Assessment supporting learning

Friday 8 May 2015, 9:50 – 17:00

Pitt Building Conference Centre, Cambridge, UK

Conference Programme

9:15 – 9:45	Registration opens
9:50 – 10:00	Welcome & Introduction
10:00-10:30	Invited talk: 'What Do Teachers Notice in Classroom Assessment Scenarios?' - Christine Harrison (King's College London)
10:30-11:00	'How classroom tests are shaped by a high-stakes examination: teachers' voices' - Doris Froetscher (Lancaster University)
11:00-11:30	'Washback of a high stakes exam: The impact on teachers' - Gemma Bellhouse (University of Oxford)
11:30-12:00	Invited talk: 'Learning Oriented Assessment: a theory in search of a pedagogy' - Neil Jones (Cambridge Language Sciences) and Sarah Unsworth (Cambridge Assessment)
12:00-12:20	Poster introduction (Presenters have 2 minutes each to briefly highlight theme, point of interest etc.)
	Poster presentations: Poster session / Lunch
12:20-13:00	'Assessment of Communication Skills in Pathway to Medicine: methods and tools' - Dawn Cremonese and Georgina Lloyd (University of St Andrews)
	'Grammar and the CEFR: towards the validation of the British Council – EAQUALS Core Inventory' - Glyn Jones (Lancaster University)
	'Working in shades of grey: error analysis in spoken learner language' - Paula Buttery, Andrew Caines, Calbert Graham, Michael McCarthy (ALTA Institute, University of Cambridge)
13:00-13:30	Extra time for Lunch
13:30-14:00	'How useful are diagnostic vocabulary tests?' - Benjamin Kremmel (University of Innsbruck, Austria & University of Nottingham, UK) and Laura Vilkaitė (University of Nottingham, UK)
14:00-14:30	'Assessment in EAP Language Pathways 'Do students really get it?'' - Sal Consoli (UWE Bristol's International College)
14:30-15:00	'Assessment - a tool for learning in a Pre-sessional setting' - Qian Zhang, Bob Fisher and Cain Barriskill (University of Northampton)
15:00-15:15	Coffee break with refreshments
15:15-15:45	Invited talk: 'Assessing English as an Additional Language in Linguistically Diverse Schools: Challenges to Theory and Practice' - Constant Leung (King's College London)
15:45-16:15	Panel discussion
16:15-16:25	Closing
16:25-17:00	BAAL TEA SIG AGM

Abstracts

What Do in Teachers Notice Classroom Assessment Scenarios?
Christine Harrison (King's College London)

Assessment for learning (AfL) depends on feedback loops and many of the strategies that teachers incorporate into their pedagogy, under the umbrella of AfL, are designed to enhance these mechanisms. The question is what triggers a response from the teacher. Is it the size of the response in that most students can or can't answer, the frequency of mistakes, the fact that a specific child could or couldn't do a task or simply that the teacher receives a response they did or did not expect? Two of my projects support teachers in strengthening their practice in inquiry-based learning scenarios and this has been influential in developing my thinking about what teachers notice during the enactment of their pedagogy. My current research focuses on how teachers enact formative assessment and this involves exploring how they plan and prepare for classroom opportunities in which they can assess and respond to learning as it is taking place. The research also considers how teachers conceptualise their classroom assessment practices within a sociocultural paradigm.

Learning Oriented Assessment: a theory in search of a pedagogy
Neil Jones (Cambridge Language Sciences)
and Sarah Unsworth (Cambridge Assessment)

Work at Cambridge is converging upon Learning Oriented Assessment from two directions: an approach from theory, which will be published this year in the Studies in Language Testing series, and an approach to practical implementation, in the form of a multi-level course developed in collaboration with Cambridge University Press. We will touch on three questions: What is learning - a question about individual cognition; what is to be learned - a question about how formal learning is structured; and most importantly: what is the role of assessment in learning - the key question for linking classroom interaction and more formal use of tests. This talk will briefly present both theory and practice, and invite participants to consider the conditions for their successful combination.

Assessing English as an Additional Language in Linguistically Diverse Schools: Challenges to Theory and Practice
Constant Leung
(King's College London)

Since the onset of the National Curriculum in 1991 the assessment criteria and framework for the subject English have been applied to EAL. In the first part of my talk I will draw attention to some of the incommensurabilities between assessing English (as a curriculum subject) and assessing EAL (as language proficiency) in terms of construct validity and educational purposes. The focus of the second part will be on the assessment implications of the complex curriculum-embedded nature of EAL development, and explore ideas for alternative approaches that can take account diverse pupil trajectories and curriculum contexts. I will draw on relevant examples of EAL assessment frameworks developed in other English-speaking countries to support the discussion.

How classroom tests are shaped by a high-stakes examination: teachers' voices.

Doris Froetscher (Lancaster University)

Whenever an exam is reformed, changes in the classroom are likely to occur. This presentation reports on a study investigating the effect of a standardized national school-leaving exam introduced in 2008. It follows several scholars' call for research into how external tests affect assessment in the classroom (Wall and Alderson, 1993; Wall and Horak, 2006; Watanabe, 2000), and indeed addresses a research gap in our field. The study specifically focuses on the washback of the exam's reading paper on the assessment of reading in classroom-based tests.

In an earlier stage of this larger research project, 173 class test tasks from 22 teachers were analyzed using a specially designed instrument. The comparison of tasks used pre and post the introduction of the new exam established the presence of washback. Results show, for example, that after the introduction of the exam, test methods in class tests had aligned to it. Results also show an increase in task quality characteristics, e.g. instructions, example items, or distracters. In this presentation, I will report on a follow-up study that investigates teachers' individual approaches and strategies when constructing or selecting reading class test tasks. To this end, semi-structured interviews were conducted with carefully selected teachers from the original sample. The teachers' tasks already analyzed in the previous phase were used to enrich the interviews with a stimulated-recall element. The analyses of the audio-recorded interviews shed light on the reasons behind and mechanisms of washback in the study's setting.

Washback of a high stakes exam: The impact on teachers

Gemma Bellhouse (University of Oxford)

Although public examinations are used to keep teaching in check (Cheng, 1999), are teachers sufficiently equipped to prepare students for these examinations? A recent washback study on the high stakes French Baccalauréat demonstrates a lack of direct communication with the stakeholders and a deficiency in assessment training (Bellhouse, in press). Instruments of questionnaires, semi-structured interviews, and follow up questions with 8 teachers in 3 secondary schools revealed that while many teachers and students relied on test preparation activities like drilling and memorisation, a minority of teachers relied on a more interactive approach in an attempt to reflect reality. This is evidence that teachers, like their students, view high stakes tests as an 'outside' control and feel pressured to manipulate 'learning' to increase scores. 75% of teachers (n=6) stated that their schools and the French Ministry of Education could have further aided them. A teacher stated: 'They gave us nothing, no help. They said to do it, just do it.' But what do teachers need, besides brief instructions and a rating scale aligned to the CEFR? Using the French washback study to emphasise needs of language teachers in secondary schools, this presentation will further the call to action for training schemes and feedback systems to aid both teachers and students.

Assessment of Communication Skills in Pathway to Medicine: methods and tools.

Dawn Cremonese and Georgina Lloyd (University of St Andrews)

The publication of Tomorrow's Doctors in 2003 has increased the focus on teaching and learning communication skills in UK Medical schools (Brennan et al, 2010). Despite a continued rise of non-native speaker medical students in the UK, existing frameworks outlining the principles of effective communication do not consider culture or mother tongue (Woodward-Kron, Stevens and Flynn, 2011; Eseonu, Wedderburn and Maurice, 2011). Hamilton (2009) also states that although the Objective Structured Clinical Exam (OSCE) is a reliable assessment tool used in Medical Schools worldwide, it requires

intercultural competence (IC); however as yet there are no standard guidelines on how or when to teach or assess IC to non-native speaker students. These issues are clearly relevant to the ESAP classroom, particularly in pathway to Medicine programmes, yet there is limited research in this area (Hughes, 2011). This presentation focuses on two studies, which investigate the assessment of communication skills on a Pathway to Medicine course, with a focus on IC. The studies were carried out at English Language Teaching, University of St Andrews between 2012 and 2015. The first focuses on the assessment of prosody using Audacity (2013) software, and examines its role in IC in the doctor-patient interview. The second explores how video feedback on students' assessed role-plays using agenda-led outcomes-based analysis (ALOBA) facilitation (Kurtz, Silverman and Draper, 2005) provides an effective feedback tool as well as a platform where IC can be addressed, thus raising students' awareness and promoting changes in language use and communication strategies.

Grammar and the CEFR: towards the validation of the British Council – EAQUALS Core Inventory

Glyn Jones (Lancaster University)

The aim of the British Council – EAQUALS Core Inventory for General English is "to make the Common European Framework of Reference for Languages [CEFR] accessible to teachers and adult learners" by addressing the question as to what the CEFR levels mean in terms of classroom aims. At the core of the document is a listing of learning objectives for English for each of the CEFR levels from A1 to C1, together with illustrative exponents. The inventory was derived by synthesising data from a number of sources, including the content of popular course book series, teacher surveys and the syllabuses of public examinations that claim to be linked to the CEFR. The aim of the current study was to investigate the extent to which these findings are borne out by analysis of written language actually produced by learners of English at the respective levels. A corpus of learner writing was compiled from test taker responses to the writing component of APTIS. This was then analysed with a view to finding, and counting, exponents of the grammar objectives listed in the CIGE, and tagging these by CEFR level of the respective writers as reported by APTIS. In order to provide an independent check on the test results, a sub-set of responses was presented to a panel of experienced EFL teachers, who were asked to assign a CEFR level to each of the test taker responses.

Working in shades of grey: error analysis in spoken learner language

Paula Buttery, Andrew Caines, Calbert Graham, Michael McCarthy

(ALTA Institute, University of Cambridge)

Our project is situated in the field of computer-assisted language learning (CALL), working towards a system for the automated assessment of learner speech. To support learning, we intend to provide automated feedback to system users. But in order to do so, we need a fuller understanding of `error' and how it relates to spoken (rather than written) language. We view error as a gradient, context-sensitive feature of language, and we present a multi-dimensional analysis framework based on this view. The idea is that the framework may be applied to all constructions, with each dimension being scalar, thereby in combination placing each construction on a continuum from what has traditionally been termed 'grammatical' to 'ungrammatical'. It is clear that there are prototypical examples of each class of grammaticality: for example, in English we presume that, 'I am a learner', would unfailingly be classed grammatical, whereas the agreement error in, 'I is a learner', would always mark out this construction as ungrammatical. However, there are constructions, such as, 'I'm gonna leave', which might be classed as an error in written language but not in speech. Moreover, 'whatcha doing', might be deemed more acceptable in more informal than more formal

registers of speech, illustrating the context-sensitive nature of error.We have designed a multi-dimensional framework to capture these subtle distinctions, and we firstly explain what the dimensions of the framework are, secondly report on our experiments to operationalise the dimensions, before finally considering the implications of a dynamic, gradient notion of error for learner assessment and feedback.

How useful are diagnostic vocabulary tests?
Benjamin Kremmel (University of Innsbruck, Austria & University of Nottingham, UK)
and Laura Vilkaitė (University of Nottingham, UK)

Vocabulary tests are a key tool in the assessment repertoire of classroom teachers. They hold great potential for integrating learning and assessment as they could be used for diagnosis and tailoring instruction to the learners' needs. Several tests have been developed and suggested for diagnostic and pedagogic use, such as the Vocabulary Levels Test (VLT) (Nation, 1990; Schmitt et al., 2001) or the Vocabulary Size Test (Nation & Beglar, 2007). Most of these have been commonly labelled measures of vocabulary size, suggesting that this is a well-defined construct, useful for pedagogic practitioners in their daily work with language learners. However, the interpretation of scores generated by tests of vocabulary size is not as straightforward as it seems. The paper will report on two studies that have investigated the score interpretation of item formats used in these vocabulary measures. Study 1 examined the interpretability of VST scores and found a considerable mismatch between test scores and the word knowledge learners demonstrated in face-to-face interviews. Study 2 compared the VLT and the VST format against a criterion of meaning recall and showed that both these recognition formats heavily overestimated learners' word knowledge. The paper will thus argue that vocabulary size figures need to be interpreted in conjunction with the degree of knowledge tapped into in order for score information to be useful for teachers and learners in pedagogic settings.

Assessment in EAP Language Pathways 'Do students really get it?'
Case Study: UWE Bristol's International College
Sal Consoli (UWE Bristol's International College)

Feedback on L2 writing skills has long been viewed as a powerful learning tool (Hyland & Hyland, 2006). Nonetheless, providing and receiving feedback is not unproblematic, as each educational context is shaped by a number of idiosyncratic factors such as learners' backgrounds, tutors' beliefs, external bodies' power (Higgins, 2006). We should, therefore, view feedback as a contextual and social system if we wish to investigate its nature and processes within the realities of its specific context(s). This study aims to investigate students' perceptions of and attitudes to feedback on their written EAP assignments at UWE Bristol's International College. The findings should establish: a) whether students feel an active interest in feedback; b) to what extent they understand the institutional discourses underpinning feedback; c) what types of commentaries students find most useful; d) what type of feedback they like to receive.The research instruments include a large-scale questionnaire and a series of semi-structured interviews with students which provide a richer and deeper understanding of the overall results. The focus is on the very first academic essay students need to produce to fulfil the requirements of a compulsory module at Foundation or Diploma level. While we believe in the quality of our teaching, with this study, we hope to obtain useful data which may inform the design of new feedback strategies to ensure students fully understand action plans for improvement and, at the same time, feel motivated to do so.

Assessment - a tool for learning in a Pre-sessional setting
Qian Zhang, Bob Fisher and Cain Barriskill
(University of Northampton)

UK higher education is facing lots of challenges and changes. In the language education sector, there is an increasing demand for changing the attitude towards assessment (Green, 2014). It is important to ensure that 'effective assessment practice should focus on enhancing student learning processes, but needs to be informed by the awareness that assessments do double duty' (Carless, 2014, p2). To respond to changing and the structural incoherence of the programme, together with poor learning and assessment outcomes, a fundamental revision of the Pre-sessional Programme was carried out in summer 2014 at the University of Northampton.The objectives of this ongoing study are to evaluate the effectiveness current assessments formats and explore teachers' perspectives about the potential use of 'learning-oriented model'(Carless, 2007) and 'learning-oriented assessment cycle' (Cambridge English, 2014) to strength teaching, learning and assessment by reflecting their teaching practice.

Research questions:
1. To what extent do pre-sessional teachers support the use of learning-oriented assessment in the department? Why or why not do they support it?
2. What are the challenges they have faced?
3. How can learning-oriented assessment be implemented?

Methods
Data will be collected by using semi-structure interview with seven pre-sessional teachers at the University of Northampton and analysed using themes that emerged from the research questions. Our tentative findings are that 1) teachers support the use of LOA 2) there are potential strategies and suggestions for implementation 3) continuous professional development on LOA are essential.