

### ***Abstract***

While universities devote great effort to initial EAP instruction, many question the effectiveness of such instruction on student production. The present study seeks to determine whether EAP instruction results in a longitudinal linguistic variation in the direction of the established norms of an academic register, which by extension, would provide a quantifiable linguistic measure of EAP course effectiveness. This paper adopts a multidimensional analysis (Biber, 1988) of a longitudinal corpus of written EAP essays and reports totalling 213,408 words, collected from freshman Chinese undergraduate students at a university in Hong Kong. The data was collected over a semester's EAP training at three data points (pre-EAP training, immediate post-training and final written examination). The results of the multidimensional analysis exhibit considerable variation between data points in the direction of academic discourse across all five dimensions analysed, including a drop in the use of first person pronouns and the mechanical use of discourse connectives, alongside an increased emphasis on nominalisation and more careful, hedged, presentation of stance. The findings suggest a warmly positive effect of EAP instruction on learner production after only a single semester. A number of pedagogical opportunities for the data are also outlined including the benefits of such analysis for written corrective feedback and future analysis of discipline-specific L2 discourse.

**Keywords - Multidimensional analysis; Learner corpora; EAP course effectiveness; Longitudinal data**

#### ***1. Introduction – The EAP context in Hong Kong.***

As English is the dominant language in Hong Kong (HK) tertiary settings, university students must quickly adopt the communicative skills required to participate in an academic environment. Universities devote great effort to providing training in English for academic purposes (EAP) to bridge the gap between secondary and tertiary expectations of academic discourse before adopting the discipline-specific nuances of particular subject areas (e.g. Hyland, 2000; Hyland and Hamp-Lyons, 2002), marking the development of academic literacy up to graduation and beyond. Where English is the medium of instruction (as in HK), the success (or failure) of EAP training is as crucial to a student's eventual progress as is training in subject-content knowledge. However, in HK, differences in local and international standards for EAP (i.e. the HK secondary school exam [HK-DSE] versus international tests such as IELTS®), a shift from English to Chinese as the medium of instruction in HK secondary schools, as well as a

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

shift from a 3-year to 4-year undergraduate curriculum (resulting in one year less of secondary education), have prompted many HK universities to revisit their EAP provision. Students have been shown to lack the linguistic means to fully participate in tertiary education (Bruce and Hamp-Lyons, 2015) and to struggle to understand technical vocabulary, listen to lectures, write in an appropriate style and conform to the conventions of academic discourse (Evans and Morrison, 2011).

Despite these issues, expectations regarding the effectiveness of initial EAP programmes remain 'unrealistic' (Bruce and Hamp-Lyons, 2015), and both teachers and students in HK and beyond express dissatisfaction with the effectiveness of EAP training (e.g. Abdolrezapour and Tavakoli, 2013). EAP course effectiveness is typically measured by test scores and student evaluations of course content and teacher performance. However, such measures are both summative and subjective in nature, and while face validity is an important factor when determining the authenticity of EAP courses (meaning both students and their institutions take them seriously, Fulcher, 1999), students are untrained to offer informed opinions on course effectiveness and teacher performance, and in-house tests may lack content/construct validity and reliability. Moreover, individual EAP teachers generally teach individual groups of students for a single semester, and so continual tracking of performance from initial EAP training into discipline-specific provision is lacking. In short, current measures of EAP effectiveness, at least in the HK context, do not take into account the *actual linguistic data* produced *over time* by students, tracking their development from pre- to post-training. It is thus necessary to collect and analyse data that chart which aspects of EAP training are effective and which remain problematic over a student's university life. In this regard, a corpus-based approach to measuring EAP course effectiveness is thus timely and advantageous.

## ***2. 'Multiple affordances' of corpus-based research for EAP: Multidimensional analysis***

The use of native language (L1) corpora (and their second language [L2] counterparts *learner corpora*) in English language teaching is now described as a 'marriage' rather than a 'fling' (Gabrielatos, 2005), and both native language and learner corpora are considered essential tools driving innovation in English language research, pedagogy, assessment and publishing (Hyland and Wong, 2013). In particular, the 'multiple affordances' (Lenko-Szymanska and Boulton, 2015) of language corpora-driven research on EAP have 'increased dramatically' over the past ten years (Csomay, 2015) in terms of materials development (e.g. Chuang and Nesi, 2006, 2007; Alexander, 2007; Jones and Durrant, 2010), analysis of academic genre and register (Biber, 1988, 1995; Hyland and Milton, 1997; Flowerdew, 2006; Hyland and Tse, 2007; Gardner & Nesi, 2012), contrastive interlanguage analysis of the processes behind L2 acquisition (Granger, 1996, 2015, Crosthwaite, 2016, accepted), dictionaries and reference grammars (e.g. Biber et al., 1999), data-driven learning via corpus-based in-class activity (e.g. Johns, 1991; Charles, 2007; Cotos, 2014) and the revision of EAP writing (Tono, Satake and Miura, 2014; Quinn, 2015).

In particular, the work of Doug Biber and his colleagues represents a fundamental corpus-based understanding of the practice, process and product of EAP in the form of quantitative multidimensional analyses of linguistic variation in register and genre. Biber's (1988) framework exploring variation across speech and writing made explicit the situational, functional nature of discourse, the impact of said function on linguistic form, and the connection between variation, situation and language use. Biber, Conrad, Reppen, Byrd, Helt, Cortes,

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

Csomay and Urzúa (2004), Biber (2006) and Biber and Conrad (2009) have all performed multivariate statistical studies investigating the linguistic features found in tertiary settings. In particular, Biber (1988) established a set of ‘dimensions’ of the variation of the normalised frequencies of particular linguistic features, along which a particular set of text types (according to genre or register) can be said to adhere (Biber, 1989). These dimensions and text types are briefly described below:

*Table 1: Summary of Biber’s (1988, 1989) dimensions and text types, adapted from Nini (2015:6-8).*

Dimension <sup>1</sup>	Description	Associated text types (high/low score)
1- Involved vs. Informational production	High scores – Affective / interactional (many verbs, pronouns) Low scores - Informationally dense (many nouns, adjectives)	High=Intimate interpersonal interaction, informational interaction Low=Scientific exposition, learned exposition, general narrative exposition
2 - Narrative vs. Non-Narrative Concerns	High scores – Narrative text (many past tenses, third person pronouns)	High=Imaginative narrative, general narrative exposition
3 – Explicit vs. Situation dependent reference.	High scores – context-independent, e.g. academic prose (many nominalisations) Low-scores – context-dependent, e.g. sports broadcast (many adverbs)	High=Scientific exposition, learned exposition Low=Intimate interpersonal interaction, informational interaction, imaginative narrative, situated reportage
4 - Overt Expression of Persuasion	High scores – Author’s point of view is explicitly marked, contains hedging and boosting of stance (many modal verbs)	High=Involved persuasion Low=Situated reportage
5 -Abstract vs. Non-Abstract Information	High scores – text is highly technical, abstract or formal, e.g. scientific discourse (many passive clauses and conjuncts)	High=Scientific exposition, learned exposition Low=Intimate interpersonal interaction, Informational interaction

<sup>1</sup> There is a 6<sup>th</sup> Dimension in Biber (1988), ‘On-Line Informational Elaboration’, but this is not featured in Nini’s (2015) MAT tagger’s graphical output and will not be referred to again in this paper.

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

The normalised (e.g. instances per 1000 words) frequencies of certain linguistic features thus correlate (or not) with certain dimensions, and the variation inherent determines the classification of an individual text type. Studies under a multidimensional approach have either used Biber's or other's (e.g. Biber and Conrad, 2009; Hardy and Römer, 2013; Gardner, Biber & Nesi, 2015) dimensions, or have performed exploratory factor analysis on the linguistic features present in their own data to derive their own dimensions (e.g. Grieve, Biber, Friginal and Nekrasova, 2010; Weigle and Friginal, 2015).

In terms of L2 multidimensional analyses, the derivation of dimensions for L2 academic discourse 'offers unique information about the linguistic choices of L2 writers that have not yet been extensively surveyed in corpus linguistics' (Friginal and Weigle, 2014:82). Friginal and Weigle (2014) performed a longitudinal multidimensional analysis using exploratory factor analysis of 209 L2 essays across four L2 dimensions, noting a significant impact of time on variation. Students produced texts that were more informational, elaborate and impersonal in nature as their studies progressed, 'reflecting the nature of academic writing' (2014:94). Weigle and Friginal (2015) used multidimensional analysis for the validation of L2 writing assessment comparing L2 texts with those of L1 writers, and found significant effects of L1 background and L2 proficiency on the linguistic variation found in academic essays. They also suggested that writing produced under timed test conditions was significantly different from that produced with no time constraints. Biber, Conrad, Reppen, Byrd and Helt (2002) looked at register variation in tertiary TESOL materials using the TOEFL 2000 Spoken and Written Academic Language Corpus, and found that the oral register used by teachers during classroom teaching was similar to conversational registers rather than the norms of academic language. This suggested that EAP students have to deal with a wide variety of registers in an English-medium of instruction tertiary

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

setting. Aguado-Jiménez, Pérez-Paredes and Sánchez (2012) explored a multidimensional analysis of L1-L2 English pairs completing the same tasks, and after becoming ‘researchers of their own learning’ (2012:92), learners were able to consider the importance of register, many of them for the first time. In following this corpus-based, data-driven approach, teachers are able to ‘define areas that need special attention in specific contexts and at different levels of competence, and so devise syllabi and materials’ (Gabrielatos, 2005:6), and should develop more insight into the *process* of learning as evidenced in authentic L2 production. However, Aguado-Jiménez et al. suggest that the use of multidimensional analysis is still ‘underused’ as a tool for pedagogy (2012:92).

With the exception of Friginal and Weigle (2014), one area in which multidimensional analysis (and corpus-based research generally) remains scarce is that of longitudinal studies. These would allow for the quantification of linguistic *development* and, by extension, allow researchers to assess the impact of instruction and/or learner context situations on such development. However, corpus-based research is beginning to address this issue. Li and Schmitt (2009) tracked a single student over one semester of EAP training as they acquired 166 new lexical phrases (despite relying on a limited range of phrases overall). Meunier and Gentil (2014) tracked French and Spanish L2 English learners’ development of nominalisation using the newly constructed Longitudinal Database of Learner English (LONGDALE, Meunier and Littré, 2013), noting key developmental milestones. Gardner & Nesi (2012) have used Biber’s (1988) dimensions to show that student writing (by native and non-native speakers) becomes increasingly informational and elaborated, with fewer narrative and explicitly persuasive features from first to final year in undergraduate studies, although with significant variance between disciplines and genres.

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

However, existing learner corpora of Asian/HK EAP data still lack a truly longitudinal scope, with corpora including the International Corpus Network of Asian Learners of English (ICNALE; Ishikawa, 2011, 2013) and the Cambridge Learner Corpus (Nicholls, 2003) adopting a pseudo-longitudinal (i.e. by L2 proficiency level) approach to linguistic development. Whilst useful for categorising L2 language norms at the macro-level, such corpora do not fully capture the language *learning* process of individuals at a more contextualised level, which can only be captured longitudinally. In doing so, one could theoretically assess EAP course effectiveness via a multidimensional analysis – charting the longitudinal impact of EAP instruction on L2 learner production from pre-training to the student’s final tests, and where the context, materials and method of instruction are held constant for all data in the corpus. Such a study necessitates the construction of a new EAP corpus, one that tracks students across their crucial first few months of tertiary education. This is the gap that the present study aims to fill.

In summary, the present study seeks to analyse the longitudinal impact of EAP instruction on undergraduate writing, comparing linguistic variation across three data points (pre-EAP, post-training, and final assessment) using Biber’s (1988) dimensions for multidimensional analysis as utilised in the recently released Multidimensional Analysis Tool (MAT, Nini, 2015). The research question to be answered is whether EAP instruction results in longitudinal linguistic variation between the three points in the direction of the established norms of an academic register. A move in this direction would be quantifiable evidence – in concrete linguistics terms – of the impact of EAP course provision on student writing.

### **3. Method**

#### **3.1 Corpus sample**

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

The HKU-CAES learner corpus was constructed as part of a pilot study into the effectiveness of undergraduate EAP provision, initially focusing on 20 error types, use of textual and interpersonal metadiscourse (following Hyland, 2000), and development of citation and referencing practices. The corpus data is drawn from three tasks taken at three data points over the semester-long EAP course. The total word count across all three data points was 213,408 words.

*Table 2: Corpus Sample*

Data point	Description	Word limit	Number of texts	Total/Av. Word count
1 – Pre-EAP training	A ‘diagnostic’ writing task (Essay or report) taken in the first week of EAP classes. Exam conditions for 90 mins. Not graded.	800	87 Essays =56 Reports = 31	46659 / 536.31
2 – Post-EAP training	Participants take the essay question if they did the report at data-point 1, or vice versa if they did the essay. Done as homework assignment to be submitted via <i>Turnitin</i> ® anti-plagiarism software, worth 25% of total EAP course grade. Submitted in week 10 of the EAP course, after the final writing class of week 9.	800	84 Essay = 29 Reports =55	70090 / 834.40
3 – Final assessment	A written essay or report produced under exam conditions in week 13, for 3 hours. Worth 35% of the total EAP course grade.	1500	86 Essays = 28 Reports = 58	96659 / 1123.94

Of note is that the data collected on data points 1 and 3 are produced under timed test conditions, which, as Wiegler and Friginal (2015) suggest, can result in texts that are different to those produced with no time constraint. Therefore, we need to be careful when considering the significance of variation between data points 1 and 2 (and also between 2 and 3). However, given



Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

that both data points 1 and 3 use data under test conditions, the linguistic variation between these two points should still be reliably comparable, and if the variation between data points 1 and 2 (and between 2 and 3) are not significantly different, then we may discount the effect of test conditions on our interpretation of the results.

The prompts for data points 1 and 2 depended on whether the students took the essay or report prompt for data point 1 (they would do the other available prompt for data point 2). The essay prompt:

*“Write an academic essay answering this question: ‘Should the death penalty be restored in Hong Kong?’ Before its concluding paragraph, your essay should answer the following questions:*

- 1) What are the main arguments for and against the death penalty?*
- 2) Should the death penalty be restored in Hong Kong?”*

The report prompt for data points 1 and 2 was:

*“Write an academic report answering this question: ‘What should be done about child labour in China?’ Before its concluding paragraph, your report should answer the following questions:*

- 1) What are the factors contributing to the problem of child labour in China?*
- 2) What should policymakers and private enterprises do (or continue doing) to solve this problem?”*

For data point 3, the students took either an essay or report prompt. The essay prompt read:

*Write an academic essay about the following topic: ‘Experiential Learning in Higher Education’. Before its concluding paragraph, your essay should:*

- 1) Explain the concept of experiential learning;*

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

- 2) *Discuss the benefits and challenges of implementing experiential learning in higher education; and*
- 3) *Evaluate the extent to which experiential learning can be incorporated as part of the curriculum at the University of Hong Kong*

The report prompt read:

*Write an academic report about the following topic: 'Campus Sustainability'. Before its concluding section, your report should:*

- 1) *Identify three good practices in campus sustainability at the University of Hong Kong*
- 2) *Discuss the potential challenges of sustainable development on campus; and*
- 3) *Recommend and fully justify three ways to improve sustainable practices on the University of Hong Kong campus.*

The issue of different writing prompts is also a factor in Wiegler and Friginal (2015), and in the present study, data-points 1, 2 and 3 are collected under different prompts for both essay and report task types. However, given that this data was collected as part of the students' regular EAP training (and not under experimental conditions), the impact of prompt on linguistic variation across data points could not be controlled for in the present study.

In terms of the potential effect of genre on the data (essay vs. report), both types contain some accompanying pre-reading with data and sources that should be cited in the written product. The conclusion sections of each are intended to be different, with the essay questions ask for the main stance to be rephrased, while the report questions stress the need for recommendations. This difference would potentially influence the MDA findings, especially along dimension 4 (overt expression of persuasion). However, students generally also provided recommendations in their essay conclusions, and also produced essay-type conclusions after the recommendations

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

in their reports, so the potential impact of genre on the MDA results was minimised. This was born out in an analysis of the impact of genre (report vs. essay) on scores for Dimension 4, which was deemed insignificant (Mann-Whitney  $U=7,985$ ,  $Z=-.255$ ,  $p=.798$ ). In reality, the main difference between the two genres in this corpus was that of structure, with reports following a numbered system of sections with clear headings and subheadings, while essays were organised into paragraphs with organisation managed via lexical and logical means.

### *3.2. Participants*

The same participants submitted data at each data point, and so the three data points track the participants longitudinally from pre-EAP training to their final EAP tests. There were 87 participants spread across five classes (although some students did not hand in their data at data points 2 and 3). Three of the five classes were taught by one teacher, and the other two classes were taught by a second teacher (neither class was taught by the researcher). The students are from a range of faculties and majors, as all eligible students must take the course and so grouping by faculty/major is impractical. For eligibility, all students must take the course unless they pass a 'native English standard exemption test', which lasts for one hour and can be taken pre-semester, although only around 20-30 out of around 1,000 students pass this test.

Both teachers were native speakers of English with over ten years' experience teaching EAP, with one a female in her late 30s and the other a male in his late 50s. All classes followed an identical EAP program. All participants were undergraduates between 18-19 years of age, a mix of men and women. The L2 proficiency of the students ranged from IELTS 6.5-8 (6.5 being the minimum requirement to enter the institute in question), with no one class significantly higher or lower in L2 proficiency than any of the others. All participants agreed to have their

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

original data - normally submitted as part of the EAP program - copied and analysed for research purposes, and were told that (non-) participation in the corpus project would not affect their grades.

### *3.3. Instruction – the EAP Curriculum*

The EAP curriculum in question is one of a skills-based course, intended for undergraduate students of upper-intermediate levels (i.e. greater than IELTS 6.5) taking their first course in academic English for post-secondary education. The EAP program in question runs for 13 weeks, covers both academic speaking and writing, and totals an average 1,500 students per semester over an average 30 teachers, and students can take the course in semester 1 or 2 of their first year. Each weekly session has three hours of in-class instruction following an in-house produced text book, and some mandatory online modules covering ‘avoiding plagiarism’ (n=6), ‘academic vocabulary’ (n=10) and ‘grammar’ (n=10). The in-class component covers academic skills, and thus has a ‘focus on form’ rather than ‘focus on forms’ (Long, 1991), while the online components deal more with specific linguistic issues such as ‘nominalisation’, ‘cohesion’, ‘pronominal referencing’, etc. Details of the program week-by-week are given below.

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

*Table 3: EAP curriculum: Core University English*

Session	Description	Tasks/Content <sup>2</sup>	Homework modules
Week 1	Pre-EAP training. Course introduction.	<b>Data point 1 – diagnostic writing task</b>	Plagiarism X 1 Vocabulary X 1
Week 2	Writing session 1	Recognising features of academic writing. Evaluating academic sources. Identifying types of supporting evidence.	Plagiarism X 1 Vocabulary X 1 Grammar X 1
Week 3	Speaking session 1	Integrating academic sources. Recognising purpose of tutorial discussion	Plagiarism X 1 Vocabulary X 1 Grammar X 1
Week 4	Writing session 2	Analysing topics. Synthesizing and linking ideas. Note-taking and paraphrasing. Referencing multiple sources.	Plagiarism X 1 Vocabulary X 1 Grammar X 1
Week 5	Speaking session 2	Identifying differences between spoken and written texts. Transforming written language into spoken language.	Plagiarism X 1 Vocabulary X 1 Grammar X 1
Week 6	Writing session 3	Identifying features of successful academic stance. Writing with an academic tone. Integrating counter-arguments and rebuttals.	Vocabulary X 1 Grammar X 1
Week 7	Speaking session 3	Expressing agreement and disagreement with other's stance. Using questions to make discussion more critical.	Vocabulary X 1 Grammar X 1
Week 8	Writing session 4	Logically connecting ideas in a paragraph. Writing appropriate section headings. Connecting ideas through cohesive devices	Vocabulary X 1 Grammar X 1
Week 9	Writing session 5	Applying a range of structural features to organise a text. Recognising differences between essays and reports. <b>Data point 2 – Post EAP training</b>	Vocabulary X 1 Grammar X 1
Week 10	Mock speaking test	Mock tutorial speaking test (25 mins per group).	Vocabulary X 1 Grammar X 1
Week 11	Speaking test	Speaking test	
Week 12	Feedback on written task for data point 2	One-on-one feedback on writing produced for data point 2	
Week 13	Writing test	<b>Data point 3 – Final test</b>	

The online grammar components include units on nominalisation, noun phrases, active vs. passive voice, pronoun referencing, tenses, hedging/boosting, linking words, prepositions, articles, and subject-verb agreement. The vocabulary components include units on academic word lists, prefixes/roots/suffixes, compound nouns, synonyms, collocations, and

<sup>2</sup> Wording taken verbatim from course book.

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

selecting/revising words. Each unit is divided into 5 tasks, and each unit takes approximately 1 hour to complete.

Only 9 of the 13 weeks involve in-class instruction by teachers (weeks 2-9, 12). Due to the relatively higher L2 proficiency level of the students at this institute (minimum IELTS 6.5, with most averaging 7-7.5) compared with that of other local institutes, this comprises the entire EAP training an undergraduate student receives in their first year before they take on more specific English-in-the-discipline courses (e.g. 'English for Dentistry', 'English Communication for Business', etc.) from their Sophomore years. Students take this course alongside their regular main degree subjects, although at this stage many students have not settled on their major and are free to swap, and very few faculties/majors have any kind of written/spoken assessments to work towards in year 1. There is also no additional organised orientation to studying at the university before students take the course.

### *3.4. Multidimensional Analysis Tagger*

All analysis was performed using Nini's (2015) Multidimensional Analysis Tagger (MAT), version 1.3. Three analyses were performed (one for each data point) using the 'Tag/Analyse' function, which uses the Stanford Parser (Socher, Bauer, Manning and Ng, 2013) to parse syntactic and part-of-speech information, then performs a multidimensional analysis on the parsed files replicating Biber's (1988) algorithms. The tagger has been tested for reliability on L1 data using the LOB and Brown corpora in Nini (2015), and the results of both were considered 'largely successful' (2014:13) in replicating Biber's (1988) analysis of these corpora. The Stanford parser has also been tested on L2 data, with Geertzen, Alexopolou and Korhonen (2013) suggesting that current natural language processing tools 'are robust to a significant part

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

of learner errors' (2013:13) due in part to the generally simplified nature of learner language compared to native norms. There will, of course, be some margin of error in any automated parsing process particularly with L2 data, but the Stanford parser is currently one of the most-widely used and accurate parsers available, and as such is the one included in Nini's (2015) MAT tagger.

The procedure produces tab delimited text files containing statistics for each of Biber's (1988) Dimensions, and statistics for each of the linguistic features analysed, which were then imported into Excel and then SPSS for analysis. The tagger also produces visual outputs of the corpora's data in comparison with the dimension features of Biber's (1988) text types, which have been included in the results.

The analysis followed Biber's (1988) default of 400 tokens for which the type-token ratio should be calculated. Each analysis used the 'Z-score correction' option in MAT, where the z-scores used to calculate the Dimension scores are checked for their magnitude. According to Nini (2015:4), 'If the absolute value of the magnitude is higher than 5, the program will change it to 5. This correction avoids the problem of few infrequent variables affecting the overall Dimension scores'.

One potential issue with running the multidimensional analysis on our corpus is related to the small number of files in each subcorpus. Friginal and Weigle (2014) used 207 tagged texts and suggested that ideally 350 texts should be used, although Bryant and Yarnold (1995) suggest that exploratory factor analysis can be performed with samples as low as 100. As our largest subcorpus is composed of 87 texts, claims regarding adherence to a particular dimension should be treated with caution, although the average word count per file and overall word count is

considerably larger than seen in Friginal and Weigle (2014). In addition, the MAT tagger is not producing an exploratory factor analysis but is determining how close a corpus is to Biber's (1988) dimensions, and as such a smaller sample size in this context is acceptable. Finally, any investigation / comparison of specific linguistic features across data-points is of course still perfectly viable.

#### **4. Results**

##### *4.1. Data point 1: Pre-EAP training*

###### *Figure 1: Pre-EAP training data*

[Figure 1 here]

The line in the figure highlights the dimension scores for the corpus at data point 1. The pre-EAP training data most closely resembles Biber's (1988) 'scientific exposition' type, where texts are 'typically informational expositions that are formal and focused on conveying information and [are] very technical' (Nini, 2015:7). This is categorised by low scores on Dimension 1 (Informational discourse), and high scores on Dimensions 3 (Context independent discourse) and 5 (Abstract information). To investigate further, a stepwise (via Wald) logistic regression analysis was performed using the normalised frequencies of linguistic features found in the 'scientific exposition' text type, to determine which features were the best predictors of that type when compared with all other text types (i.e. the constant). After 8 steps, the regression analysis identified 'Amplifiers' (*absolutely, extremely, etc.*,  $\beta = -2.996$ ,  $\exp \beta$  odds ratio = 0.05, Sig. = .003,  $M=0.15$ ,  $SD=.26$ ) and 'Causative adverbial subordinators' (*because, etc.*,  $\beta = -2.778$ ,  $\exp \beta$  odds ratio = 0.06, Sig. = .022,  $M=.14$ ,  $SD=.20$ ) as significant negative predictors of this text type. 'Conjuncts' (*however, rather, etc.*,  $\beta = 2.857$ ,  $\exp \beta$  odds ratio = 17.41, Sig. <.001,  $M=.89$ ,  $SD=.48$ ), 'Other adverbial subordinators' (*while, whereas, etc.*,  $\beta = 6.530$ ,  $\exp \beta$  odds ratio =



Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

685.66, Sig.<.001,  $M=.25$ ,  $SD=.23$ ), and ‘Possibility modals’ (*can, may*, etc.,  $\beta = 1.471$ , exp  $\beta$  odds ratio = 4.532, Sig. <.001,  $M=1$ ,  $SD=.60$ ) were significant positive predictors. These features are typical of academic discourse, although they may be used mechanically by novice EAP writers (Leedham and Cai, 2013; Granger and Tyson, 1996; Chen, 2015).

#### *4.2. Data point 2: Post-training (week 9)*

##### *Figure 2: Post-training data*

[Figure 2 here]

As with the pre-training data, the post-training data most closely resembles Biber’s (1988) ‘scientific exposition’ type. The mean z-scores for Dimension 1 are lower than were found pre-EAP training, while the mean z-scores for Dimension 3 are higher than were found pre-training, suggesting that the texts produced at this data point are more informationally dense and context-independent – both qualities are associated with academic genres.

#### *4.3. Data point 3: Final test*

##### *Figure 3: Final test data.*

[Figure 3 here]

For the third data point (final test), the data most closely match Biber’s (1988) ‘learned exposition’ type. Texts of this type are ‘typically informational expositions that are formal and focused on conveying information’ (Nini, 2015:7), and, like ‘scientific exposition’, is categorised by low z-scores on Dimension 1 and high z-scores on Dimensions 3 and 5. To investigate further, a stepwise (via Wald) logistic regression analysis was performed against the linguistic features of the ‘learned exposition’ text type to determine which features were the best predictors of that type when compared with all other text types (i.e. the constant). After 13 steps, the

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

regression analysis identified ‘Amplifiers’ (*absolutely, extremely, etc.*,  $\beta = -9.629$ ,  $\exp \beta$  odds ratio = 0.00, Sig. <.001,  $M=.10$ ,  $SD=.14$ ), ‘Conjuncts’ (*however, rather, etc.*,  $\beta = -4.354$ ,  $\exp \beta$  odds ratio = 0.13, Sig. <.001,  $M=.72$ ,  $SD=.32$ ), ‘Demonstrative pronouns’ (*this, these, etc.*,  $\beta = -4.812$ ,  $\exp \beta$  odds ratio = 0.008, Sig. =.001,  $M=.29$ ,  $SD=.23$ ), ‘Necessity modals’ (*ought, should, etc.*,  $\beta = -2.378$ ,  $\exp \beta$  odds ratio = 4.532, Sig. <.001,  $M=.19$ ,  $SD=.19$ ), ‘Pronoun it’ ( $\beta = 1.907$ ,  $\exp \beta$  odds ratio = -0.149, Sig. =.003,  $M=1.02$ ,  $SD=.48$ ), and ‘Possibility modals’ (*can, may, etc.*,  $\beta = -3.036$ ,  $\exp \beta$  odds ratio = 0.048, Sig. <.001,  $M=1.13$ ,  $SD=.46$ ) as significant negative predictors, and ‘Nominalizations’ (*-tion, -ment, etc.*,  $\beta = 0.433$ ,  $\exp \beta$  odds ratio = 1.542, Sig. <.012,  $M=7.46$ ,  $SD=1.89$ ) as a significant positive predictor. The negative values listed here do not suggest that conjuncts or demonstrative pronoun, etc. are *not* found in ‘learned exposition’ texts, as these texts will also have elements of ‘scientific exposition’ and other text types included (i.e. the constant against which the regression analysis was performed). Rather, the defining feature of ‘learned exposition’ in this regard is nominalisation, i.e. the more nominalisation occurs, the closer a text type will match ‘learned exposition’ and not any other text type. As evidence, the mean frequency of nominalisation in ‘learned exposition’ texts is  $M=6.40$ ,  $SD=2.28$ , compared with  $M=4.50$ ,  $SD=1.72$  across all other text types. Thus, by the time the students reach their final test, they are much more likely to produce texts with a high frequency of nominalisations, compared with their production at earlier data points.

#### 4.4. Measuring the effect of instruction: Cross-data point analysis

The following table compares the dimensional z-scores across the three data points. As the data is not always normally distributed when comparing dimensions between particular data

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

points, non-parametric Kruskal-Wallis tests were used in each case, and where significant, post-hoc multiple pairwise comparison is performed using Dunn's correction for multiple tests.

*Table 4: Cross data point comparison of dimension scores*

Dimension	DP1 (Pre-EAP)	DP2 (Post- training)	DP3 (Final test)	Kruskal-Wallis	Pairwise Comparison
1- Involved vs. Informational production	M=-11.64 SD=5.59	M=-15.19 SD=4.90	M=-15.53 SD=4.1	$H(2)=29.20, p<.001$	DP2<DP1, $t(2)=49.37, p<.001$ DP3<DP1, $t(2)=56.39, p<.001$
2 - Narrative vs. Non-Narrative Concerns	M=-2.21 SD=2.24	M=-2.98 SD=1.66	M=-3.85 SD=1.07	$H(2)=34.86, p<.001$	DP1>DP3, $t(2)=66.45, p<.001$ DP2>DP3, $t(2)=41.76, p=.001$
3 – Explicit vs. Situation dependent reference.	M=5.79 SD=2.73	M=7.68 SD=2.32	M=10.75 SD=1.89	$H(2)=117.57, p<.001$	DP1<DP3, $t(2)=-121.24, p<.001$ DP2<DP3, $t(2)=-80.47, p<.001$ DP1<DP2, $t(2)=40.76, p=.001$
4 - Overt Expression of Persuasion	M=1.92 SD=3.20	M=0.99 SD=2.28	M=-0.003 SD=2.16	$H(2)=18.26, p<.001$	DP1>DP3, $t(2)=48.14, p<.001$ DP2>DP3, $t(2)=29.98, p=.027$
5 -Abstract vs. Non-Abstract Information	M=5.36 SD=2.92	M=5.16 SD=2.50	M=4.29 SD=2.32	$H(2)=7.28, p=.026$	DP1>DP3, $t(2)=29.96, p=.033$

The results show clear and very significant effects of instruction across the three data points and across all five dimensions. For Dimensions 1,2,3 and 5, the general trend is that the student's production more closely matched the norms of academic discourse from pre-EAP training to the final test. This is evidenced in the significantly lower dimension scores for

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

dimensions 1 and 2 and the higher scores for Dimension 3 in the final test data compared to the pre-EAP data, which follow that of Gardner & Nesi (2012) under the same 5 dimensions.

### ***Dimension 5***

The lower score for Dimension 5 in the final test data compared with the pre-EAP data is interesting in that the writers appear to have included more features associated with interpersonal interaction in their final test production, which does not suggest a particularly academic register. To investigate further, a Pearson correlation matrix was performed with each linguistic variable analysed in MAT against the scores for Dimension 5 across the entire corpus, to determine which features were positively or negatively associated with that dimension in the data. Significant positive correlations included the features ‘Conjuncts’ (*however, rather, etc.*) and ‘Other adverbial subordinators’ (*while, whereas, etc.*), while significant negative correlations included ‘Downtoners’ (*almost, nearly, etc.*), ‘First person pronouns’, ‘Second person pronouns’, and ‘That relative clauses on object position’ (*the dog that I saw*). Kruskal-Wallis comparison was used to compare the occurrence of these features between data points, and found significant effects of time for the features ‘Conjuncts’ ( $H[2]=6.14, p=.046$ ), ‘Other adverbial subordinators’ ( $H[2]=7.85, p=.020$ ), and ‘First person pronouns’ ( $H[2]=36.49, p<.001$ ), each of which were more frequent pre-EAP training ( $M=.89/.72, SD=.48/.32$  for conjuncts between data point 1 and 3,  $M=.25/.16, SD=.23/.13$  for other adverbial subordinators, and  $M=.60/.24, SD=.64/.46$  for first person pronouns). The drop in the use of first person pronouns over data points is an obvious effect of EAP training, but perhaps more interestingly, it appears as though the students are using different grammatical means to organise their utterances instead of the mechanical use of lexical conjuncts (*however, rather, etc.*) and adverbial subordinators (*while,*

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

*whereas*) which might be considered more typical of their high school essays and reports, and are commonly overused by L2 writers (Leedham and Cai, 2013; Granger and Tyson, 1996; Chen, 2015). By way of example, this sample from data point 1 demonstrates a high frequency of lexical conjuncts (highlighted in bold):

[0082Q6-1.txt] - **Besides**, implementing death punishment is not a common and easy decision to make solid evidence with strong agreement of public should be one of the requirements. **Especially**, advanced science and technology with modern forensic techniques help us a lot in finding offenders, without catching innocent people. **Moreover**, the standards of implementing death penalty are much strict than before. **For example**, Ted Herring who was a murder that had nearly been executed because of his intellectual disability in the past. He, **however**, was given a chance to avoid death penalty, **since** intellectual disability had longer be the standard of execution. (Alvarez L. and Schwartz J., 2014)

This example from data point 3 from the same student shows a different approach to conjunction, where lexical conjunctions are now replaced with phrasal or left-dislocated sentence structures:

[0082Q6-3.txt] - 3.1 *Insufficient support and resources*

*For the success of sustainability development on campus, there are so many factors which will affect the possibility of success, such as financial support, government policy, culture and climate (Clugston, 2004 and Giulio, 2006, as cited in Kildahl and Liao, 2013). **Apart from the above considerations**, the priorities for the limited resources*

would also hinder the development. **Because of the competition among other international universities**, universities may choose to develop other aspects such as improving quality of teachers and facilities that the importance of sustainability development would be ignored (Kildahl and Liao, 2013) and only little resource for sustainability is gained.

#### **Dimension 4**

An unexpected finding is that there is significantly more evidence of overt expression of persuasion (Dimension 4) in the pre-EAP training data compared with the later data points. This suggests significantly fewer occasions where the author explicitly marked their point of view, or at least tended not to hedge or boost their claims where necessary. Given that providing evidence of the author's stance is supposed to represent 40% of the final writing (and speaking) assessment, as well as being a central feature of the EAP course from weeks 6-9, this finding suggests that, when taken at face value, the student's stance is as (or less) discernible post-training as it was in their pre-EAP training data.

To investigate further, a Pearson correlation matrix was performed using each linguistic variable analysed in MAT against Dimension 4 across the entire corpus. For dimension 4, significant positive correlations included the features 'Conditional adverbial subordinators' (use of *if* or *unless* clauses), 'Downtoners' (*almost*, *nearly*, etc.), 'Necessity modals' (*ought*, *should*, etc.), 'Predictive modals' (*will*, *would*, etc.), 'Time adverbials' (*afterwards*, *later*, etc.), 'Infinitives' and 'Third person pronouns', while significant negative correlations included the features 'Independent clause coordination' (*and this*, etc.), and 'Nominalizations' (*-tion*, *-ment*, etc.). Kruskal-Wallis comparison was used to compare the occurrence of these features between

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

data points, and found significant effects of instruction for the features ‘Conditional adverbial subordinators’ ( $H[2]=16.17, p<.001$ ) and ‘Necessity modals’ ( $H[2]=84.35, p<.001$ ), where these were more frequent pre-EAP training ( $M=.17/.07, SD=.19/.12$  for conditional adverbial subordinators between data point 1/3, and  $M=.67/.19, SD=.42/.19$  for necessity modals). There were also significant effects of instruction on ‘Nominalizations’ ( $H[2]=146.37, p<.001$ ) and ‘Predictive modals’ ( $H[2]=9.56, p=.008$ ), where these were more frequent in the final test data ( $M=3.60/7.46, SD=1.15/1.89$  for nominalisations between data points 1/3, and  $M=.42/.49, SD=.38/.23$  for predictive modals).

These findings suggest that in their final tests, students were less likely to make strong conditional statements or to come up with strong recommendations to resolve any problems posed by the essay or report prompts, avoiding sentences such as ‘*If [we do not do something about this]*’ or ‘*[The government] must [do something about this]*’, which are typical of emotional statements and perhaps less typical of careful, measured academic discourse. By way of example, this sample of a text from data-point 1 shows how strongly-worded student recommendations at this stage are:

[0102Z4-1.txt] ***For sure** adopting the death penalty again could reduce the violent crime rate. However, if the government choose to do so, **it would change** our moral standard completely. **We would thereafter ask** should a more severe punishment be given to robbers, fullies, even those who just get late for school. The secondary consequences of restoring death penalty **are large and should not be neglected. I think** unless the violent crime rate is so high that we cannot put up with anymore, say, everyday our lives are threatened, **death penalty should not be restored** for the sake of a more controversy topic*

*secondary consequences. Though the violent crime rate in Hong Kong is low and therefore the capital costs suggested by other side can be somehow neglected, **the idea should be abolished first** before the death penalty because of the value of lives that can be saved suggested by the deterrent effect. Nevertheless, the death penalty **should by no means** be restored by a more powerful argument of secondary effect. **And I think** the question lasting years **should end** sometime here.*

Rather, students at data point 3 were more likely to make statements such as ‘*[this] would [effectively resolve the problem]*’, which are arguably more logical and cautious statements after the presentation of evidence in support of the writer’s stance. This is exemplified by a sample of the same student’s recommendations at data point 3

[0102Z4-3.txt] - *In the light of the above constraints and challenges, the followings are recommendations to improve such development in campus. Offering lessons the university dwellers **could** raise their awareness towards the issues. And this has been run in other universities such as University of Georgia (Levy and Marans, 2012). Planning for long-term sustainable development in itself means long-term development with balancing economic, social and environmental interests (Walter, 2014). Therefore, planning for long-term **may be** of paramount importance **when** it comes to sustainable development in campus. As mentioned, governments **often** fail to provide adequate sources to universities for sustainable development. Therefore, the report suggests HKU **could** ask more from the government, especially on the issue of cross-regional pollution.*

Thus, even though there is less ‘evidence’ of stance in terms of frequency or the emotion provoked by the writer’s statements, it appears as though the EAP training has in fact resulted in



a more cautious presentation of stance overall - a key feature of the instructional component of the EAP course for a number of weeks.

## **5. Discussion**

Despite the ‘unrealistic’ expectations of (and associated dissatisfaction with) the effectiveness of initial EAP programmes, the findings of the present study may be interpreted as warmly positive for both educators and students alike. Namely, that in a period of only 13 weeks (of which only 9 were ‘taught’), student’s written production exhibited sufficient and significant variation towards a more academic register, at least as proposed under the multidimensional framework adopted. While the improvement is evidenced in the student's writing, it is also possible that the students' spoken production will exhibit a more academic-like register over time, and further data collection is forthcoming in that regard.

It is, however, also possible that there is more to the variation than a simple effect of EAP instruction alone. Namely, students are also studying their subject content at the same time, although they are not typically asked to write about them at this stage. Moreover, the initial undergraduate experience allows for students to join a range of academic clubs and societies, and as the institution in question has a large number of non-Chinese speaking international students, interaction between domestic and international students both in and out of class may also be a source of variation (although this is more likely to be evidenced in the spoken register). Despite these concerns, given the analysis provided here, one could be reasonably satisfied that the EAP course in question has certainly contributed to the direction the students' production has now taken.

If this is the case, the question that remains is not ‘whether the course worked’, but ‘what’ it was about the course that ‘worked’. Certainly, the ‘focus on forms’ activities for many of the linguistic features involved (nominalisation, etc.) were sourced in the online components. There has been a continued shift in EAP language education towards the ‘flipped classroom’ model (e.g. Milman, 2012; Sung, 2015; Leis, Cooke, and Tohei, 2015) where instruction on grammar is performed online outside of class time, freeing up class time for actual writing practice, feedback and discussion. This is also the case at the institution featured in the present study (Hyland, 2014). However, the online component only represents the focus on linguistic form, and in fact, not all students sampled in the present study took the same vocabulary or grammar modules for the online component. Rather, the focus on form, where the individual linguistic items are contextualised, is managed via careful selection and editing of the authentic academic reading materials included in the course textbooks and examination papers, alongside the choice and sequencing of tasks that ask students to consider the key rhetorical and organisational features EAP writers need to master, and the feedback given to students’ spoken and written production. The evidence from the present study suggests that by ensuring that the online and in-class materials, tasks and feedback practices are synthesised to promote as many opportunities for using the appropriate target language forms as possible, a positive outcome can be achieved.

The data provided here should also be useful for educators who wish to conduct student-led data-driven analysis of register variation, as with the approach taken by Aguado-Jiménez, Pérez-Paredes and Sánchez (2012). The author of this paper, upon presenting the findings to the course co-ordinator and teachers involved, discussed ideas with these stakeholders regarding how students could be made aware of register variation using the data collected. As the Nini

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

(2015) MAT tool drastically cuts down the amount of time and analysis needed to conducted multidimensional analysis, it was suggested that teachers and even students could perform their own analyses using the tool - both in and out of class – so that students would have visual feedback about how closely their writing follows that of established text types (of which ‘learned/scientific exposition’ should be the targets for EAP). The teachers also suggested that with enough statistical support, they would be able to use the data in the form of targeted written corrective feedback to students, if a suitable ‘benchmark’ of comparable professionally-written English essays and reports (using the same task prompts) could be generated. Students could then get an idea about how far (or how close!) their writing was from that of professionally-written discourse, and have this information presented in concrete linguistic terms, together with the more general holistic comments on overall performance that students receive from their EAP tutors. This was considered as particularly useful for the EAP course in question, if the multidimensional analysis between learner and professionally-written texts could be performed using data collected at data point 2 (the first assessed written task) and the feedback received in time for students to consider before their final tests. Given the increased importance of register/genre variation (and, hence of multidimensional analysis research) for discipline-specific linguistic concerns as students reach their post-EAP English-in-the-discipline training in their sophomore years, teachers also considered the potential usefulness of multidimensional analysis for feedback on writing as extending far beyond the initial EAP training analysed in the present study.

Aside from increasing the number of texts for analysis, one obvious consideration for improving on the present study is that the Nini (2015) MAT tagger uses Biber’s (1988, 1989)

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

dimensions and text types for analysis. It would be preferable in future analyses of EAP discourse to use Biber and Conrad's (2009) dimensions for L1 academic discourse, or, preferably, Gardner, Biber & Nesi's (2015) dimensions for successful student writings, so as to more carefully delineate the variation exhibited as a result of instruction along the established norms of academic text types. The approach of analysing 'scientific' vs. 'learned' exposition undertaken in the present study may then be more accurately linked to established academic registers. It may also be advisable to include a comparable L1 set of data (making the study more of an Integrated Contrastive Model, e.g. Granger, 1996), although as the L1 students at our institute are exempt from the EAP course in question they would have to be contacted individually to participate. However, these concerns should not distract too much from the findings of the present study which have suggested a significant statistical impact of instruction on language variation, regardless of which of Biber's text types (or L1 data) were eventually considered as benchmarks.

Another option is for educators to conduct an exploratory factor analysis on their own EAP courses in order to determine a set of dimensions that characterise the production of their specific context. This does not mean that the data should be used to determine context-specific strategies of how to improve course provision without trying to meet any *a priori* established norms of academic text types, which would lead to a circular approach to course development where improvements are only made to what students are already producing, rather than what they could produce. However, the opportunity for using EFA so as to create a benchmark by which variation between one EAP context to another can be determined would be very useful for educators as statistical evidence for new directions in materials development and pedagogy.

A final comment about the methodology taken in this paper is that it is not the author's intention to suggest that all that is needed to improve L2 academic production is to guide students to reproduce in their production a set number of linguistic features so as to bring about a statistical correlation with an L1 or professional academic English corpus. This would be potentially worrying given that a lot of the 'focus on forms' instruction about these linguistic features takes place online, and so the case could be made for replacing the EAP teacher entirely. Rather, the value of the MDA approach is to complement, rather than replace, existing holistic and formative appraisals of student performance and of EAP course effectiveness, given that EAP is more than simply giving students a list of linguistic features of memorize.

## **6. Closing comments**

The present study has used a multidimensional analysis of a longitudinal corpus of undergraduate EAP essays and reports in order to determine whether the EAP course in question was effective in helping students to develop a more academic register. The positive results, gained from only a single semester's instruction, should be hugely encouraging to educators and students alike, and the potential for further analyses as students enter their discipline-specific English language programs is certainly promising. The multidimensional approach taken in this study has provided a useful and quantifiable window into the effects of EAP instruction over time, and by extension, the data can be used to determine the relative effectiveness of such instruction as compared with other forms of instruction on other courses, at least in terms of the effectiveness of instruction on L1 or L2 language related issues. Such a comparison would allow for the curriculum and pedagogical format of such courses to be fine-tuned and validated based

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

on actual linguistic evidence, rather than holistic judgement alone. This may lead to more effective, valid, and reliable practice in both instruction and assessment.

## 7. References

- Abdolrezaipoor, P. & Tavakoli, M. (2013). University teachers and students' perceptions of EAP methodologies and their effectiveness. *The Social Sciences*, 8(1), 49-54.
- Aguado-Jiménez, P., Pérez-Paredes, P. & Sánchez, P. (2012). Exploring the use of multidimensional analysis of learner language to promote register awareness. *System*, 40(1), 90-103.
- Alexander, O. (2007). Introduction. In O. Alexander (Ed.), *New approaches to materials development for language learning* (pp. 9-14). Oxford, UK: Peter Lang.
- Crosthwaite, P.R. (2016). Definite article bridging relations in L2: A learner corpus study. *Corpus Linguistics and Linguistic Theory*, awaiting print version, online version available Feb 2016, DOI: [10.1515/cllt-2015-0058](https://doi.org/10.1515/cllt-2015-0058)
- Crosthwaite, P.R. (accepted). L2 English Article Use by L1 Speakers of Article-less Languages: A Learner Corpus Study. *International Journal of Learner Corpus Research*.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. (1989). A typology of English texts. *Linguistics*, 27(1), 3-43.
- Biber, D. (1995). Dimensions of register variation: A cross linguistic comparison. Cambridge, CUP.
- Biber, D. (2006). *University Language. A Corpus-Based Study of Spoken and Written Registers [Studies in Corpus Linguistics, 23]*. Amsterdam: John Benjamins. DOI:10.1075/sci.23
- Biber, D. & Conrad, S. (2009). *Register, Genre and Style*. Cambridge, CUP. DOI:10.1017/CBO9780511841358.
- Biber, D. Conrad, S. Reppen, R. Bryd, P. & Helt, M. (2002). Speaking and Writing in the University: A Multidimensional Comparison. *TESOL Quarterly*. 36:9-48. DOI: 10.2307/3588359.
- Biber, D. Conrad, S., Reppen, R. Byrd, P. Helt, M. Cortes, V. Csomay, E. & Urzúa, A. (2004) *Representing Language Use in the University: Analysis of the TOEFL 2000 Spoken and Written Academic Language Corpus [TOEFL Monograph Series MS-26]*. Princeton, NJ: Educational Testing Service.
- Biber, D., Johansson S., Leech G., Conrad S. and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. London: Longman.
- Bruce, E. & Hamp-Lyons, L. (2015). Opposing tensions of local and international standards for EAP writing programmes: Who are we assessing for? *Journal of English for Academic Purposes*, 18, 64-77. DOI: 10.1016/j.jeap.2015.03.003
- Bryant, F., & Yarnold, P. (1995). Principal components analysis and exploratory and confirmatory factor analysis. In L. Grimm & P. Yarnold (Eds.), *Readings and Understanding Multivariate Statistics* (pp. 99-136). Washington, DC: American Psychological Association.
- Chen, M.Y. (2015). Spoken Discourse Markers in English Writing of Taiwanese Students. *Journal of Modern Education Review*, 5(5), 453-460.

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

- Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes*, 6(4), 289-302.
- Chuang, F-Y., & Nesi, H. (2006). An analysis of formal errors in a corpus of L2 English produced by Chinese students. *Corpora*, 1(2), 251-271.
- Chuang, F-Y., & Nesi, H. (2007). GrammarTalk: Developing computer-based materials for Chinese EAP students. In O. Alexander (Ed.), *New approaches to materials development for language learning* (pp. 315-330). Oxford, UK: Peter Lang.
- Cotos, E. (2014). Enhancing writing pedagogy with learner corpus data. *ReCALL*, 26(2), 202-224.
- Csomay, E. (2015). A corpus-based analysis of linguistic variation in teacher and student presentations in university settings. In Cortes, V & Csomay, E. (eds.) *Corpus-based Research in Applied Linguistics: Studies in Honor of Doug Biber* (pp. 1-23). Amsterdam, John Benjamins.
- Evans, S., & Morrison, B. (2011). The first term at university: implications for EAP. *English Language Teaching Journal*, 65(4), 387-397.
- Flowerdew, J. (2006). Use of signaling nouns in a learner corpus. *International Journal of Corpus Linguistics*, 11(3), 345-362.
- Friginal, E. & Weigle, S. (2014). Exploring multiple profiles of L2 writing using multi-dimensional analysis. *Journal of Second Language Writing*. 26, 80–95. DOI: 10.1016/j.jslw.2014.09.007.
- Fulcher, G. (1999). Assessment in English for Academic Purposes: Putting content validity in its place. *Applied Linguistics*, 20(2), 221-236.
- Gabrielatos, C. (2005). Corpora and Language Teaching: Just a fling or wedding bells? *TESL-EJ*, 8(4), 1-39.
- Gardner, S., & Nesi, H. (2012) 'A classification of genre families in university student writing'. *Applied Linguistics* 34 (1), 1-29. DOI:10.1093/applin/ams024
- Gardner, S., Biber, D., and Nesi, H. (2015) 'MDA perspectives on Discipline and Level in the BAWE corpus'. *Proceedings of the 8th International Corpus Linguistics conference, 'CL2015'*. Held 21-24 July 2015 in Lancaster, UK. 126-128
- Geerzten, J., Alexopoulou, T., & Korhonen, A. (2013). Automatic linguistic annotation of large scale L2 databases: The EF-Cambridge Open Language Database (EFCAMDAT). In *Proceedings of the 31st Second Language Research Forum (SLRF)*. Cascadilla Press.
- Granger, S. (1996), From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora, in K. Aijmer, B. Altenberg, and M. Johansson (eds), *Languages in contrast* (pp.37-51), Lund, Sweden: Lund University Press.
- Granger, S. (2015). Contrastive Interlanguage Analysis: A reappraisal. *International Journal of Learner Corpus Research* 1:1 (2015), 7–24. doi 10.1075/ijlcr.1.1.01gra
- Granger S. and Tyson S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English, *World English*, 15, 19–29.
- Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2010). Variation among blogs: A multi-dimensional analysis. In A. Mehler, S. Sharoff, & M. Santini (Eds.), *Genres on the web: Corpus studies and computational models* (pp. 45–71). New York: Springer-Verlag.

- Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.
- Hardy, J. A., & Römer, U. (2013). Revealing disciplinary variation in student writing: A multi-dimensional analysis of the Michigan Corpus of Upper-level Student Papers (MICUSP). *Corpora*, 8(2), 183–207.
- Hyland, K. (2000). *Disciplinary Discourses: Social Interactions in Academic Writing*. London, UK: Longman.
- Hyland, K. (2014). Re-imagining literacy: English in Hong Kong's new university curriculum. In Coniam, D. (ed.) *English Language Education and Assessment* (pp.139-151). Singapore: Springer.
- Hyland, K. & Hamp-Lyons, L. (2002). EAP: Issues and directions. *Journal of English for Academic Purposes*. 1, 1-12.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6(2), 183-205.
- Hyland, K. & Tse, P. (2007). Is there an 'academic vocabulary'? *TESOL Quarterly*, 41(2), 235-253.
- Hyland, K & Wong, L. (2013). *Innovation and Change in Language Education*. London: Routledge.
- Ishikawa, S. (2011). A New horizon in learner corpus studies: The aim of the ICNALE Project. In G. Weir, S. Ishikawa, & K. Poonpon (Eds.), *Corpora and language technologies in teaching, learning and research* (pp.3-11). Glasgow, UK: University of Strathclyde Press.
- Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In Ishikawa, S. (ed.), *Learner Corpus Studies in Asia and the World*, 191-118.
- Johns, T. (1991) From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10: 14–34
- Jones, M., & Durrant, P. (2010). What can a corpus tell us about vocabulary teaching materials? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics*. London, UK: Routledge.
- Leedham M. & Cai G. (2013). "Besides ....on the other hand": Using a corpus approach to explore the influence of teaching materials on Chinese students' use of linking adverbials. *Journal of Second Language Writing*, 22(4), 374–389.
- Leis, A., Cooke, S., & Tohei, A. (2015). The Effects of Flipped Classrooms on English Composition Writing in an EFL Environment. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 5(4), 37-51.
- Leńko-Szymańska, A. & Boulton, A. (eds.) (2015). *Multiple Affordances of Language Corpora for Data-Driven Learning*. Amsterdam: John Benjamins.
- Li, J. & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*. 18, 85-102.
- Long, M.H. (1991). Focus on form: A design feature in language teaching methodology. In de Bot, K., Ginsberg, R. & Kramsch, C. (eds.) *Foreign Language Research in Cross-Cultural Perspective*. Amsterdam: John Benjamins.
- Meunier, F. & Littré, D. (2013). Tracking learners' progress: Adopting a dual 'Corpus-cum-experimental data' approach. *Modern Language Journal*. 97(1), 61-76. DOI: 10.1111/j.1540-4781.2012.01424.x
- Meunier, F. & Gentil, G. (2014). Tracking learners' progress in nominalization use: a quantitative and qualitative longitudinal corpus analysis. Paper presented at *GURT 2014-*



Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

*Georgetown Roundtable in Linguistics: Usage-based approaches to language, language learning and multilingualism*, Washington, USA (14/03/2014-16/03/2014).

Milman, N. B. (2012). The flipped classroom strategy: What is it and how can it best be used?. *Distance Learning*, 9(3), 85-87.

Nicholls, D. (2003). The Cambridge Learner Corpus: Error coding and analysis for lexicography and ELT. In *Proceedings of the Corpus Linguistics Conference* (pp. 572–581). Lancaster University:University Centre for Computer Corpus Research on Language.

Nini, A. (2015). *Multidimensional Analysis Tagger (Version 1.3)*. Retrieved from: <http://sites.google.com/site/multidimensionaltagger>.

Quinn, C. (2015). Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, 69(2), 165-177.

Socher, R., Bauer, J., Manning, C.D., & Ng, A.Y. (2013). Parsing With Compositional Vector Grammars. *Proceedings of ACL 2013*.

Sung, K. (2015). A Case Study on a Flipped Classroom in an EFL Content Course. *Multimedia-Assisted Language Learning*, 18(2), 159-187.

Tono, Y, Satake, Y & Miura, A. (2014). The effects of using corpora on revision tasks in L2 writing with coded error feedback. *ReCALL*, 26(2), 147-162.

Urzúa, A. (2015) Corpora, context and language teachers: Teacher involvement in a local learner corpus project. In Cortes, V & Csomay, E. (eds.) *Corpus-based Research in Applied Linguistics: Studies in Honor of Doug Biber* (pp. 99-122). Amsterdam, John Benjamins.

Weigle, S.C. & Friginal, E. (2015). Linguistic dimensions of impromptu test essays compared with successful student disciplinary writing: Effects of language background, topic, and L2 proficiency. *Journal of English for Academic Purposes*, 18(3), 25-39. DOI: 10.1016/j.jeap.2015.03.006.

Figure 1:

Crosthwaite, P. (2016). A longitudinal multidimensional analysis of EAP writing: Determining EAP course effectiveness. *Journal of English for Academic Purposes*.

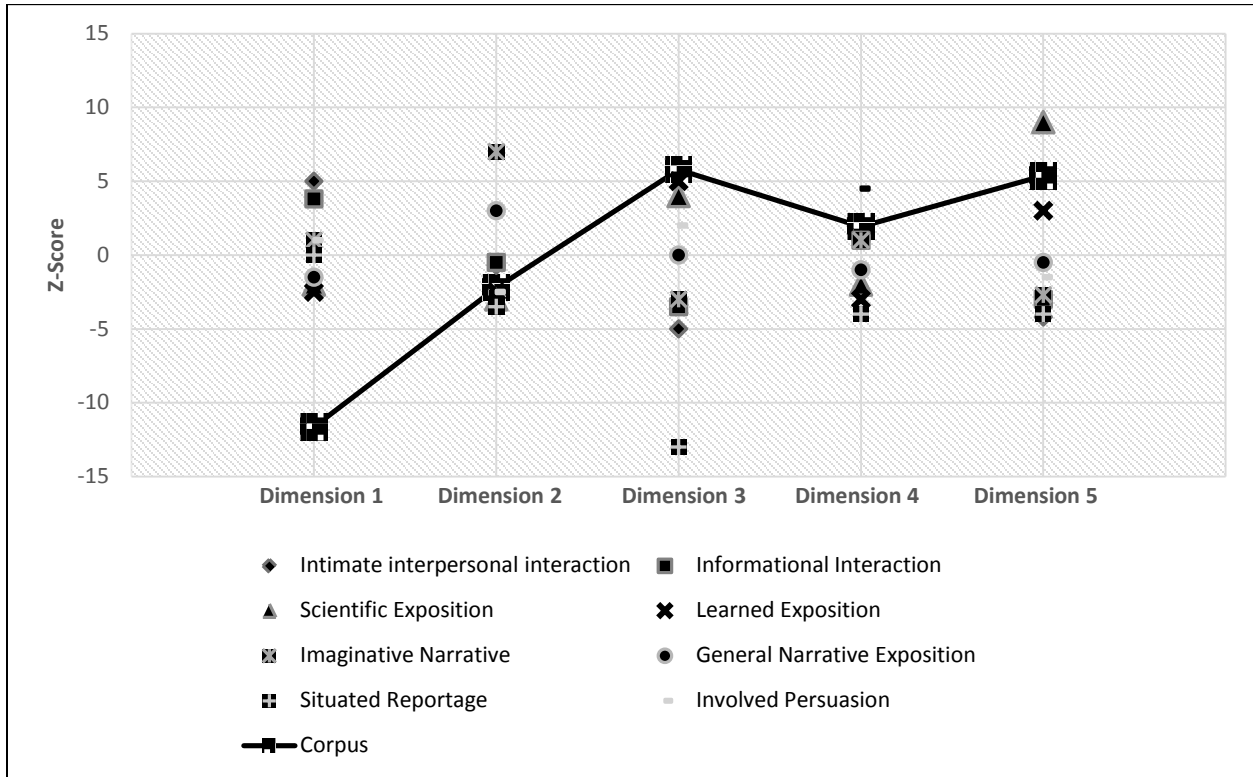


Figure 2:

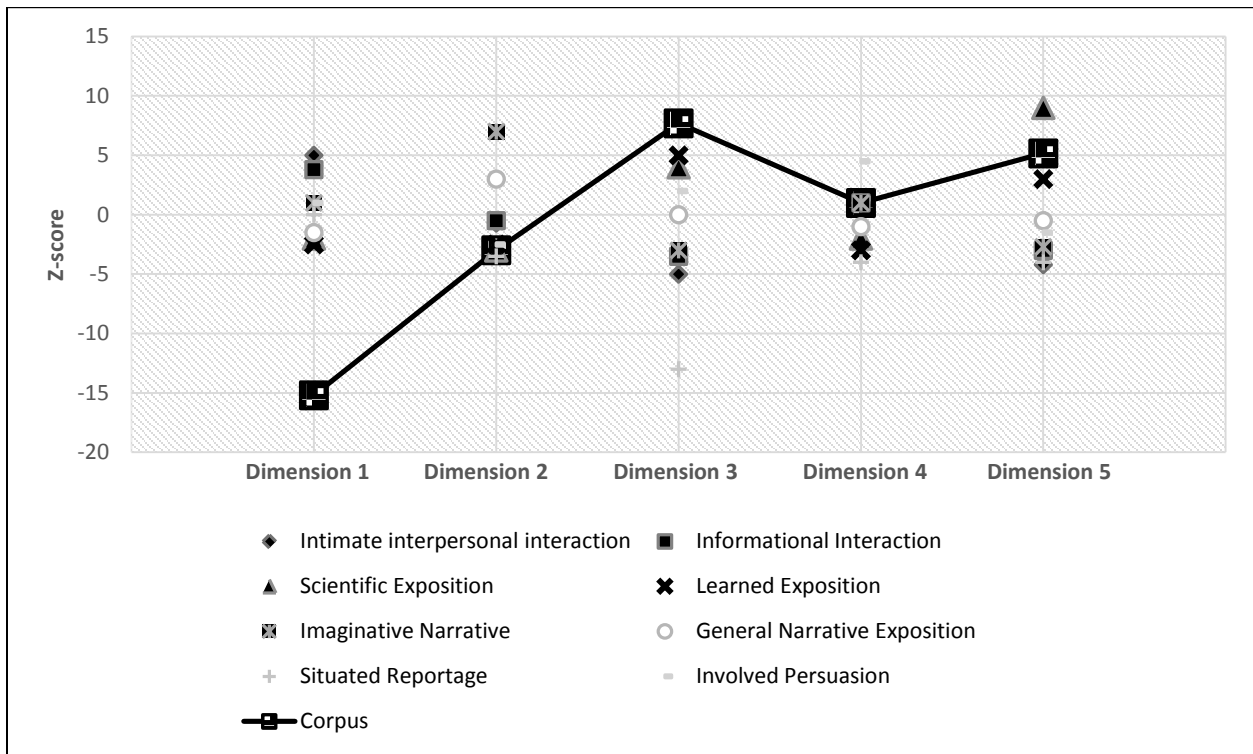


Figure 3:

