Exploring the relationship between language assessment and second language acquisition (SLA) research

Friday, 29 March 2019, 9:45 – 16:45

University of London, UK

Conference Programme

9:45	Registration
9:55 – 10:00	Welcome & Introduction
10:00 – 10:40	**Invited talk:** Fluency in language assessment: What testers may learn from research in different disciplines Nivja De Jong (Leiden University)
10:40 – 11:10	Bridging the gap: The effectiveness of intensive IELTS writing preparation in China and Japan and 'relearning' academic conventions Tony Clark (Cambridge Assessment English)
11:10– 11:30	Coffee Break / Group Photo
11:30 – 12:00	Assessing Fluency construct from a second language acquisition perspective: The case of TEEP speaking tests Parvaneh Tavakoli, John Slaght, Gill Kendon (University of Reading) Ann-Marie Hunter (Leeds University)
12:00 – 12:30	Exploring the order of acquisition of lexical collocations among Chinese learners of English: A cross-sectional study Jing Xu (Cambridge Assessment English)
12:30 – 13:30	Lunch Poster Discussion
13:30 – 14:10	**Invited talk:** Interactional competence: At the crossroads of SLA and Language Testing Evelina Galaczi (Cambridge Assessment English) Lynda Taylor (University of Bedfordshire)
14:10 – 14:40	Exploring the relationship between language learning motivation and proficiency for 15-year-old Spanish learners of English Karen Dunn (British Council) Janina Iwaniec (University of Bath)
14:40 – 15:10	Measuring cognitive fluency: The underlying processes involved in second language production Tannistha Dasgupta (VIT-AP)

15:10 - 15:30	Coffee Break
15:30– 16:15	**Panel Discussion:** The relationship between language assessment and second language acquisition (SLA) research (in alphabetical order of surnames) Nivja De Jong (Leiden University) Evelina Galaczi (Cambridge Assessment English) Ana Pellicer-Sanchez (University College London) Lynda Taylor (University of Bedfordshire) Parvaneh Tavakoli (University of Reading)
16:15 – 16:45	Closing, then TEA SIG AGM

Paper Abstracts

10:00 – 10:40 Invited talk
Fluency in language assessment: What testers may learn from research in different disciplines
Nivja De Jong (Leiden University)

This presentation critically evaluates the current conceptualization of fluency in language testing, namely that it is a concept that should be sought in the ear of the beholder. Most of the research on fluency that is usually cited and used as empirical basis for the current conceptualization comes from the field of applied linguistics. This presentation seeks evidence from additional research disciplines: not only from applied linguistics, but also psycholinguistics, discourse analysis, and sociolinguistics. To summarize the insights from these research fields, it can be concluded that (1) fluency behavior is in part dependent on personal speaking style, (2) disfluencies are not only signals of trouble in formulating but are also helpful signals for the listener, and (3) disfluencies are part of communicative effective speech. I will discuss the implications of the insights from different disciplines for language testing practices.

10:40 – 11:00
Bridging the gap: The effectiveness of intensive IELTS writing preparation in China and Japan and 'relearning' academic conventions
Tony Clark (Cambridge Assessment English)

Existing research has investigated the IELTS exam preparation process, and monitored the progress of previously successful candidates on a UK course of study. However, the important relationship between the two remains underexplored, as they have not been included - together - in a project thus far. This case study explores IELTS exam preparation in candidates' home countries. Specifically, it explores the challenges Chinese and Japanese students face when learning to write IELTS essays in English for the purpose of university entrance.

There are two sections in the study, which emerged from extensive experience of teaching intensive exam preparation classes in Asia. The first part investigates the IELTS exam preparation process for learners at two language centers – one in China and one in Japan. Conducted over a nine-month period, classroom observations included an IELTS preparation course in its entirety at each data site. Furthermore, students and course instructors were interviewed about the IELTS preparation experience. Finally, written homework essays were also collected, to explore the development of students' work. Although this is not a comparative study, it documents the variations in approach between the institutions in each country. Thirdly, the research focused on Chinese and Japanese students studying for a postgraduate degree in the UK. The challenges these candidates faced, and how these were overcome, were explored in depth. The overall results demonstrate the complexity of test preparation for international study, and highlight key recommendations to improve the process for learners from early to later stages.

11:30 – 12:00
Assessing Fluency construct from a second language acquisition perspective: The case of TEEP speaking tests
Parvaneh Tavakoli, John Slaght, Gill Kendon (University of Reading)
Ann-Marie Hunter (Leeds University)

Fluency is a key construct in the assessment of second language oral proficiency and an important indicator of the degree of automaticity in the process of second language acquisition (SLA). From a SLA perspective, different aspects of fluency, i.e., speed, breakdown and repair reflect important stages of L2 processing and development (Skehan, 2009). From a language testing perspective, fluency is a key construct that highlights the candidate's communicative ability. Despite this important connection between the two fields, investigating fluency from both perspectives is an under researched area (de Jong, 2018; Tavakoli, Nakatsuhara & Hunter, 2017). The current study, aimed at making a connection between the two, examines the relationship between assessment of fluency in the Speaking paper of Test of English for Educational Purposes (TEEP) and a linguistic analysis of fluency by using detailed analytic measures of speed, breakdown and repair. Investigating fluency construct across four assessed levels of proficiency (A2-C1), The data comprised 60 candidates' task performances, 15 at each level of proficiency. Using PRAAT software, the data were analysed for a range of speed, breakdown and repair measures. A repeated measures ANOVA was used to investigate the differences in the candidates' performances across the four levels of proficiency. The results confirm the findings of Tavakoli et al., (2017) suggesting that although speed measures distinguish between different levels, B2 and C1 level are not different in terms of speed. End-clause pausing distinguishes lower from higher levels of proficiency, whereas mid-clause pausing behaviour is a useful aspect of fluency to distinguish C1. No statistically significant differences were observed for repair measures across the four levels. These findings have significant implications for both fields. The results imply that speaking test rating scales and rater training should be reconsidered to reflect representations of fluency across different levels of proficiency.

12:00 – 12:30
Exploring the order of acquisition of lexical collocations among Chinese learners of English: A cross-sectional study
Jing Xu (Cambridge Assessment English)

Collocation, known as words that commonly co-occur, is a prominent category of formulaic language. There has been growing consensus among second language acquisition (SLA) and language testing researchers that collocation use is essential to effective verbal communication and a strong indicator of L2 proficiency and pragmatic competence (Schmitt, 2010; Voss, 2012; Wray, 2002). A lexical collocation concerns the co-occurrence of two or more content words that contribute almost equally to its whole meaning. Misuse of lexical collocations or lexical deviation, according to many linguists, is more likely to obscure meaning than structural deviation (Barnbrook, 2007; Bolinger, 1976; Sinclair, 1991). Psycholinguistic research has also suggested that lexical deviation in L2 speech may cause more processing difficulty than structural deviation for native-speaker listeners (Millar, 2011).
The present study analyzed 2344 lexical collocations extracted from spontaneous L2 English speech produced by sixty Chinese learners of English at four proficiency levels. The speech was collected in two distinctive contexts of language use: 1) conversing with an interlocutor on daily topics and 2) giving a mini-lecture on an academic topic. The extracted collocations were double-coded for semantic accuracy, grammatical accuracy, transparency, restrictedness, and fluency. It was found that the learners used verb-noun (44%), adjective-noun (25%), and noun- noun collocations most frequently. Their overall collocational performance tended to be stronger in the mini-lecture than in the conversation. Most importantly, a clear pattern of improvement in collocational performance was observed across the four proficiency levels. The implications of the findings for language assessment and learning will be discussed.

13:30 – 14:10 Invited talk
Interactional competence: At the crossroads of SLA and Language Testing
Evelina Galaczi (Cambridge Assessment English)
Lynda Taylor (University of Bedfordshire)

Interactional competence in second/foreign (L2) language education has assumed greater importance in recent decades, resulting in communicative approaches to developing and assessing speaking skills. Social interaction has been positioned as 'the matrix in which acquisition takes place' (Ellis 2005:219) and a prerequisite of language learning both in socio-cultural and interactionist theories (Lantolf 2000, Long 1996). That widely accepted premise presents opportunities as well as challenges from a testing point of view, since tests based on an interactional construct - which is essential from an SLA point of view - need to balance the tension between a broad interactional construct and test variability. In this presentation we will briefly overview key historical influences which have shaped the evolution of the interactional competence construct and the theoretical dilemma of defining it as a psycholinguistic construct residing within an individual versus a socio-cognitive construct residing with a social context. We will present an example of a task from a Cambridge Assessment English speaking test, and will use it as a case study for exploring a number of challenges associated with tests tapping into an interactional construct, such as the tension between conceptualising interactional competence as a validity asset vs. a validity threat, the trade-offs in embedding tasks in a rich communicative contexts, the effect of interlocutor variables, the role of non-verbal behaviours, and the challenges of standardisation and reliability.

14:10 – 14:40
Exploring the relationship between language learning motivation and proficiency for 15-year-old Spanish learners of English
Karen Dunn (British Council)
Janina Iwaniec (University of Bath)

