - b. No

2. In your first year at the university, where did you live? Circle more than one answer if you moved during the first year.
 - a. in a hostel/college
 - b. with a host family
 - c. with a relative
 - d. with friends
 - e. in student accommodation
 - f. alone

3. What language did you usually speak at home during your first year at the university
 - a. English
 - b. Other _____

4. In your first year at the university, how often did you have a friendly conversation in English?
 - a. Very often – more than once each day
 - b. Often – at least daily
 - c. Sometimes – at least weekly
 - d. Rarely
 - e. Never

5. In your first year at the university, did you have friends with whom you only communicated in English?
- a. Yes – most of my friends
 - b. Yes – about half my friends
 - c. Yes – only some of my friends
 - d. No

6. Did you experience any difficulties with writing in English during your first year of study?
- a. Yes,
 - b. No

If yes, what kind of difficulty did you experience?

7. If you answered ‘Yes’ to the question above, what actions did you take to address these difficulties (circle as many as apply)
- a. no action
 - b. I sought help from my lecturer/tutor
 - c. I sought help from a fellow student
 - d. I got a native speaker of English to edit my work
 - e. I asked for an extension of time for my assignment/s
 - f. Other _____

8. Did any of your lecturers comment on your English language ability when giving feedback on your written work?
- a. Yes
 - b. No

If yes, what was the nature of this feedback?

9. Do you believe that your skills in English may have affected the results you received for some assignments?
- a. Yes, certainly
 - b. Yes, possibly

c. No, not at all

10. In your first year at the university, did you receive any English writing support?

- a. Yes
- b. No

11. If you answered 'yes' above, what type of support?

- a. Academic Skills Unit
- b. ESL (English as a second language) credit subjects (e.g., Academic English 1)
- c. Friends
- d. Someone in my host family
- e. Other (please specify)

12. Do you feel your academic English writing improved during your first year at university?

- a. Yes
- b. No
- c. Not sure

If yes, why do you think you improved? If no, why not?

13. In your assessment tasks (tests, assignments, exams) in your first year at the university, how much writing did you have to do?

- a. None (e.g., mostly multiple choice questions)
- b. Hardly any
- c. A fair amount
- d. A lot (all assessments required long written responses)

14. What proportion of your first year assessments were exams?

- a. 100%
- b. 70%
- c. 50%
- d. 20%
- e. 0%
- f. Other _____

15. The assignments last year were mostly

- a. Group assignments
- b. Individual assignments

16. What was the average lengths of your assignments

- a. Less than 2000 words
- b. 2000 words
- c. 3000 words
- d. More than 3000 words
- e. Other _____

17. In the last year, if you searched the internet for pleasure (not for university work), how often was it in English?

- a. always
- b. often
- c. sometimes
- d. rarely
- e. never

18. In the last year, if you searched the internet for university work, how often was it in English?

- a. always
- b. often
- c. sometimes
- d. rarely
- e. never

19. In the last year, if you wrote an email, how often was it in English?

- a. always
- b. often
- c. sometimes

- d. rarely
 - e. never
20. In the last year, if you sent a text message, how often was it in English?
- a. always
 - b. often
 - c. sometimes
 - d. rarely
 - e. never
21. In the last year, if you visited a chat room, how often was it in English?
- a. always
 - b. often
 - c. sometimes
 - d. rarely
 - e. never
22. In the last year, if you read a book for pleasure, how often was it in English?
- a. always
 - b. often
 - c. sometimes
 - d. rarely
 - e. never
23. In the last year, if you watched TV, how often was it in English?
- a. always
 - b. often
 - c. sometimes
 - d. rarely
 - e. never
24. How would rate your ability to speak and understand English?
- a. very strong, like a native speaker
 - b. strong, but occasional problems
 - c. acceptable, can usually express myself and understand others but not as well as in my native language
 - d. weak, frequent problems
 - e. very weak, can say and understand very little

25. How much of your daily communication with other people in Australia is in English?
- a. all of it or nearly all of it
 - b. most of it
 - c. about equal use of English and another language
 - d. most of it is in another language
 - e. all of it or nearly all of it is in another language