

Innovations in English for Academic Purposes (EAP) assessment

Friday 31 March 2017, 10:00 – 16:45
C126 (ground floor) Cottesbrook, Park Campus
University of Northampton, UK

Conference Programme

9:55 – 10:00	Welcome & Introduction
10:00 – 10:30	Invited talk: Are you looking for what I'm looking for? Evidence-led writing criteria, assessment literacy and the teaching of writing Daniel Waller (University of Central Lancashire)
10:30 – 11:00	Encouraging better preparation: a new Test of Academic Literacy for entry onto postgraduate EMI course Jocelyn Wyburd & Karen Ottewell (University of Cambridge)
11:00– 11:20	Coffee Break / Group Photo
11:20 – 11:50	Long-term strategies? Test-taker reading behaviour in a university foundation programme Richard Spiby (British Council)
11:50 – 12:20	An examination of British students' knowledge of polysemous academic English vocabulary Sophia Skoufaki (University of Essex) & Bojana Petric (University of London)
12:20 – 13:45	Lunch / Poster Session
13:45 – 14:15	Invited talk: Measuring the Effectiveness of EAP: What's in our Toolkit? Diane Schmitt (Nottingham Trent University)
14:15 – 14:45	Alternative ways of assessing listening and speaking skills on an in-session EAP course Elizabeth Allen (University of Bristol)
14:45 – 15:15	The evolution of EAP assessment tasks towards critical application of learning and sustainable feedback Maxine Gillway (University of Bristol)
15:15 - 15:35	Coffee Break
15:35 – 16:20	Panel Discussion on EAP Practice Qian Zhang (University of Northampton) Susie Cowley-Haselden (University of Northampton) Chris Smith (University of Sheffield) Diane Schmidt (Nottingham Trent University) Daniel Waller (University of Central Lancashire)
16:20 – 16:50	Closing, then TEA SIG AGM

Paper Abstracts

10:00 – 10:30 **Invited talk**

Are you looking for what I'm looking for? Evidence-led writing criteria, assessment literacy and the teaching of writing

Daniel Waller (University of Central Lancashire)

The benefits of analytic instruments over holistic ones for writing assessment have been extensively discussed in the literature, and the main points highlighted include a better reliability (Weigle, 2002) as well as improved feedback to test takers and teachers (Knoch, 2011). Experience also suggests that when using holistic instruments raters might come to similar scores for rather different reasons, therefore even an apparently satisfactory inter-rater reliability might in fact mask different opinions and weighting of the different components within the rating pool.

Such considerations led the EFL testing team at UCLan to implement a research project to produce new criteria for writing assessment, both grounded on theory and empirically developed so that the resulting scale might match as closely as possible the features of writing found in essays written by real students (North, 2003; Fulcher, 1987, 1996; Knoch, 2009, 2011). However, in addition to developing robust criteria, the project presents the opportunity to develop criteria which maximise usability and accessibility for all users as well as enhancing positive washback. This presentation will explore how we are attempting to develop academic literacy by actively highlighting to teachers the connections between models of writing, the marking criteria and assessment practices (such as holistic v. analytic marking) in order for them to better teach good writing practices and achieve our goal of positive washback. An example of this is recent research into the use of metadiscourse markers in writing by learners at different levels of proficiency; research which can be incorporated into the criteria and communicated to teachers.

10:30 – 11:00

Encouraging better preparation: a new Test of Academic Literacy for entry onto postgraduate EMI course

Jocelyn Wyburd & Karen Ottewell (University of Cambridge)

The influence of standardised English language tests has never been greater due to increased internationalisation and the ever-growing number of EMI courses on offer around the world. Yet while such tests are clearly important in confirming whether students have reached a minimal language competence, as Wingate has recently noted, such tests that are largely 'focused on language proficiency, such as grammar, structure and the use of cohesive devices [...] have been shown to have little predictive value of students' ability to use language in an academic context.' So what are the other linguistic skills and competencies beyond language proficiency that need to be gauged, especially for PG entry?

The Admissions Testing Service, part of Cambridge Assessment, together with the University of Cambridge Language Centre are currently working on a possible solution to this sector-wide issue – namely, with the development of a Test of Academic Literacy. In this presentation, I will provide an overview into the nature of the test itself – how the test will be structured, what will be assessed, how it will be assessed etc. – as well as the underpinning pedagogy behind it: namely, that what we are hoping to create with this new assessment is a far more positive washback effect where students are better aware of and prepared for the challenges of doing postgraduate study through the medium of English.

11:20 – 11:50

Long-term strategies? Test-taker reading behaviour in a university foundation programme
Richard Spiby (British Council)

In order to cover the large amounts of core course materials they face at university, L2 students are encouraged to develop the ability to read flexibly and strategically. This entails the training and effective assessment of students in a variety of reading styles. Accordingly, two subtests, of careful and expeditious reading, were developed with reference to the sociocognitive framework of Khalifa and Weir (2009) in an attempt to reflect reading demands with university texts more closely. The behaviour of 88 university students and 3 teachers during the completion of the two reading subtests was investigated. Both quantitative and qualitative data were collected through questionnaires and interviews in order to ascertain test-taker behaviour and strategy use on the two tests. The results of the study indicate that there were significant differences between several strategies used on the tests. However, there was also considerable overlap in some areas, particularly in terms of the incorporation of expeditious strategies into careful reading. Qualitative data also revealed other issues of interest related to metacognition and difficulties faced during the tests by test takers. The implications are discussed with reference to wider theory and practice in the teaching and testing of reading at university level in terms of the compensatory role of strategies and student awareness of suitability of strategies according to purpose. Recommendations are also made for further research into students' reading speed and the extent to which actual strategy use is effective on university courses.

11:50 – 12:20

An examination of British students' knowledge of polysemous academic English vocabulary
Sophia Skoufaki (University of Essex) & Bojana Petric (University of London)

Despite evidence that many academic words are polysemous (e.g., Cobb 2010; Hyland & Tse 2007; Durrant 2016), university students' knowledge of the meaning senses of polysemous academic words has not been examined. Yet this examination would help estimate students' learning needs and would, hopefully, promote and inform the to-date rare explicit academic vocabulary instruction in pre-sessional and in-sessional EAP courses. This presentation will report on a study testing 48 British English native speaker university students' knowledge of English academic polysemous words. The tests examined the knowledge of the meaning senses of 20 words selected from the first 1000 lemmas of Gardner and Davies's Academic Vocabulary List (2014). First a test of productive knowledge with a gap-fill task, which tested all the most frequent senses of each word were, was administered. Then a test of receptive knowledge with a multiple-choice task, which tested only the most frequent sense of each word, was administered. Findings suggest that students can achieve scores up to around 90% in the receptive test whereas they can at best achieve a score of around 70% in the productive test. The effect of factors such as the students' year of study and broad academic discipline on English academic vocabulary test scores will be discussed.

References

- Cobb, T. (2010). Learning about language and learners from computer programs. *Reading in a Foreign Language* 22(1), 181–200.
- Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes* 43, 49–61.
- Gardner, D. and Davies, M. (2014). A new Academic Vocabulary List. *Applied Linguistics* 35(3), 305–327.
- Hyland, K. and Tse, P. (2007). Is there an 'academic vocabulary'? *TESOL Quarterly* 41(2), 235–

253.

13:45 – 14:15 **Invited talk****Measuring the Effectiveness of EAP: What's in our Toolkit?****Diane Schmitt** (Nottingham Trent University)

As the number of students studying at university through the medium of English grows, so too does the need to provide English for academic purposes tuition to ensure that students have the necessary level of English to benefit from instruction and achieve their full potential. In the UK and elsewhere, universities expend considerable resource to provide pre- and in-session courses, workshops and other types of support, yet there is little data available to demonstrate how much or in what ways students' English language proficiency (ELP) develops as a result of this provision. Language gain has traditionally been measured using test-retest methods (e.g. Elder and O'Loughlin, 2003; Green, 2005; Birrell, 2006; Humphreys et al, 2012) and results tend to show little gain. This appears to be true whether students are in pre-university EAP programmes or in what might be considered more "language-rich" in-session environments where students are immersed in English-medium university courses. Are we to believe then that the language proficiency of L2 university students does not improve?

Most language gain studies make use of standardized proficiency tests such as IELTS, TOEFL or TOEIC (see above and Ross, 1998) as the measurement tool. These tests are tools that aim to assess a wide range of learners from a wide range of contexts and by their nature need to capture proficiency in a very general way. The law of the instrument, commonly attributed to Maslow, refers to the practice of overly relying on a familiar tool to try to solve any problem. In this talk, I will argue that EAP practitioners and language testers rely too heavily on too small a range of testing tools and that we need to work together to reconceptualise language development in the EAP context and create more appropriate tools for measuring language gain.

14:15 – 14:45

Alternative ways of assessing listening and speaking skills on an in-session EAP course**Elizabeth Allen** (University of Bristol)

It is difficult to create valid assessments for EAP listening and speaking skills, such as individual student's seminar skills and the ability to take effective notes during university lectures. However, this presentation will show some of the ways that these difficulties have been addressed on an undergraduate, in-session, academic listening and speaking skills course. The course aims to improve students' ability to listen critically to academic lectures and take effective notes; understand the structure and development of academic talks; provide oral summaries of lectures and seminar contributions; build on the contribution of others in seminars; justify personal contributions with appropriate evidence; and prepare and deliver short academic presentations. In line with these aims, three different assessment tasks are given. The most innovative one is the task that assesses seminar skills; instead of trying to holistically assess a group seminar, the subskill of being able to build on the contribution of others is evaluated by having individual students record themselves responding to a speaker and then submit the recording. To assess presentation skills, students submit a recording of themselves giving a three-minute presentation. To assess critical listening and note-taking, students sit an exam in which they listen twice to a recorded talk, take notes and then answer questions based on these notes and write a critical response to the talk.

Overall these assessment methods have been created to reflect not only the learning outcomes of the course, but also the broader academic context to help students develop relevant,

transferable skills.

14:45 – 15:15

The evolution of EAP assessment tasks towards critical application of learning and sustainable feedback

Maxine Gillway (University of Bristol)

The classic pre-post-graduate EAP assessment task is a discipline-specific piece of extended writing of around 2500 words demonstrating the ability to use sources effectively and ethically, which is redrafted after feedback at various points (e.g. McCormack & Slaght, 2012). This is often linked to an oral presentation of the written work in some way (e.g. Salter-Dvorak, 2016).

As a result of research carried out among pre-sessional students and teachers in the summers of 2014 and 2015, our assessment has evolved from such a task, which asks students to incorporate feedback into a second draft, to one that asks students to transfer feedback from one task to the next with a view to encouraging ‘sustainable feedback’ (Carless et al, 2011). This has resulted in an innovative task format for both written course work and oral presentation, integrating teaching and assessment and helping to design out plagiarism. We have moved from the replication of one social genre to the critical application of various cognitive genres (Bruce, 2008).

After a quick overview of the whole pre-sessional assessment system, I will focus on the new tasks and present data from student and teacher evaluations. The audience will be invited to critique the proposed 2017 version of the tasks and criteria in the spirit of enabling continuous improvement.

Poster Abstracts (in alphabetical order of surnames)

Ebtesam Abdulhaleem (University of Warwick): *Identifying misalignments in language proficiency using CEFER scales in a context outside of Europe*

The Preparatory Year Programme (PYP) was established for students enrolled in universities who deliver instructions in English. The aim of the programme is to prepare matriculated students to meet the English language demands required of first year students. However, the students’ proficiency in English was still not up to the standard required, despite the intervention of the PYP. Therefore, this study aims to identify the misalignment between the student English language proficiency towards the end of the PYP with the actual language level required. In a longitudinal study, ten CEFER scales were used to identify the perceived levels reached towards the end of PYP. One year later, the same participants used the CEFER scales to identify the levels the students required in their first year at the university. Detailed analyses of each stage were considered. Then, the results were compared using paired sample t-test, which identified differences in means between the two stages. The aim was to identify any significance difference between the level reached and the level needed. It was found that there is a noticeable gap between the levels reached and needed. The gap is more obvious with lower proficiency levels compared to higher ones. To conclude, CEFER scales were proven to be a useful tool to indicate the students’ CEFER proficiency levels and to identify possible gaps in the curriculum, especially in areas where using the CEFER is new and there are no other means, such as examinations, of establishing the language levels of students.

Lizbeth Morales Berlanga (University of Southampton): *Assessing speaking: The importance of proving validity in assessment case study in an EAP course*

Assessing speaking is an important part of language teaching, however, is one of the most difficult skills to assess due to its nature of spontaneity. Speaking skills could not be evaluated based on a number of errors, therefore, evaluators should be aware of several considerations that usually are part of an assessment criteria (Luoma, 2004). Teachers' perspectives play also an important role in the assessment of the speaking skills; teachers should be aware of the objectives of the test and it is relevant as well, that they were all pursuing the same goals (Lazaraton, 1996; Lumley & O'Sullivan, 1995), otherwise, the validity of the test could be a subject of judgment (Bachman & Palmer, 1997).

The aim of this study is to delimit the validity of a speaking test in an EAP course to make sure that the test is measuring the skills and features that are intended to measure. Thirteen EAP students voluntarily participated in the study, to undertake a questionnaire and know their thoughts about the course, the test and the use of the scales, two EAP teachers were interviewed with the aim of knowing their own perspectives, at the end an analysis of the tasks and the final results of the test were made using quantitative methods. Overall the results showed that there are relevant differences in the final marks depending on the teachers' perspectives about what should be the main goal of the test, highlighting the importance of teachers' training and possible advantages of standardization.

Helen Grinsell & Clare Albans (University of Newcastle): *"Making it Real": innovation in Year 1 EAP speaking assessment*

Previous speaking assessment for Year 1 international students of Business involved consensus-reaching tasks. Students 'presented' monologues of opinion with little genuine interaction or construction of a developed argument. This highlighted the difficulties that international students of different linguistic ability can have with 'interaction' in a seminar setting.

The aim was to identify how assessment of speaking can happen to not disadvantage students with lower language ability and how such assessment can be made meaningful and add value to using English.

A new academic seminar task was developed and piloted to move towards truly assessing students' ability to fully participate in academic seminars. The students need to use subject content, demonstrate understanding of it and how it relates to other reading. They must relate to and build on the contributions of other students to collectively build an argument and construct their understanding of the seminar topic. Students' attention is also drawn to features of academic speaking, using language in situationally appropriate ways.

Findings over 2015/16 indicate more interaction is happening, more naturally in most cases, and student talk is more content responsible, resulting in students coming away with new insights into the content and more understanding of effective seminar behaviour.

This has necessitated further investigation into the suitability of current assessment criteria and the positive backwash effects on semester seminar skills practice to support students in extending their repertoire of language in an academic context and developing their seminar behaviour.

Chris Smith (University of Sheffield): *Listening-into-speaking seminars: design, implementation, problems and solutions*

Over the last 2 years at the ELTC, we have been redesigning our summer pre-sessional EAP course, moving away from a model using a high stakes end of course exam and in its place



developing bespoke assessments better integrated into the course. A new development for 2016 was a listening-into-speaking seminar for which students attended a live lecture and then participated in a student-led seminar the following day.

This presentation will describe the task in terms of time, groupings, input material and preparation. It will also outline the role of the teacher/assessor, assessment procedures and marking criteria for the seminar from an assessment perspective. The talk will also show how the task was integrated into the course, the practice and preparation, the teaching materials and the feedback processes that were established. End of course feedback on the seminars will also be discussed.

There are various problems inherent in assessing these seminars: the number of people makes assessing them more difficult; stronger students may dominate and weaker students struggle; if there are different roles, some may lend themselves to better performance (e.g. seminar leader); students may prepare to the point of memorisation; there is the danger that it can be too indirect so that students could pass without having attended the lecture; and in combatting that point there is the danger of asking overly specific questions. The solutions and responses to these problems will be discussed.

Qian Zhang (University of Northampton): *Propose a Framework for EAP Programme Evaluation*

This poster proposes a framework for evaluating EAP programme. During the last few years more and more EAP practitioners presented their research and good practice at the BAAL TEA Sig conferences, BALEAP PIMs and also at the recent LTF 2016. It clearly shows there has been an increasing interest in assessments in EAP covering from EAP programme redesign, assessment supporting learning, learning oriented assessment or innovative assessments. In addition, the form of the BALEAP Testing working party has shown the demands from EAP practitioners for information and advice for the development of valid and reliable EAP language testing practices. Although the EAP assessment and EAP programme are considered important, there is little mention of EAP programme evaluation to measure the effectiveness of the current programme in order to implement the programme and better fulfil different stakeholders' need. It hopes that this proposed framework will contribute to EAP programme coordinators and EAP practitioners, to measure whether their goals of EAP program have been met and to ensure they continuously improve their programme. Furthermore, it is expected this proposed framework will benefit the university management and other professional services (admission and international office), which are considered as one of the shareholders.