Research on language learning motivation was initially spurred by finding that variance in language achievement could not be fully explained by language learning aptitude (Gardner & Lambert, 1972). Yet since then, collaboration between experts in the fields of assessment and motivation has not been pursued to its fullest extent, resulting in mixed evidence for the extent and nature of the relationship between L2 proficiency and motivation. This paper presents an innovative approach to exploring the relationship between language achievement and motivation using latent profile analysis.
1773 learners of English were sampled randomly from schools located in the Madrid region. These learners completed an eight-scale questionnaire, which operationalised constructs from influential theories of language learning motivation (Dornyei, 2005)), plus a multi-skill English language test (Aptis). To explore patterns of motivation and achievement amongst participants, a Latent Variable Mixture Modelling (LVMM) analysis was undertaken in Mplus7. This person-centred analysis divides participants into groups according to shared attitudes and performances (Muthén & Muthén, 2010) rather than assuming homogeneity (Marcoulides & Heck, 2013).
Results indicated five distinct groups amongst the Spanish students, each characterised by varying proficiency/motivational profiles. Amongst other observations, it was noted that two highly motivated, highly achieving groups were distinguished by internalised motives – their conception of their future and present L2 selves – whilst a discernible difference in levels of interest and effort separated high and low performing students with otherwise similar motivational profiles. Teasing out the relationships in this way goes some way to explaining previous mixed findings in this area.

14:40 – 15:10
Measuring cognitive fluency: The underlying processes involved in second language production
Tannistha Dasgupta (VIT-AP)

Despite the extensive use of standardized speaking tests and rubrics to measure fluency, issues with such measures have entailed latest research to suggest that more valid assessment may be derived from the measurement of the cognitive processes underlying L2 speech production. Levelt's (1989) 'blueprint' discusses the linguistic, psycholinguistic, and cognitive issues underlying the act of speaking. However, L2 speaking assessments have largely neglected the underlying mechanism involved in language production. Although a few studies have constructed measures to examine cognitive abilities in L2 speaking, further research is needed to establish the standardized measures. In order to determine the cognitive processes underlying L2 speech production- reflecting how efficiently a speaker can plan and produce speech- it is important to examine L2 cognitive fluency. This paper studies the measures that can be employed to assess the cognitive processes involved in language production. Thus, it relates cognitive fluency to utterance fluency. Data is collected from 39 adult L2 learners of English with intermediate level of proficiency in the language. Tasks measuring cognitive abilities are designed based on De Bot's L2 language production model. Automaticity and attention-based mechanism using tasks that tapped a) speed and efficiency of lexical processing and attention control are examined in this study. Data analysis shows that among the measures of utterance fluency, speed and length of pauses could be related to L2 cognitive fluency. Thus, assessment criteria can be designed to capture the best measures of the underlying cognitive processes in order to make testing of speaking more comprehensive.

Poster Abstracts

Measuring intelligibility, comprehensibility and pronunciation accuracy of L2 speech by Chinese adult learners of English
Yuanyue Hao (University of Oxford)

Assessment of second language (L2) pronunciation has been established as an integral component of L2 speaking assessment, usually combined with other rating criteria such as fluency, lexical resource and topic development to generate an overall score for the speaking section in major language tests. It wasn't until recently that researchers began to focus on assessment of pronunciation per se, which plays a critical role in the pedagogical context, such as pre-service teacher training and screening of international teaching assistants (Farnsworth, 2014). This study attempts to explore the relationship between intelligibility, comprehensibility and pronunciation accuracy of Chinese learners of English with the aim to provide suggestions for pronunciation instruction in China.

Participants in this study were first-year undergraduate students enrolled in a pre-service English teacher education programme in a university in eastern China. They were rated on a read-aloud task by two tutors using the rating scales of pronunciation accuracy and comprehensibility. The analytic scale of pronunciation accuracy was developed and validated by the researcher, while the scale of comprehensibility was adapted from the existing scale developed by Isaacs, Trofimovich and Foote (2018). Moreover, recorded speech samples were automatically transcribed by natural language processing (NLP) tools. The transcribed scripts were compared with the original text which they had been instructed to read aloud by text similarity analysis to measure the intelligibility of students' speech. Multivariate regression analysis was conducted to identify robust predictive pronunciation features that contributed to the intelligibility and comprehensibility of L2 utterance by Chinese students.

Investigating the influence of recast on the L2 speaking test performance
Pooneh Heydari (University of Bedfordshire)

Recast as a type of corrective feedback in the context of EFL has received much attention in the literature. However, there are still unanswered questions about if recast best facilitates L2 language acquisition in the context of Iran. Along the same lines, the effectiveness of recast in a speaking test in the context of Iran was investigated in this study. Two groups of learners were recruited as experimental and control groups in which the learners received recast and no recast during the treatment session, respectively. The pretest and posttest were run before and after the treatment session. In the treatment session of experimental group, the students' errors in using articles 'a' and 'the' were corrected using declarative full sentence recasts by the teacher. Consequently, the effect of recast on the performance of the students in a picture story speaking test was investigated. The data collected from 40 participants at intermediate proficiency level were analysed using independent sample t-test. When the differences of the pretest and posttest of the experimental and control groups were analysed; it was found that the former group significantly outperformed the latter one. The findings of the study suggested that recasts help learners overcome their errors and consequently improve their performance in the second language. Such findings could help improve our knowledge of corrective feedback in Speaking assessment context.

Sponsors: