

# ALTE 6th



## INTERNATIONAL CONFERENCE

BOLOGNA, ITALY | 2017

### Learning and Assessment: Making the Connections

3–5 May 2017



In collaboration with



and supported by



**CAMBRIDGE ENGLISH**  
Language Assessment  
Part of the University of Cambridge

# Book of Abstracts

---

**The presentations appear by format,  
by alphabetical order according to the first letter of the title,  
and colour-coded according to the strand.**

<b>Plenary presentations .....</b>	<b>1</b>
Connecting policy and practice at European level .....	1
Construct and content in context: Implications for language learning, teaching and assessment in China .....	1
Language policy and social cohesion: What links between social environment and regimes of learning and assessment? .....	2
Language testing washback and impact in our globalized world .....	3
Making the connections: digital innovation and diagnostic feedback .....	3
<b>Panels.....</b>	<b>5</b>
Insights from research on sign language tests .....	5
Lessons learnt from Italy in language learning, teaching and assessment .....	9
The challenges of a learning oriented and multilingual school assessment policy .....	11
The development of China's Standards of English and its potential application .....	15
<b>Workshops .....</b>	<b>19</b>
Erasmus+: Funding opportunities for language education.....	19
Languages in Education & CLIL .....	19
Making multilingual language teachers digital in Denmark: ensuring quality in digital language teaching.....	20
Mediation and exploiting one's plurilingual repertoire: Exploring classroom potential with proposed new CEFR descriptors.....	21
Online text analysis tools for test development and validation.....	22
Thumbs up or thumbs down? Putting apps to the test.....	22
<b>Paper presentations .....</b>	<b>24</b>
<b>WIP presentations .....</b>	<b>100</b>
<b>Poster presentations .....</b>	<b>114</b>

# Plenary presentations

---

## Connecting policy and practice at European level

The EU guideline for language teaching in school was agreed between the Heads of State and Government fifteen years ago: every European citizen should have the opportunity to learn two foreign languages from a very early age. However, European surveys indicate that around half of all European citizens are still unable to speak even one foreign language. This is sadly the case even for the generation that has spent the past fifteen years in school. New and better connections therefore have to be made between policy and practice. The European Commission is working together with the Member States to make language teaching and learning more efficient and at the same time more inclusive. Recent initiatives take account of technological progress and new trends in assessment. They also reflect an increased awareness of global developments, such as the role of English as lingua franca and the influence of languages from outside Europe. One common challenge, which is reinforced through the current refugee crisis, is that an increasing number of pupils speak a different language at home than in school. Educators must be prepared to deal with the specific needs and expectations of a more heterogeneous population of learners. Attention to the linguistic background of each individual is an essential element of successful teaching in school. If it is ignored, diversity can lead to learning difficulties and ultimately to school failure. But if it is properly recognised, it can create a more inclusive environment, paving the way to better learning outcomes across all subjects in later years and to a smoother acquisition of better language competences for all.

**Kristina Cunningham**, Senior policy officer, Directorate General for Education, Culture, Youth and Sport, European Commission, Brussels

Kristina Cunningham is the senior policy officer in charge of multilingualism at the European Commission. She previously worked as a sales and marketing manager in the private sector. She holds a Masters degree in Business Administration and Modern Languages from the University of Gothenburg, Sweden, and a Certificate of Political Studies from the Institute of Political Sciences in Paris, France. She is involved in the development of strategies for making multicultural and multilingual classrooms more efficient and socially more inclusive, and to support teachers in their evolving roles.

## Construct and content in context: Implications for language learning, teaching and assessment in China

Context is vitally important in conceptualizing the construct and specifying the content of language learning, teaching and assessment; however, the unique features of local contexts are often difficult to identify and capture. In a rapidly changing globalised world, the experience of China will be used to discuss these issues. In this presentation, I will illustrate the importance of contextualized construct definitions and content specifications on language education in China with a specific focus on English and the reform programme that is currently underway. Following the theme of the

conference and in making connections, I will highlight the current contextual conditions instead of historically embedded structures or practices (Marginson & Rhoades, 2002). To be specific, I will discuss the impact of China's fast-growing economy, its multilingualism, and the advancement in information and communication technology on learning, teaching and assessment.

**Dr Jin Yan**, Professor of Applied Linguistics at the School of Foreign Languages, Shanghai Jiao Tong University, China

Dr Jin's research focuses on the development and validation of large-scale and high-stakes language assessments. She is currently Chair of the National College English Testing Committee in China and Vice President of the Asian Association for Language Assessment. She is also co-editor-in-chief of the Springer open-access journal Language Testing in Asia and is on the editorial board of international and local journals, including Language Testing, Classroom Discourse, The Journal of Asia TEFL, and the International Journal of Computer-Assisted Language Learning and Teaching.

## Language policy and social cohesion: What links between social environment and regimes of learning and assessment?

This talk will discuss links between language, and specifically multilingualism, and social conflict, violence and disruption of the civil order. While many of the examples and evidence will be drawn from Asian settings, and an account given of the Language, Education and Social Cohesion initiative of UNICEF, and related activities of UNESCO, links will be made with European settings of contemporary challenges to social cohesion. It will be argued that while historians, social scientist of different kinds, sociologists and many educators, including language educators, fail to see the relations between questions of language and the social order close examination of social conflicts, and a non-reductive understanding of language, reveals deep links between them. Language conflicts however are much more amenable to resolution, through research and dialogue, than other causes of social disorder. The talk will trace ways in which effective education practice focused on maximising learning, and overcoming intergenerationally inherited language based disadvantage, including robust roles for assessment, can make a powerful and distinctive contribution to social cohesion.

**Dr Joseph Lo Bianco**, Professor of Language and Literacy Education, Melbourne Graduate School of Education, University of Melbourne, Australia

Professor Lo Bianco wrote Australia's National Policy on Languages in 1987, the first multilingual national language policy in an English speaking country, and was Chief Executive of the National Languages and Literacy Institute of Australia until 2002. His language policy advice was used at the Sydney, Athens and London Olympic Games, and by many countries including South Africa, Hawaii, Italy, and by Pacific Island countries. Among his many roles, since 2011, he has been senior research advisor for LUCIDE, a European Commission project on Languages in Urban Communities - Integration and Diversity for Europe, and in 2012, became Research Director of the UNICEF Language and Peace Building Initiative in Malaysia, Myanmar and Thailand. In 2014, he was appointed academic advisor to the National Research Centre for Foreign Language Education at Beijing Foreign Studies University, and is also Immediate Past President of the Australian Academy of the Humanities, the first educator elected to this role. He is also the author of 130 publications.

## Language testing washback and impact in our globalized world

It has been over twenty years since Wall and Alderson (1993) published their landmark paper about the Sri Lankan impact study and posed the question, “Does washback exist?” (Alderson & Wall, 1993). Two decades ago, influenced by ideas put forth by Arthur Hughes (1993), I proposed a simple visual image of washback, which depicted the participants, processes and products involved in the washback phenomenon (Bailey, 1996), though the column for “processes” was left blank at that time. Now, after many investigations of washback and test impact, we know much more about the phenomena involved. In this presentation, I will revisit the research literature on washback and test impact, in hopes of making further connections between learning and assessment in our globalized world.

**Dr Kathleen M Bailey**, Professor of Applied Linguistics, Middlebury Institute of International Studies at Monterey, USA

Dr Bailey works in the TESOL-TFL Program at MIIS (where she has taught since 1981), and also teaches in the EdD and MA Programs at Anaheim University. She is President of both The International Research Foundation for English Language Education, and American Association for Applied Linguistics, and has received a number of awards for her services to TESOL, including the James E Alatis Award and the Heinle Lifetime Achievement Award. Dr Bailey's research interests include teacher development and supervision, language assessment and classroom research, and she is the editor, co-editor, author or co-author of 15 books.

## Making the connections: digital innovation and diagnostic feedback

In recent years there has been a widespread change in the understanding of the relationship between learning and assessment. The perception of learning and assessment as being diametrically opposed is increasingly recognised as a false dichotomy. There is growing awareness of how assessment can be used to promote learning. This approach is encompassed in the concept of Learning Oriented Assessment (LOA).

LOA emphasises the primacy of the task in the learning cycle: intended learning outcomes, language activity, observation, feedback and adjustment. The use of new technologies can be incorporated into this cycle to make it more effective and efficient. The use of connected devices allows tasks to take place outside of the classroom to open up the possibility of learning affordances happening anywhere at any time.

We have developed, by combining artificial intelligence with task-based pedagogy, a teaching tool which uses automated writing assessment to provide diagnostic feedback. Creating a virtual learning environment to enable and encourage learners to improve their writing. We discuss the challenges of developing an automated placement model for writing. How does the artificial intelligence behind the model learn to mark as accurately as a trained examiner? What data does the artificial intelligence learn from? And, indeed, how do you create this artificial intelligence?

In the LOA approach the learner is at the heart of assessment and learning, and the tool has been designed and built to create different learning opportunities and effective learning outcomes. We will share results, of trials and feedback, which show how the tool has been used by learners to improve their writing.

**Dr Helen Yannakoudakis**, Senior Research Associate, Computer Laboratory, University of Cambridge, UK

Helen is a Senior Research Associate at the Computer Laboratory of the University of Cambridge, working on Automated Language Teaching and Assessment. She is also a fellow at Girton college, a committee member of Women@CL, and a Research & Development specialist at iLexIR. She holds a PhD in Natural Language and Information Processing from the University of Cambridge, during which she also worked on the English Profile Programme in collaboration with Cambridge English Language Assessment; an MPhil in Computer Speech, Text and Internet Technology; and a BSc in Computer Science. Her research interests lie at the intersection of Computational Linguistics, Machine Learning and Visualisation, and particularly revolve around the areas of automated language teaching and assessment, self-assessment and tutoring systems, error detection and correction, and second language acquisition.

**Dr Ardeshir Geranpayeh**, Head of Automated Assessment & Learning, Cambridge English Language Assessment, UK

Ardeshir Geranpayeh is the Head of Automated Assessment & Learning at Cambridge English Language Assessment, a department of the University of Cambridge. Ardeshir holds a PhD from the University of Edinburgh on the comparability of language proficiency testing. He has 27 years' experience of test validation and has contributed to the design, development, revision and evaluation of several internationally recognised language proficiency tests. Ardeshir has published extensively on language proficiency testing and cheating detection.

# Panels

---

## Insights from research on sign language tests

Sign language test research and application is an emerging field in the European and global context. Existing sign language tests, such as for Swedish or French Sign Language, target children or adult learners. One of the biggest challenges sign language test developers face in many European countries is that their respective sign language is often not well documented/researched. Collaborations and exchange between the communities of spoken language and sign language, such as this panel or previous presentations at the LTRC, are currently rare and their importance cannot be highlighted enough. In this panel, four different presentations will be delivered on the development and use of sign language tests for different group of learners, including the potential of web-based technologies for sign language test delivery.

### Coordinator:

**Prof Tobias Haug**, Interkantonale Hochschule für Heilpädagogik, Switzerland

Professor Tobias Haug is director of the sign language interpreting training program in Zurich since 2004 and owner of the company Sign Language Assessment Services. He has been involved in numerous European and national research and development projects on sign language testing, with a focus on new technologies. He is also currently a student in the distance master programme in language testing at Lancaster University.

### Discussant:

**Mark Wheatley**, Executive Director, European Union of the Deaf

Mark Wheatley has been the Executive Director of the European Union of the Deaf (EUD) which is based in Brussels, Belgium since 2007 and is the co-author of the EUD book Sign Language Legislation in the European Union. Under his leadership EUD has grown to an organisation that now has six full-time member of staff, including a new internship programme for young Deaf leaders. He was regularly asked to deliver presentations and evidences in European Parliament, European Commission, European Economic and Social Committee and the United Nations.

## Presentation 1 – Teaching and Assessing British Sign Language

The topic of this session is the UK community of people who are pre-lingually deaf and who use British sign language (BSL) to communicate. Ninety percent of deaf children are born in hearing families and hence the learning of sign language is by no means a matter of course. Although sign languages share many similarities with spoken ones, they do vary in significant ways. We will look at some of the key features of BSL and its varieties, both in terms of structure and dialect. The second focus is on the accreditation of sign language skills. The vast majority of candidates consist of hearing people who learn BSL as a second language in order to communicate with the deaf. Again, sign languages have unique features. While there are

systems to annotate sign language, there is no written form of sign language as such. In the English educational system this fact has implications for the national status of BSL and its equivalence with spoken languages. But it also has implications for language learning and testing. The question is whether and how we can transfer practice established in spoken languages to the assessment of sign languages and create positive washback on the teaching and learning of BSL.

**Dr Philida Schellekens**, Schellekens Consultancy, UK

Dr Philida Schellekens is a specialist in adult foreign language learning. While her work has been largely in spoken languages, especially English, she has 20 years' experience of working with the deaf. She has worked on the accreditation of BSL and wrote the UK National Occupational Interpreting Standards, which are used by both spoken and sign language interpreters. She has also been a trustee of Signature, the national awarding body for communication with deaf people, since 2012.

## Presentation 2 – Elicited imitation tasks (EITs) as a tool for measuring sign language proficiency in L1 and L2 signers

In previous literature, elicited imitation tasks (EITs) have been discussed with regard to the effect that memory skills have on performing tasks. More recent studies have shown, however, that EITs are a reliable tool for measuring language proficiency for L1 users and L2 learners (Klem et al., 2015; Gaillard & Tremblay, 2016). There have also been recommendations for minimising the negative impact of poor memory skills, for example, by shortening sentence structures.

In contrast to spoken languages, which are merely linear in structure, sign languages operate in the gestural-visual mode, which relies on a visual pattern that allows for a degree of simultaneity in production. For instance, when signing a single lexical sign, the shape, movement and location of the hand combine to express phonological properties at the same time. Additionally, there are more complex signs with internal morphological structures that involve multiple hand shapes, movements and locations. Such features need to be taken into account when valid and reliable EITs are developed for signed languages, and in recent years, there have been a growing number of sign language tests developed within the framework of EITs, e.g. American Sign Language, ASL-SRT (Hauser et al., 2008), and Swedish Sign Language, SSL-SRT (Schönström, 2014).

In this talk, we will discuss sentence structure as well as the scoring method of the tests we have developed on two EITs for Swedish Sign Language: SSL-SRT, which is targeted for L1 signers, and SignRepL2, targeted for L2 signers. We found that for the L2 group, complex (single) signs can be used as test items, and there are qualitative differences related to the linguistic properties of signs. We will also describe different scoring paradigms for the respective tests. Our results will be presented and discussed in relation to the EIT theoretical framework.

## **Prof Krister Schönström, Stockholm University, Sweden**

Professor Krister Schönström is an Associate Professor in Linguistics at Stockholm University. He obtained his doctoral degree in Swedish as a second language for the deaf from Stockholm University in 2010. His research interests include issues related to bilingualism in deaf individuals and language learning, with a special focus on both the bimodal aspect of their bilingualism and the visual aspect of language learning. He has been involved in several projects related to sign language test development.

## **Prof Ingela Holmström, Stockholm University, Sweden**

Professor Ingela Holmström is an assistant professor and researcher at Department of Linguistics at Stockholm University. She obtained his doctoral degree in Education from Örebro University in 2013. Her research is directed towards communication issues in the interactions between deaf, hard-of-hearing, and hearing people, both inside and outside of school contexts. She has a special interest in deaf education and bimodal bilingualism.

## **Presentation 3 – Assessing morphosyntactic skills in LSF (French Sign Language): How and why focus on predicative structures?**

Up to now, no reference tools can be found to assess LSF competence and identify potential SLI. This is due to i) the paucity of linguistic descriptions of LSF, in terms of first language acquisition and developmental stages in children, and ii) the failure of previous attempts to adapt tests from other Sign Languages (SL) (no cross-linguistic, standardised tests are available, Courtin & al., 2010; Haug, 2008). Just as in spoken language, SLI in SL is characterised by heterogeneous language skills. Yet the speech modality induces a number of differences. Morphosyntactic disorders in SL can be linked to the way the signer uses: semantico-syntactic space (Quinto-Pozos, 2011), agreement morphology and classifier system (Morgan et al., 2007). In a previous study (Puissant-Schontz, 2013), we created a pilot assessment tool, which proved insufficient to investigate predicative structures, due to the lack of overall description of the predicative system. This paper aims at filling this gap, and proposes a more fine-grained classification of predicative structures in SL. After a corpus analysis of different speech situations, we select formal features in order to classify predicates: i) action predicates: manual contact with the body, manual orientation, manual movement (with a change of grammatical space), and configuration; ii) existence predicates: standard sign, gaze, chest movement, pointing, classifier; and iii) property assignment predicates: standard sign, facial expression and classifiers. We present hypotheses on the impact of the type and the number of clues in the acquisition. We then proceed to work out an assessment tool for 4 to 10 year-old children, with reception and production tasks, with a view to test sign language acquisition and diagnose potential SLI or delayed acquisition. The assessment tool could also be used as a basis for remediation protocols. And the classification of predicative structures could be used in other SLs.

**Laetitia Puissant-Schontz, Université Paris Nanterre & Laboratoire MoDyCo – CNRS UMR7114, France**

Laetitia Puissant-Schontz is a speech and language therapist working with deaf children (native signers with sign language deficit/disorders). As a PhD student, she is working on the description of LSF predicative structures from a linguistic and developmental point of view.

## Presentation 4 – Web-based sign language assessment: Challenges and innovations

Due to the increased demand to assess sign language learning in the context of secondary and tertiary education in many European countries, web- or mobile-based assessment solutions can provide a valuable resource for sign language learning and assessment. Most of the existing sign language online tests have been designed for specific tests only, whereas the newly developed sign language assessment portal aims at integrating different kinds of tests (e.g., receptive, expressive) with different kinds of tasks and response formats. A preliminary version of this portal has been used within the EU-funded project SignMET. In the first part of the presentation, some key features of the portal will be presented, followed by a discussion on advantages and disadvantages as well as ethical issues in mobile-based sign language testing. Additionally, the potential of natural sign language processing (mostly automatic sign language recognition) and its value for sign language assessment will be addressed.

**Prof Tobias Haug**, Interkantonale Hochschule für Heilpädagogik, Switzerland

Professor Tobias Haug is director of the sign language interpreting training program in Zurich since 2004 and owner of the company Sign Language Assessment Services. He has been involved in numerous European and national research and development projects on sign language testing, with a focus on new technologies. He is also currently a student in the distance master programme in language testing at Lancaster University.

## Lessons learnt from Italy in language learning, teaching and assessment

### Presentation 1 – European Policies and Language Education in Italy: managing the change

My talk is divided into three main parts. In the first I will focus on European language policies and strategic objectives as set out by the European Commission and the Council of Europe, stemming from the EU Recommendation of 18 December 2006 on key competences for lifelong learning and the CoE proposals aimed at ensuring quality in language education.

In the second part I will deal with the Italian decision to implement a competence based approach involving setting standards, fostering quality teaching and innovative methodologies, defining trainers' and teachers' profiles and supporting foreign language certification. I will focus on CLIL/EMILE as a change agent in education since this enables integration in the curriculum and benefits cross-section and cross-curricular dialogue, thus offering an example of an innovative educational model.

In the third part I will focus on the difficulties in implementing language policies and will analyse key areas to be faced by different actors: a) identification of needs and resources; b) training and qualifications for teachers; c) development of classroom applications; d) certification for students and teachers. To conclude, I will propose three guiding criteria for "managing the change": convergence, flexibility, sustainability.

**Gisella Langé**, Italian Ministry of Education, Italy

Gisella Langé is a Foreign Languages Inspector with the Italian Ministry of Education, advising on internationalisation and foreign languages. A researcher, writer and consultant on issues relating to Foreign Language teaching, curriculum development and intercultural education, she has extensive experience of working on culture and language learning solutions and web-based teacher training. An expert within the European Commission and the Council of Europe, Gisella is involved with special interest groups, studies and projects. Her work has included the development of language policies, early language learning and bilingual education. She is currently involved in national groups organized by the Ministry of Education on teacher training and in developing national guidelines for curricula and CLIL.

### Presentation 2 – Linking policy and practice: the example of e-twinning

eTwinning [www.etwinning.net](http://www.etwinning.net) offers a platform for staff (teachers, head teachers, librarians, etc.), working in a school in one of the European countries involved, to communicate, collaborate, develop projects, share and, in short, feel and be part of the most exciting learning community in Europe.

eTwinning promotes school collaboration in Europe through the use of Information and Communication Technologies (ICT) by providing support, tools and services for schools. eTwinning also offers opportunities for free and continuing online Professional Development for educators.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

In 2016 the word that characterised eTwinning Italy was MORE. eTwinning Italy was very successful; a growth was registered in all aspects of eTwinning. – more registrations, more projects, more quality labels, more teachers trained, and more visibility.

According to data provided by EACEA Italy had an increase of 87% of registered users; an increase of 51% of teachers involved in European projects; an increase of 34% of teachers involved in national projects. The number of users involved in online training organised at national level increased as well by 104%. Apart from the general appeal of this action, in which participation is not ruled by competition and bureaucracy is nil, part of its success can be explained with the promotion carried out at national level.

In 2016 eTwinning was promoted together with the National Plan for the Digital School (Piano nazionale Scuola Digitale) which encourages innovation in schools through the use of digital technologies; eTwinning was promoted, as well, in relation to the National Guidelines for the School Curriculum (Indicazioni nazionali per il Curricolo della scuola di base) which appeared particularly relevant to attract schools of the pre-primary and primary level: the methodology advocated in this policy paper reflects very closely what normally happens in eTwinning projects. In promotion we also made clear reference to the professional development of teachers seen as key to School Innovation as it is reflected by Law 107 "La buona scuola"

**Donatella Nucci**, National Agency INDIRE, Italy

Donatella Nucci is the Head of the Italian eTwinning NSS. She has been involved in eTwinning since its very beginning and she worked as well for the Central Support service for one year in Brussels. She likes to define herself as an eTwinner of the Jurassic era. Before eTwinning, she was involved in many European research projects in the 5, 6 and 7 Framework Programme. At the beginning of her career in Indire, she worked in the Eurydice Unit.

## **Presentation 3 – Promoting quality in language assessment at the national level: the CLIQ association**

To be confirmed

## The challenges of a learning oriented and multilingual school assessment policy

The last two decades a worldwide increased mobility can be observed. This mobility – among other challenges – impacts also the linguistic and socio-cultural diversity in schools. Exploiting this diversity, providing equal opportunities for all pupils and raising their language proficiency is a major challenge for many school principals and teachers. In this context, a powerful school assessment policy which monitors the (language) learning processes of all pupils and which takes into account the multilingual realities of many pupils is pivotal. In this panel 4 papers will focus on these challenges.

### Coordinators:

**Prof Piet Van Avermaet**, Ghent University, Belgium

Professor Piet Van Avermaet is head of the Centre for Diversity and Learning, at the Linguistics Department of Ghent University. He teaches 'multicultural studies', 'multilingualism in education' and 'language policy' at the same university. His expertise and research interests cover topics related to multilingualism in education, social inequality in education, language policy and practice in education, language policy and practice in contexts of (social) inclusion, language assessment, diversity and inclusion, integration and participation, discrimination in education, migration.

**Dr Koen Van Gorp**, Michigan State University, USA

Dr Koen Van Gorp received his PhD in 2010 from the University of Leuven and is currently Head of Foreign Language Assessment at the Center for Language Teaching Advancement (CeLTA) at Michigan State University. From 2010 till 2015 he was Director of the Certificate Dutch as a Foreign Language (CNaVT), the international test of Dutch as a foreign language. His research interests are task-based language teaching and assessment.

### Discussant:

**Jim Purpura**, Associate Professor, Columbia University

James E. Purpura is Associate Professor of linguistics and education in the Applied Linguistics and TESOL Program at Teachers College, Columbia University, where he teaches L2 assessment and L2 research methods. Jim is currently co-authoring a book (with Carolyn Turner) entitled Learning-oriented assessment in language classrooms: Using assessments to gauge and promote language learning (Routledge). He is also deeply immersed in Scenario-Based Assessment research. Jim is the Co-Editor-in-Chief of Language Assessment Quarterly, and the series co-editor of both New Perspectives on Language Assessment and Language Assessment at ETS: Innovation and Validation. Former President of the International Language Testing Association and expert consultant for the EALTA, he is currently a Fulbright Scholar at the University for Foreigners of Siena.

## Presentation 1 – Linking learning and assessment to support plurilingualism

The presentation will be organised around three core questions – What is learning? What is to be learned? What is the role of assessment in language learning? – and reports on projects that seek to implement a learning-oriented approach in practice.

Language learning is one of the first things we do in life and is a natural and engaging process. Children in multilingual communities can acquire several languages from an early age leading to extensive plurilingual competence. However, school-based learning too often fails to

equip youth with real-world, plurilingual skills needed in contemporary society (European Survey on Language Competences, 2012). Policies on educational assessment can be part of the problem. While assessment is a process bound up with learning, large-scale educational assessment can become systematised in ways that fail to promote language learning effectively (Jones and Saville, 2016).

Learning Oriented Assessment (LOA) as outlined in this talk seeks to address this problem. It is an *ecological* approach that seeks to get all aspects of the system to work in harmony to serve the most important goals. In this systemic view, different stakeholders are enabled to work to a common purpose putting tasks at the centre. This highlights the role of meaningful communication in the cognitive development of learners, and the practical utility of tasks for aligning the different parts of the system. It also favours the acquisition of plurilingual repertoires which link school-based learning with society in more effective ways.

**Dr Nick Saville**, Cambridge English Language Assessment, University of Cambridge, UK

Dr Nick Saville is Director of Research & Thought Leadership in Cambridge English (University of Cambridge) and Secretary-General of ALTE. He has a PhD in Language Assessment, an MA in TEFL and a BA in Linguistics. Before joining Cambridge English he taught at the University of Cagliari (Italy) and worked in Japan. He is joint Series Editor of the Studies in Language Testing (SiLT) series (jointly published by Cambridge English and Cambridge University Press) with Professor Cyril Weir and has recently completed a volume on Learning Oriented Assessment (LOA) with Neil Jones (SiLT 45).

## Presentation 2 – Towards a learning oriented language assessment: The role of a Toolkit Alternative Assessment

In order to identify students at risk of failing academically and to raise the language proficiency of every student, the Flemish Minister of Education wanted to introduce a compulsory standardised language test at the beginning of secondary education. An expert committee that was consulted on this issue advised against the idea of one standardised test and instead recommended to develop an assessment toolkit that would support schools in setting up a language assessment policy.

During 2012-13, a team of the University of Leuven and Ghent University created a framework for learning-oriented assessment. The framework should allow teachers to measure students' language proficiency in more depth and width than one single test does. First, all the language competences that students need to function academically were inventoried. Second, all available assessment instruments (tests, observation instruments, etc.) were linked to these language standards. Each instrument was evaluated and annotated on its validity, reliability, impact and practicality by a team of assessment experts. Third, scenarios and videos of good practices were developed and included in the toolkit to support teachers and school teams in setting up a language assessment policy. A team of language experts, teachers, school advisors and school inspectors reflected on the framework and evaluated its practicality. The toolkit is available free of charge on the website of the Flemish Department of Education.

This paper presentation will discuss how with the support of important stakeholders (the Flemish Department of Education, the inspectorate, the educational networks, headmasters and teachers) a Toolkit Alternative Assessment for primary and secondary education was developed helping teachers and schools develop a coherent language assessment policy that takes into account the dynamic and complex language development of each individual learner. We will address the challenges of implementing the toolkit in Flemish education.

**Dr Koen Van Gorp**, Michigan State University, USA

Dr Koen Van Gorp received his PhD in 2010 from the University of Leuven and is currently Head of Foreign Language Assessment at the Center for Language Teaching Advancement (CeLTA) at Michigan State University. From 2010 till 2015 he was Director of the Certificate Dutch as a Foreign Language (CNaVT), the international test of Dutch as a foreign language. His research interests are task-based language teaching and assessment.

## Presentation 3 – Schools as laboratories for exploring multilingual assessment policies and practices

The unprecedented increase in linguistic diversity in Flemish schools (Belgium) is assessed in many different ways. At least two dimensions can be identified: the assessment of children whose linguistic functioning in educational spaces is being subjected to evaluations and the assessment by different educational stakeholders of children's multilingual realities in school contexts. In this paper the dynamic interactions between these two dimensions will be discussed. How are the multilingual realities in different urban spaces assessed? How does this affect the valorisation of (language) proficiencies and practices? What are the implications of contemporary multilingual realities for assessment in education, more generally, and in (language) learning, in particular? Can alternative approaches to assessment succeed in providing a more complete picture of students' linguistic repertoires and contribute to better achievement, a heightened sense of self-efficacy, and more well-being? These questions will be addressed in this paper.

**Prof Piet Van Avermaet**, Ghent University, Belgium

Professor Piet Van Avermaet is head of the Centre for Diversity and Learning, at the Linguistics Department of Ghent University. He teaches 'multicultural studies', 'multilingualism in education' and 'language policy' at the same university. His expertise and research interests cover topics related to multilingualism in education, social inequality in education, language policy and practice in education, language policy and practice in contexts of (social) inclusion, language assessment, diversity and inclusion, integration and participation, discrimination in education, migration.

**Prof Stef Slembrouck**, Ghent University, Belgium

Professor Stef Slembrouck is Senior Full Professor in the Linguistics Department of Ghent University. His main research interest is in the role of communicative and interactive processes in the functioning of institutions, and with particular reference to globalisation-affected multilingualism. He has published widely on such processes in contexts of education, social work and health care.

**Fauve De Backer**, Ghent University, Belgium

Fauve De Backer is a PhD student at the Centre for Diversity and Learning (Ghent University). Currently she is working on a project of Multilingual Assessment in Education (MulAE). Her PhD focuses on the evaluation of multilingual pupils in primary education. Before she started her PhD, she worked as an assistant national manager for the Programme for International Student Assessment (PISA), co-developed the Toolkit Alternative Assessment for secondary education and researched the multilingual reality in schools for the MARS-project.

## Presentation 4 – Pupils' perceptions of multilingual accommodations in science assessments

The linguistic abilities of multilingual pupils impact their results on content-related tests. It is therefore unlikely that multilingual pupils who take a test on content (e.g. mathematics or science) are able to demonstrate fully what they know and are capable of (Menken, 2010). This leads to the question to what extent the assessment procedure succeeds in measuring the intended construct (Wolf, Kim, & Kao, 2012). To overcome this validity issue, accommodations may offer a fruitful solution. However, research on which accommodations are effective under what conditions is limited and often contradictory (Pitoniak et al., 2009). Even less research is available on pupils' perceptions of assessment in general and accommodations in multilingual assessment in particular. In Flanders, a study on multilingual assessment was conducted in 2016-17. A population of pupils was divided into three groups with specific test conditions: a condition with a written science test in the language of instruction (Dutch); a condition where pupils took a bilingual test in the school language and a translation in their L1 and, finally, a third condition in which pupils received the same bilingual test with extra audio support in both languages. In addition, for each condition, two pupils were interviewed (N=12) afterwards so as to find out (a) how they experienced the assessment; (b) why they chose (not) to make use of the read-aloud and/or translated accommodation (2<sup>nd</sup> and 3<sup>rd</sup> conditions only); (c) which extra support or accommodation they perceive to be necessary, as well as (d) their perceptions of fairness in the case of accommodations. Transcripts of the interviews were analysed using NVivo 11 Pro. The outcomes of these interviews will be presented in this paper. Preliminary findings already indicate that pupils perceive testing accommodations positively, though much depends on their proficiency in the language of instruction and their academic language proficiency in their L1.

**Fauve De Backer**, Ghent University, Belgium

Fauve De Backer is a PhD student at the Centre for Diversity and Learning (Ghent University). Currently she is working on a project of Multilingual Assessment in Education (MulAE). Her PhD focuses on the evaluation of multilingual pupils in primary education. Before she started her PhD, she worked as an assistant national manager for the Programme for International Student Assessment (PISA), co-developed the Toolkit Alternative Assessment for secondary education and researched the multilingual reality in schools for the MARS-project.

**Prof Stef Slembrouck**, Ghent University, Belgium

Professor Stef Slembrouck is Senior Full Professor in the Linguistics Department of Ghent University. His main research interest is in the role of communicative and interactive processes in the functioning of institutions, and with particular reference to globalisation-affected multilingualism. He has published widely on such processes in contexts of education, social work and health care.

**Prof Piet Van Avermaet**, Ghent University, Belgium

Professor Piet Van Avermaet is head of the Centre for Diversity and Learning, at the Linguistics Department of Ghent University. He teaches 'multicultural studies', 'multilingualism in education' and 'language policy' at the same university. His expertise and research interests cover topics related to multilingualism in education, social inequality in education, language policy and practice in education, language policy and practice in contexts of (social) inclusion, language assessment, diversity and inclusion, integration and participation, discrimination in education, and migration.

## The development of China's Standards of English and its potential application

In 2014, the Chinese government launched a national project “Developing China's Standards of English (CSE)”, aiming to develop a transparent, coherent and comprehensive proficiency scale for Chinese English learners, and a common framework of reference for English language learning, teaching, and assessment in China. This panel provides an overall introduction to the development of CSE, including its background, theoretical framework, procedures and updated progress. This panel also includes three in-depth presentations on the development of translation, interpretation and pragmatic proficiency scales. The potential application of CSE is discussed with examples from a new English test system in China.

### Coordinator:

**Sha Wu**, National Education Examinations Authority, Ministry of Education, China

Ms. Sha Wu is now serving as Director of Foreign Language Assessment Department, National Education Examinations Authority, Ministry of Education, China. She got her MA in English and American Literature from Beijing Normal University and is now pursuing her PhD degree in language testing in Tsinghua University. She has participated in the test development and administration of language test programs for many years and is now engaged in the development of China's Standards of English (CSE) and National English Test System.

### Discussant:

**Dr Jin Yan**, Professor of Applied Linguistics at the School of Foreign Languages, Shanghai Jiao Tong University, China

Dr Jin's research focuses on the development and validation of large-scale and high-stakes language assessments. She is currently Chair of the National College English Testing Committee in China and Vice President of the Asian Association for Language Assessment. She is also co-editor-in-chief of the Springer open-access journal *Language Testing in Asia* and is on the editorial board of international and local journals, including *Language Testing*, *Classroom Discourse*, *The Journal of Asia TEFL*, and the *International Journal of Computer-Assisted Language Learning and Teaching*.

## Presentation 1 – The Development of China's Standards of English and its Role within China's National Policy towards English Education

English learning and teaching is implemented at all educational levels in China. However, the educational system in China is segmented with different governmental jurisdictions. National curriculum objectives, syllabuses and assessment criteria are developed and implemented under the auspices of these different departments. In 2014, the Chinese government launched a national project “Developing China's Standards of English (CSE)”, aiming to develop a transparent, coherent and comprehensive scale for Chinese English learner, a common framework of reference for foreign language learning, teaching, and assessment in China. In addition, CSE also takes into consideration the current situation of China's foreign language education and aims at its future demand and development. This paper first introduces the purpose, principles and theoretical framework of CSE development. Then results of the descriptors of the CSE are reported. Finally, the role the CSE will play in China's national policy

towards English education and assessment is discussed with some examples of the application of the CSE.

**Prof Jianda Liu**, Guangdong University of Foreign Studies, China

Professor Jianda Liu is a professor and director of the National Key Research Center for Linguistics and Applied Linguistics at Guangdong University of Foreign Studies, China. He was vice director of National Education Examination Authority, Ministry of Education, China. He received his PhD in Applied Linguistics from the City University of Hong Kong. He is now leading the development of China's Standards of English. His recent book entitled *Measuring Interlanguage Pragmatic Knowledge of EFL Learners* was published by Peter Lang.

## Presentation 2 – The Development of the pragmatic ability scales in CSE

The development of the pragmatic ability scales is part of the national project “Developing China’s Standards of English (CSE)” launched by the Chinese government launched in 2014. In this talk, I will first briefly talk about how pragmatic ability is defined in the CEFR and CLB, then I will discuss in detail how pragmatic ability is defined in CSE and the descriptive framework of pragmatic ability scales in CSE. I will also discuss the principles we follow and the steps involved in developing the pragmatic ability scales. Some issues or problems concerning the operationalisation of the descriptive framework of pragmatic ability will be discussed. The distinguishing features of the language ability framework for CSE, especially for pragmatic ability scales will also be addressed in this talk.

**Prof Han Baocheng**, Beijing Foreign Studies University, China

Professor Han Baocheng is now serving as Deputy Director of the National Research Center for Foreign Language Education of Beijing Foreign Studies University, China. He is also Secretary-General of the Foreign Language and Literature Committee of the China Council for Self-taught Higher Education Examinations. His main research areas include language education and language assessment. He is particularly interested in the understanding of the nature of language ability, second language proficiency development and conducting research on teacher education and development.

## Presentation 3 – Interpreting Competence in China’s Standard of English

With reference to Chomsky’s “competence”, Bachman’s “communicative language ability” and “overall language proficiency” from the CEFR, this paper reviewed the definitions, notions and concepts of “language competence”. By analysing the relations between interpreting competence and language competence, the paper positioned interpreting in the scope of overall language proficiency, put forward a proposal on interpreting proficiency scales and interpreting descriptor frameworks, so as to provide a theoretical basis for the interpreting proficiency scale of China’s Standards of English (CSE) by NEEA (National Education Examinations Authority).

**Dr Wang Weiwei**, Guangdong University of Foreign Studies, China

Dr Wang Weiwei, Research Fellow of Guangdong University of Foreign Studies, Committee Member of China Translators’ Association. Dr Wang published over 10 journal articles and is now leader or key member of several research projects of both provincial and national level including China's Standards of English (CSE) Project by

NEEA (National Education Examinations Authority). As an active conference interpreter, she provided simultaneous and consecutive interpreting service to over 100 international conferences.

## **Presentation 4 – A Report on the Development of Translation Proficiency Scales under the China's Standard of English Project: Construct, Methodology and Process of Research**

As part of the project of China's Standards of English (CSE), the development of (Chinese to English) scales of translation proficiency is aimed at providing the national reference framework of the teaching/training syllabus as well as assessment of the translation proficiency. While there has been little research on can do description of such proficiency (not to mention a Chinese-English one), this project takes the initiative.

This paper will report the construct, research methods, process and the latest development of the translation proficiency scales currently under the project of China's standards of English sponsored by NEEA. The following topics will be covered:

- 1) The construct of translation proficiency based on the literature review and needs analysis;
- 2) The parameter system of translation proficiency scales designed for the current project, ensuring observability as the fundamental principle;
- 3) The semantic and syntactical norms for can do descriptions;
- 4) Use of various research methodologies, including corpus-based method, intuition, qualitative surveys, quantitative analyses as well as the methods involved in the rewriting and compiling the descriptors and identification of typical translation activities;
- 5) Validation of all the descriptors in terms of categorization and calibration as well as the problems therein;
- 6) A few thoughts on the above steps and suggestions for the studies that will follow.

**Dr Feng Li**, Heilongjiang University, China

Dr Feng Li obtained her PhD of Chinese linguistics from Heilongjiang University. She has published books and papers on Chinese-English translation, business translation, contrastive studies of English and Chinese and linguistic typological studies.

## **Presentation 5 - Developing National English Test System on the Ground of China's Standards of English**

With the backdrop of economic globalisation, China's 2020 Education Reform and Development Plan has highlighted the need to improve students' ability to participate in international communication activities. Being essential in international communication, language ability, especially English language ability, becomes a key word in the on-going education reform. In the 2014-2020 Reform of Examinations and Enrollment System, launched by China's State Council,

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

one important measure is to develop the Foreign Language Assessment Framework to promote the overall quality and efficiency in foreign language education. Within the Framework, there are two important components, i.e. China's Standards of English (CSE), a comprehensive proficiency scale, and National English Test System (NETS), a new English test programme. Beginning with an introduction to NETS's purposes, test objectives and test formats, this paper will attempt to illustrate how the new test system is developed on the basis of CSE, from test objectives outlining to score reporting. Finally this paper will discuss the school admission policy changes NETS will help to accelerate and the potential impact on English teaching and learning.

**Dr Han Yu**, National Education Examinations Authority, Ministry of Education, China

Dr Han Yu is Vice President of National Education Examinations Authority (NEEA), Ministry of Education, China. He obtained his PhD degree from Tsinghua University and had been working as Director of Tsinghua University Admission Office for many years before serving as Deputy Governor of Ganzi Tibetan Autonomous Prefecture in Sichuan Province. He joined NEEA in 2015, and is now supervising the reform and development of some important national examinations. His current research interests involve education policy and management.

**Sha Wu**, National Education Examinations Authority, Ministry of Education, China

Ms. Sha Wu is now serving as Director of Foreign Language Assessment Department, National Education Examinations Authority, Ministry of Education, China. She got her MA in English and American Literature from Beijing Normal University and is now pursuing her PhD degree in language testing in Tsinghua University. She has participated in the test development and administration of language test programs for many years and is now engaged in the development of China's Standards of English (CSE) and National English Test System.

# Workshops

---

## Erasmus+: Funding opportunities for language education

After a brief introduction to the latest European policy developments in the field of education, with special reference to languages, the main opportunities for funding in Erasmus+ are presented, both for European Programme countries and for international worldwide mobility.

Schools, higher education institutions and adult education organisations have a wide range of opportunities in Erasmus+: from Key Action 1 and 2 Partnerships for Mobility (including in-service training of teachers) and Strategic cooperation, to eTwinning for virtual cooperation among schools, to the Online Linguistic Support for Erasmus HE students to the Electronic Platform for Adult Learning (EPALE).

A special focus will be dedicated to analysis of impact and the support of quality projects, with initiatives such as the European Language Label and other awards.

**Lorenza Venturi**, National Agency INDIRE, Italy

Lorenza Venturi is Head of Unit at the Erasmus+ National Agency INDIRE in Italy, where she coordinates the activities of communication, information, dissemination and exploitation of results at national level, including initiatives for the valorisation of best practices such as the European Language Label. She is also in charge of the Studies and Analysis Unit, focussing on the analysis of the impact of the Erasmus+ Programme on the participating individuals and institutions: schools, universities and adult education organisations. Finally, she is Head of the EPALE Unit, in charge of implementing the Electronic Platform for Adult Learning in Italy.

## Languages in Education & CLIL

CLIL is a dual-focused approach in which an additional language is used for the learning of both content and language with the objective of promoting both content and language mastery to pre-defined levels. It does not simply denote the context (for example teaching science through French in an Australian High School), but the teaching and learning methodologies used in the context. This presentation describes the situation when the term CLIL was launched in the 1990s; how and why the definition has remained stable; and the ways in which socio-educational forces, evidence of impact, and media-rich learning environments remain powerful drivers for the future of teaching and learning languages.

**David Marsh**, Director Innovation & Outreach, EduClusterFinland, University of Jyväskylä Group, Finland

David Marsh has experience of tasks in over 40 countries, contributed to over 150 publications, and received 5 degrees from the United Kingdom, Finland, and Spain. Instrumental in developing processes for transforming language learning, he co-launched Content and Language Integrated Learning (CLIL) and now works on developing the implementation of CLIL models globally. His current work-in-progress *The Children of Cyberspace: Towards a New Understanding* is due for publication in 2018.

## Making multilingual language teachers digital in Denmark: ensuring quality in digital language teaching

This workshop will explore how we can upgrade/educate our teachers to become competent digital users when teaching languages. It will focus in particular on the interplay between new digital methods for teaching and the pedagogical approach to using them in the classroom. The workshop will present different digital and pedagogical initiatives developed by Studieskolen, Denmark's leading foreign language school. Studieskolen's foreign language department offers language training in more than 20 different languages and has a multilingual team of 75 teachers. All of our teachers have disparate backgrounds regarding IT skills and knowledge of how to use them in the classroom.

As a way of overcoming this challenge, Studieskolen has focused on the quality of our language teaching and thus our teachers for the past few years. When using digital tools in language teaching it is of vital importance to know how to use them in a meaningful and pedagogical way. To make sure our teachers can integrate both elements in their teaching practice, we have developed in-house courses and online accessible tool boxes. In this workshop we will present our courses and show how they have had an impact on the quality of digital language teaching at Studieskolen and the way our teachers have integrated and implemented digital and pedagogical tools into their teaching practice. Furthermore, we will give you an insight to the challenges we have experienced during this process and our plans to overcome them. The workshop will present the following:

- IT-pedagogical certificate – all teachers with limited knowledge of how to use IT in their teaching practice are required to complete four basic modules.
- Digital toolbox – an online site where instruction videos and texts show how to use digital tools for teaching.
- Pedagogical toolbox – an online site with basic pedagogical teaching tools.

### **Stine Lema**, Studieskolen, Copenhagen

Stine Lema is Head of the English Department at Studieskolen in Copenhagen. Stine obtained her MA in Adult Education and Cultural & Language Studies at Roskilde University and Macquarie University, Sydney. Stine spent five years in the United Kingdom where she gained a sound knowledge of British English language and culture. In the United Kingdom she taught the exam *Cambridge English: ESOL Skills for Life* and acquired her CELTA teaching qualification and experience. She is also the Cambridge English Team Leader and manages the development of Cambridge English examinations in Denmark.

### **Charlotte Lorenzen**, Studieskolen, Copenhagen

Charlotte Lorenzen is Managing Director of Studieskolen's Foreign Languages and Language Test Centre. She holds an MA in Education from Durham University and has worked in the Danish educational system most of her career. Over the past five years, Charlotte Lorenzen has worked strategically to develop possibilities to offer language teachers education in didactic and IT skills. Moreover, she has worked strategically on developing Denmark's first international language test centre.

## Mediation and exploiting one's plurilingual repertoire: Exploring classroom potential with proposed new CEFR descriptors

This workshop will give participants the opportunity to familiarise themselves with the product of a 2014–17 Council of Europe project to update the CEFR's 2001 illustrative descriptor scales and to develop descriptors for areas that were not covered in the original set, namely mediation activities, mediation strategies, online interaction, reactions to literature, and plurilingual and pluricultural competences. The focus will be on giving participants an opportunity to (a) discuss the potential that the descriptors offer for enriching language pedagogy and assessment by encouraging collaborative interaction, the co-construction and mediation of ideas and the exploitation of learners' plurilingual profiles in class, and (b) envisage the kinds of teaching and assessment tasks that operationalisation of the descriptors suggests. The approach taken to mediation in the project reported on is far broader than in some interpretations, including pedagogic and social mediation as well as conveying received meaning – the latter being the main sense in which mediation was introduced in the CEFR in 2001. The development of the extended categories of descriptors followed the same three-phase process as in the original CEFR descriptor research in the 1993-96 Swiss research project (intuitive authoring and editing, qualitative validation in workshops, quantitative validation through Rasch model scaling), but on a larger scale. Some 150 institutions and 1,300 people took part in the three validation phases that took place between February 2015 and February 2016, with a formal consultation phase from October 2016 till February 2017. The definitive version is expected to be published in December 2017. A brief report on the piloting of the descriptors to date will also be given, together with an invitation to join the network for further piloting in the academic year 2017-18.

### **Dr Brian North**, Eurocentres, Switzerland

Dr Brian North has extensively collaborated with Council of Europe during his long career at Eurocentres. He developed the CEFR levels and illustrative descriptors, co-authored the CEFR itself, the prototype European Language Portfolio (ELP), the Manual for relating examinations to the CEFR, the European Association for Quality Language Services (Eaquals) CEFR Core Inventories for English and French, and the European Profiling Grid for language teacher competences (EPG). His most recent project is the 2014-17 update and extension of the CEFR descriptors.

### **Prof Enrica Piccardo**, OISE – University of Toronto, Canada/Université Grenoble-Alpes, France

Professor Enrica Piccardo is Associate Professor at OISE – University of Toronto and at the Université Grenoble-Alpes, France. She has extensive experience in language teaching, teacher education and second/foreign language education research. A CEFR specialist, she has coordinated two European Centre of Modern Languages (ECML) CEFR-related projects and collaborated in the Council of Europe project to develop descriptors for mediation, online interaction and plurilingual/cultural competences. Her other research focuses on complexity theory, emotions and creativity in language education, assessment and its role in the curriculum.

## Online text analysis tools for test development and validation

New online text analysis tools can offer test designers important new insights into the difficulty level of reading and listening texts to use in testing contexts, and also important tools for validating language tests. This talk examines a high-powered web-based text analysis tool called TextInspector.com. It discusses the tool's growing research base, namely the analysis of thousands of Reading, Listening, Writing and Speaking texts and transcripts, benchmarked to the CEFR, which lay behind the development of the metrics used in the tool. It then demonstrates the potential of the tool for test developers, and the possibilities of developing similar text analysis tools for other languages.

**Prof Stephen Bax**, Open University, UK

Professor Stephen Bax is Professor of Modern Languages and Linguistics at the Open University in the United Kingdom. His research focuses on reading and technology, and his books include 'Discourse and Genre' (2011, published by Palgrave Macmillan). He was awarded the TESOL Distinguished Researcher Award 2014 for his 2013 article in the journal Language Testing, which used eye-tracking technology to analyse L2 reading, and more recently has developed an online tool called Text Inspector for analysing text difficulty in testing and teaching.

## Thumbs up or thumbs down? Putting apps to the test

Have you ever used digital technology to learn and teach a language? Have you ever used digital technology in language assessment? If 'yes', then you're one of millions of teachers and learners who face the latest frontier in education – technology. While technology is making its mark in education, the digital competencies of teachers are not necessarily moving forward in lock step with advances in technology. As one teacher put it when asked to rate herself against a framework describing digital competence: 'I didn't realise how much I don't know. I consider myself pretty au fait with technology compared to most people my age in terms of both theoretical knowledge and practical/everyday use. Yet these are aspects I'd never even considered.'

In this presentation we will briefly explore key digital competencies that language teachers need in order to be effective professionals and will present a framework of digital teaching competencies developed by Cambridge English as a diagnostic tool. We will then move onto one specific aspect of digital competence, which relates to evaluating digital resources and addresses the basic question which many teachers face: 'What makes a good language app?' We will explore some basic tenets of the Cambridge English philosophy for digital learning, such as a strong pedagogy, a holistic approach to language learning, the need for digital to add something new, a built-in progression across levels, and teacher development to make good use of digital tools. We will discuss key evaluation criteria that teachers and learners need to consider

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

when selecting and using language learning apps, and will also explore the value-added potential of digital apps, such as the opportunity for personalised learning, instant feedback and gaming features.

**Geoff Stead**, Cambridge English Language Assessment, University of Cambridge, UK

Geoff Stead is Director of Digital and New Product Development at Cambridge English Language Assessment. He has been designing and developing digital learning products for over 20 years. He joined Cambridge English from Qualcomm in San Diego, where he was a Senior Director specialising in mobile and learning technologies. While there he won multiple awards for their internal employee app-store as well as over 40 internal apps and learning sites. Prior to this, he worked for Tribal in Cambridge, leading their digital learning business.

**Dr Evelina Galaczi**, Cambridge English Language Assessment, University of Cambridge, UK

Dr Evelina Galaczi is Head of Research Strategy at Cambridge English Language Assessment. She has worked in language education for 25 years as a teacher, teacher trainer, materials writer, programme administrator and assessment specialist. She especially enjoys working on assessment literacy, the integration of learning and assessment, and the role of digital technologies in assessment and learning. She regularly presents at international conferences and has published papers on speaking assessment, computer-based testing, and paired speaking tests.

# Paper presentations

---

## ¿Aprender en varias lenguas incide en el rendimiento en las competencias básicas?

La educación plurilingüe pretende integrar la adquisición de los contenidos de las disciplinas académicas con el desarrollo de las destrezas y de los repertorios lingüísticos en distintas lenguas. Pero, ¿tenemos evidencias sobre qué efectos tiene este tipo de enseñanza en el aprendizaje de las competencias básicas? Para responder a esta pregunta, esta investigación se ha centrado en analizar si el aprendizaje en varias lenguas afecta en el rendimiento en las competencias básicas del alumnado que participa en los proyectos de enseñanza plurilingüe. Para ello se han aplicado diferentes tipos de pruebas a lo largo de cuatro cursos escolares (2011-2014) a más de 4.000 alumnos de Educación Primaria (9-11 años) y Educación Secundaria (12-14 años) en los centros trilingües del País Vasco. Las pruebas de rendimiento, que han sido diseñadas por ACER (Australian Council for Educational Research) y Cambridge English Language Assessment, han permitido determinar cómo influye aprender en varias lenguas en el rendimiento del alumnado en diferentes competencias básicas: Comprensión lectora en lengua vasca (euskera) y lengua castellana (español), competencia matemática, competencia en cultura científica, y lengua inglesa (English). El proyecto ha sido liderado por el ISEI-IVEI (Instituto Vasco de Evaluación e Investigación) del Departamento de Educación del Gobierno Vasco. La investigación analiza de manera comparativa los resultados obtenidos por el alumnado que ha cursado sus estudios en tres lenguas (euskera, español e inglés) en relación con aquellos alumnos que lo han hecho en dos (euskera y español). Los resultados demuestran que los alumnos que cursan un programa de enseñanza plurilingüe obtienen resultados similares a quienes lo hacen en programas bilingües o monolingües en las competencias científicas: matemáticas y ciencias. Igualmente, en las competencias lingüísticas (L1 y L2) los resultados son equiparables. Sin embargo, en las pruebas de L3 (inglés) el rendimiento del alumnado plurilingüe es significativamente más alto.

**Jesús Grisaleña Urrecho**, Basque Institute for Research and Evaluation, Spain

Jesús Grisaleña has a degree in Spanish Language and Literature from the University of Barcelona and a master's degree in Applied Linguistics from the University of the Basque Country. He works in the Basque Government's Evaluation and Research Institute, designing evaluations and research for the Basque education system. He has published several papers related to language learning, school coexistence and organisation and management in educational centres. He is also the co-author of the book *Herramientas para directivos escolares*.

## A comparative study on the washback of writing tasks in two international EFL tests on Chinese test-takers

In mainland China, recent years have witnessed growing numbers of students studying abroad and undergoing testing through International English Language Testing System (IELTS) and the internet-based Test of English as a Foreign Language (TOEFL iBT) (Wang and Miao 2015); writing is often the weak link for many Chinese students (Guo & Qin, 2010) and the washback of IELTS and TOEFL iBT writing on Chinese students is worthy of researcher's attention. Adopting the mixed-method approach, this study comparatively makes an investigation into the washback of IELTS and TOEFL iBT writing on Chinese test-takers' preparation practice. A total of 242 college students who had registered for IELTS or TOEFL iBT participated in the present study. Research data were collected mainly through IELTS (N=139) and TOEFL iBT (N=103) preparation questionnaires and supplementally through semi-structured individual interview (N=23). A number of statistical analyses were performed on the questionnaire data, including exploratory factor analysis, descriptive analysis and difference tests, and qualitative interview data were coded through an inductive approach. The findings indicate that first of all, compared with TOEFL iBT writing, IELTS writing exerted significantly shorter washback during the whole preparation, but was not significantly different on a daily basis. Secondly, five types of test preparation practice were detected in IELTS and TOEFL iBT writing preparation and test-takers showed a similar pattern of preparation intensity, except in test preparation management. Moreover, this study finds that personal factors like experience, language competency and self-efficacy can shape the washback to a certain extent, while gender and academic background can hardly mediate the washback on IELTS and TOEFL iBT writing preparation. Explanations and implications of the study findings will be further discussed and it is hoped that this study offers important empirical evidence from test-takers for future consequential validity study on IELTS or TOEFL iBT.

**Xiangdong Gu**, Chongqing University, China

Xiangdong Gu is the Director of the Research Centre of Language, Cognition and Language Application at the Chongqing University in China and consultant in Cambridge English Language Assessment. She obtained her PhD in Linguistics and Applied Linguistics from the Shanghai Jiaotong University in China and furthered her study after her PhD at the University of California Los Angeles as a visiting scholar. Her interests mainly focus on content validity study and washback/impact study of large-scale and high-stakes national and international English tests.

**Yue Hong**, Chongqing University, China

Yue Hong is a master's candidate of foreign languages and literature at the Chongqing University in China. His interests mainly focus on language assessment and foreign language teaching.

**Chengyuan Yu**, Chongqing University, China

Chengyuan Yu is a master's candidate of foreign languages and literature in Chongqing University, China. His interests mainly focus on language assessment.

## Academic literacy and language proficiency in testing: Overlapping and diverging constructs

The construct of language proficiency focuses on speakers of other languages, whereas academic literacy is conceptualised as a set of skills needed for successful university study, which many native speakers struggle to develop. Clearly, there is overlap and divergence between the two and both inform English for Academic Purposes (EAP) assessment. Some areas of difference and similarity between the constructs are easily identified; however, they also intersect in ways that have not been examined closely before. By employing a socio-cognitive model of language proficiency (Weir 2005) as a basis, this presentation examines commonalities and distinctions between the two concepts. Language proficiency as operationalised by the CEFR is compared with academic literacy as defined by universities, to examine how the constructs align theoretically. Examples of how the constructs become conflated in assessments targeting higher levels will also be examined. Presenting these issues will reveal assumptions about the constructs, and questions for researchers to address. These include:

- To what extent is a certain level of general language proficiency a pre-requisite for acquiring academic literacy?
- To what extent is high-level language proficiency as conceptualised by the CEFR intrinsically academic?
- In what ways can academic literacy be considered an extension of language proficiency as conceptualised by the CEFR?
- Can language testers avoid conflating the concepts of EAP and general language proficiency?

Making the questions explicit and posing them to the language testing community will inform future EAP assessment, pedagogy and research. The recent trend of globalisation in higher education makes it important for universities to develop efficient forms of support for learners from various cultures. Programmes that improve academic literacy are possibly different to those aiming to improve language proficiency, and this potentially impacts on how assessment can support learning. The presentation will conclude by recommending research directions that begin to address these issues.

**Kevin Cheung**, Cambridge English Language Assessment, United Kingdom

Kevin Cheung is a Senior Research Manager at Cambridge English Language Assessment, where he works on assessment of writing, particularly at levels required for HE study. He is a chartered psychologist and before joining Cambridge English, lectured at UK universities in social psychology and research methods. His doctoral research focused on plagiarism and authorial identity in undergraduates. He has also taught writing as a basic skills tutor for the probation service, where he worked with offenders to improve their literacy skills.

**Mark Elliott**, Cambridge English Language Assessment, United Kingdom

Mark Elliott is a Senior Validation & Data Services Manager at Cambridge English Language Assessment. He leads data analysis for validation and research of the Cambridge English: Young Learners tests and Cambridge Assessment Admissions Testing. He also specialises in test equating, pretesting, and grading procedures. He holds an MA in Mathematics from the University of Cambridge and an MA in ELT and Applied Linguistics from King's College London.

## Aligning China Standards of English (CSE) with the CEFR

The Common European Framework of Reference for Languages (CEFR) has played a dominant role in language and education policy, language teaching, learning and assessment practice within and beyond Europe since its publication in 2001. China Standards of English (CSE) is a transparent, coherent and comprehensive scale for Chinese English learners, a common framework of reference for foreign language learning, teaching, and assessment in China. Although the CSE is developed specifically to the Chinese context rather than by adapting the CEFR to the Chinese context, it is still important and inevitable to align the CSE to the CEFR in light of international communication, the CEFR's influence worldwide, internalisation of the CSE and the improvement of Chinese students' English proficiency. This alignment study is carried out based on three research questions: Firstly, in what way are Chinese students related to the CEFR? Secondly, to what extent is the CEFR applicable to the Chinese context? And thirdly, how are the nine levels in the CSE related to the six levels in the CEFR? To answer these three research questions, linked questionnaires made up of CEFR descriptors will initially be distributed to the representative sample based on Rasch scaling, in order to establish the relationship between CEFR and Chinese students' English. Following this, some anchor items (descriptors) in CSE are anchored to the CEFR, covering all levels. Finally, when all the descriptors in CSE and CEFR are on the same scale based on Rasch analysis, the relative relationship between CSE and the CEFR can be established.

**Chuan Peng**, Guangdong University of Foreign Studies, China

Chuan Peng is currently a PhD student at the Guangdong University of Foreign Studies. She is working on the Chinese national project "Developing China Standards of English" with the National Education & Examination Authority (NEEA).

## An AUA validation study of the integrated teaching and assessment framework

How to assess students' learning of a foreign language has become one research focus worldwide for language teaching practitioners and researchers. In order to overcome the limitations of the summative one-shot assessment framework in the Chinese College English learning and teaching context, a four-stage and multi-agent integrated teaching and assessment framework was proposed as the result of synthesising the Assessment For Learning (AFL) theories and the related social learning theories, as well as taking into consideration the current educational technology development and the researcher's teaching experience. This presentation proposes the design validity argument for the integrated teaching and assessment approach whose interpretations and uses can be justified. With the application of theoretical framework of Assessment Use as Argument (AUA) put forward by Bachman, the backing for the inferences of the interpretive argument of the integrated Approach in this stage is mainly from theoretical analysis. The inferences, warrants, assumptions, and backing will be stated clearly to demonstrate the feasibility and viability of the integrated teaching and assessment approach on

students' performance and perceptions in order to explore a new way to improve the teaching of College English writing in the new era.

**Huang Jing**, China West Normal University, Nanchong,

Huang Jing is currently a lecturer in China West Normal University with eighteen years of teaching and researching experiences. She got her PhD degree in language testing from Tsinghua University in China. She has worked as a visiting scholar under the supervision of Prof. Paul Nation and Dr. Peter Gu in Victoria University of Wellington in New Zealand. Her research interests are classroom-based writing assessment and integrated feedback approach.

## **An investigation into scale descriptors for spoken English proficiency: Analysis based on descriptor pool**

In 2014, China's National Education Examinations Authority (NEEA) launched a national project for developing China Standards of English as the basic step towards establishing a national foreign language assessment system. One of the major tasks involved in developing China Standards of English-Speaking scale is to use the illustrative descriptors to portray oral English proficiency. A quantitative analysis of descriptors is conducted in this study so as to provide research evidence for speaking scale development and validation. About 4,300 existing descriptors of oral English proficiency were collected to establish the initial descriptor pool and using corpus software and statistical software for word frequency and key word analysis. This study looked into details the collocation of words, annotated, checked descriptor key words based on stylistic analysis, and interpreted and explained the features of oral English proficiency descriptors, in order to have a systemic, in-depth analysis and research upon proficiency descriptors therefore link up the descriptor analysis from lexical level to the sentence level. The study attempts to make explicit the syntactic structure of descriptors, and expatiate on the language feature of proficiency descriptors, which makes users further understand descriptors, grasps the collocation rules and principles among components, and thus provide evidence and guidance for the writing and revision of oral English proficiency descriptors.

**Wei Jie**, Shanghai JiaoTong University, Shanghai University of Business and Economics, China

Wei Jie is a PhD candidate in Language Testing and Assessment in Shanghai JiaoTong University, and associate professor in applied linguistics.

## **An investigation of the influence of age-related factors on the construction of China Standards of English**

Children within the age range of 5 to 12 are a very special group of language learners sharing distinctive characteristics that set them apart from adult learners. According to McKay (2006), these characteristics could be categorised into growth, literacy and vulnerability. For any effective evaluation systems involving young learners, considering these factors is somewhat indispensable during the developmental stage. As an essential component of China's assessment system for foreign language proficiency, China Standards of English (CSE) is

currently designed with 9 provisional levels containing a number of performance descriptors illustrating what learners at each level can do with the target language. The descriptors, however, have no apparent age-related indications on the assumption that they can be used for both young and adult learners. Since young learners constitute a large part of the population of lower-proficiency learners of English in China, it is necessary to explore whether these descriptors are feasible for young learners. Therefore, this study aims to explore how to take age-related factors into consideration when developing a nation-wide speaking performance standard without excessively restricting the scale's usability. Specifically, this study applies an exploratory sequential mixed methods design (Creswell 2014). In the qualitative phase, a series of small-scale teacher workshops were conducted to probe into the salient features of the age-related factors. The following quantitative phase aims to statistically compare teachers' and learners' scaling judgments on the collected descriptors with and without salient age-related features using a survey investigation which has been launched by the National Education Examinations Authority of China. Overall, with the integration of qualitative and quantitative analyses, this study will enhance our understanding of the influence of age-related factors on the construction of a nation-wide speaking performance standard and how to take them into account during the developmental stage.

**Jun Wang**, Shanghai Jiao Tong University, China

Jun Wang obtained her Master of Education from University of Sydney and Master of Research specialising in Language Testing from Macquarie University. She is currently a PhD Candidate in Shanghai Jiao Tong University, China. Both presenters are members of China Standards of English's research group.

## **Análisis del comportamiento de los calificadoros de una prueba de expresión escrita en el contexto de una prueba de dominio**

La comunicación describe cómo analizar el comportamiento de los calificadoros de la prueba de expresión e interacción escritas de una prueba de dominio. Para ello proponemos utilizar uno de los modelos de Rasch, el denominado Many-Facet Rasch Measurement, MFRM, con el fin de situar en la misma métrica los elementos de las facetas incluidas en el contexto de una evaluación de desempeño: candidatos, calificadoros, tareas y atributos evaluados. Para poder utilizar este método es preciso que cada examen sea calificado por dos examinadores que actúen de manera independiente y también requiere crear una red de calificadoros en la que queden conectados pruebas, candidatos y examinadores, de modo que se puedan comparar todos los datos en un mismo marco de referencia. De este modo es posible lograr un escalamiento conjunto según el nivel de competencia de los candidatos y conocer la severidad/benignidad de los calificadoros y determinar la dificultad de las tareas y de los atributos evaluados. El procedimiento también permite analizar la existencia de tendencia central o de efecto de halo en las calificaciones.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

**Juan Miguel Prieto Hernández**, Cursos Internacionales, Universidad de Salamanca, Spain

Juan Miguel Prieto Hernández holds a PhD in Spanish Philology from the University of Salamanca and is currently in charge of the Assessment Department of Cursos Internacionales de la Universidad de Salamanca. He represents the University of Salamanca in The Association of Language Testers in Europe (ALTE). He is highly experienced in running teacher training courses and seminars at Spanish and international universities and institutions. He is responsible for the teams that work on the development and marking of the Diplomas de Español como Lengua Extranjera (DELE) and the Servicio Internacional de Evaluación de la Lengua Española (SIELE). Prieto Hernández is the author of many specialised articles on the methodology of Spanish teaching and the assessment of Spanish as a foreign language. Additionally, he has contributed his expertise to several teaching materials as a co-author.

## Assessing speaking – the challenge of eliciting authentic performance

As teachers or language professionals we are required to assess students' speaking competence, but how do we get students to show us what they can do? This presentation explores how to elicit authentic speaking performance from students, and considers why this is an essential element not only of valid assessment of communicative competence, but also of best teaching practice. It is widely accepted that good assessment practices can enhance learning outcomes in second language study and help students fulfil their potential. Key to these practices is placing the test taker at the centre of the testing system; and considering the local contextual factors, cognitive demands and candidate's individual resources within the construct of the assessment. For Trinity tests this is achieved through performance-based assessment based on elicited, personalised communication subject to scrutiny through the application of rating criterion. This presentation looks at the practical challenges of eliciting real world spoken performance for assessment whilst keeping the test taker at the centre of the communication, and relates this back to the classroom, asking how teaching practices can help facilitate authentic spoken production. The presentation refers to elicitation techniques employed by examiners attempting to elicit specific language functions listed in the Trinity GESE specifications for Grade 6 and reflects on how effective these techniques can be and how they influence the student's performance. Consideration is then given to how relevant these techniques are to teaching practice and how teachers can use them in everyday practice. The presentation combines this with analysis of examiner and candidate production taken from the Trinity Lancaster Learner Corpus. The objective is to offer a practical toolkit of techniques for elicitation of authentic interaction for teaching and in turn for speaking assessment, offering general guidelines applicable to both informal classroom applications through to high-stakes summative language exam delivery.

**Alex Thorp**, Trinity College London, United Kingdom

Alex Thorp is the academic support specialist at Trinity College London and previously had over 10 years experience heading a teacher training department. With special interests in assessment, neuro-linguistics and evidence-based teaching, he writes and presents on supporting best teaching practices.

**Cathy Taylor**, Trinity College London, United Kingdom

Cathy Taylor is currently Head of examiner panel development and standardisation at Trinity College London. Cathy has worked in language testing for 10 years and before that was a teacher trainer and teacher for over 20 years. She has been working on the Trinity Lancaster Learner Corpus and is investigating examiner elicitation strategies.

## Assessing the pragmatic competence of EFL learners at B2C2 levels

As the number of overseas students and employees in English-speaking countries has exponentially increased over the last decades, the importance of pragmatic competence for L2 speakers to be successful in their studies and at work has been highlighted and the need for assessing it has become more pressing (e.g. Ross and Kasper 2013). Most currently available pragmatic tests use the Speech Act Theory as a theoretical framework, however, it has been criticised lately for overlooking the importance of the discursive side of pragmatics, which requires the use of on-line processing skills. This research, therefore, investigated an approach to assessing B2-C2 level learners' pragmatic competence in oral discourse. It aimed to identify criterial features defining the level of EFL learners' pragmatic competence and to examine the extent to which different speaking task formats allow test-takers to display their pragmatic competence. Thirty university students participated in the study, which included four monologic and two dialogic tasks, followed by a semi-structured interview. Performance on the tasks was video-recorded, transcribed and analysed quantitatively as well as qualitatively using a Conversation Analytic framework. The results indicate that with increasing language competence EFL learners not only used more but also a wider range of pragmalinguistic devices. It was also found that very high level learners tended to adjust the number and type of modifiers used, somewhat more appropriately taking the given power constellation into consideration. The presentation will conclude with two suggestions. First, tasks and criterial features that test designers and EFL teachers can use in their assessment of students' pragmatic competence. Secondly, how pragmatic competence might influence future prospects of employability in an English-speaking country.

**Edit Ficzero Willcox**, Oxford Brookes University, United Kingdom

Edit Ficzero Willcox obtained her MA from King's College London and is currently conducting her PhD research at CRELLA, University of Bedfordshire. She is an Associate Lecturer at Oxford Brookes University and has previously worked as Director of Studies, teacher trainer and teacher in several countries. Her publications include an article on the usefulness of teaching politeness strategies in English language classrooms' in the Baltic Journal of English Language, Literature and Culture, and a book review on Assessing Second Language Pragmatics.

## Assessing young learners speaking skills in primary education

In commission of the Board of Education of the Netherlands, we conducted research assessing the oral communicative competence of 11-year-old pupils in the Netherlands (8th grade). Realising that oral communication in primary education is generally a neglected area, we decided that we should not copy current educational practices with regard to the development of

oral linguistic competence, but instead develop innovative but authentic oral tasks, both for speaking and for interaction, as a tool to assess their competence in this area. This was a potential risk, because pupils may not be used to being assessed in this way, but there was also the potential to positively affect the development of curricula with regards to oral competence in primary education. In this presentation we will first present how we developed these tools, whilst we tried both to serve the research goals as defined by the Board of Education (assessing the current linguistic oral competences of pupils) and to provide educational settings with innovative means of developing speaking and interaction skills.

In the second part of this presentation, we will more thoroughly discuss how pupils' performance on these tasks should be assessed. In addition, the development of the assessment categories was driven both by the requirement to develop valid and comprehensive assessment criteria, and by the desire to deliver the right (i.e. a more valid and authentic) focus for developing oral skills.

To illustrate our arguments, we will show video recordings of pupils' performances from our research both in interactional and in monologue settings.

**Mirna Pit**, Bureau ICE, Netherlands

Mirna Pit is Head of Test Development and Research at Bureau ICE in the Netherlands.

## Assessment challenges in CLIL: new perspectives in teacher education

Content and language integrated learning (CLIL) has emerged as a major innovation for improving both language and content learning. But CLIL requires a reconceptualisation of language from language learning per se towards an integrated model which actively involves the learner in using and developing language of learning, language for learning and language through learning. This approach has raised several questions concerning what to assess and how in a CLIL based approach. Recent projects aimed at training teachers of subject matters in teaching content through English have witnessed an unprecedented number of teachers participating in order to meet a new and engaging challenge. The acknowledged role of language as a means to sustain cognitive reasoning as well as content learning has been the underlying notion of most projects developed by new groups of language teachers and learners where assessment was central in the project design. Emerging communities of practice of teachers of different subject matters are learning to collaborate and explore implications of using another language to widen the borders of their subject matter. Discovering the potential role of language in teaching and learning (linguaging) as well as the key role of dynamic assessment in CLIL, represent the emerging issues and a shift in perspective in teacher education and curriculum development. This paper will present findings and implications of reflective practice adopted in CLIL classrooms within the module devoted to assessment and evaluation.

**Lucilla Lopriore**, Roma Tre University, Italy

Lucilla has a PhD in Spoken Assessment in Italian Certifications from the University for Foreigners of Siena and an MA TEFL from Reading University. She is a scientific committee member and responsible for CERT IT (University of Roma Tre, CLIQ). She is also an Associate Professor, Foreign Languages Department at Roma Tre University, a teacher and teacher educator for pre- and in-service second language teachers and for CLIL teachers. She has published extensively in teaching methodology, and is interested in Italian L2, assessment, CLIL and ELF.

## Assessment in the future: A Cognitive Diagnostic Modelling for college english reading test

Diagnostic language assessment (DLA) has recently gained much attention from teachers, language testers and second language acquisition researchers. It seeks to promote further learning designed to address the learners' weaknesses and increase their overall growth potential with "learning-oriented assessment" as its rationale. With the empowerment of Cognitive Diagnostic Approach (CDA), the latest theory in psychometrics, DLA is possible to be achieved not only theoretically but also methodologically and practically. However, the results of most previous CDA-based research have classified learners' mastery of a set of tested skills into a dichotomous-score pattern (0/1 pattern), lacking accuracy in that the critical value of mastery is vague. Aiming to provide students with finer-grained diagnostic feedback, this paper addresses how a College English reading diagnostic assessment model is constructed via the combination of CDA and other statistical approaches. The study starts with building a hypothetical diagnostic model including cognitive reading attributes, diagnostic test design, and Q-matrix models, followed by the validation of reading attributes via regression analysis and Q-matrix via Fleiss Kappa. Then R software is adopted to generate Group-level and Individual-level mastery probability of each verified reading attribute and LOWESS approach is applied to gain a group-level mastery tendency model. Finally, all the diagnostic information is synthesised into a well-designed diagnostic feedback including, 1. Learners' relative placement in group-level diagnostic result; 2. The individual reading attribute mastery probabilities; 3. The interpretation of learners' performance in the diagnostic test. The findings are significant in achieving the goal of personalised assessment so as to fulfill the purpose of "tailored" L2 learning through DLA framework: integration of diagnosis, feedback, and remedy.

**Du Wenbo**, Xi'an Jiaotong University, China

Du Wenbo obtained his MA from School of Foreign Studies at Xi'an Jiaotong University, and is currently pursuing his PhD in the field of Language Assessment and Language Acquisition. He takes an interest in Diagnostic Language Assessment (DLA) via Cognitive Diagnostic Approach (CDA) in psychometrics and has attended several International Conferences relevant to Language testing and assessment, such as Pacific-Rim Objective Measurement Symposium (PROMS 2016), and Pacific Second Language Research Forum (PacSLRF 2016) in Tokyo.

**Ma Xiaomei**, Xi'an Jiaotong University, China

Ma Xiaomei is a professor and PhD supervisor of Applied Linguistics in School of Foreign Studies at Xi'an Jiaotong University. She also works as Executive director of the China Institute of Educational Linguistics. She devotes herself to the field of personalized English learning, language assessment, and online diagnostic assessment. She has accomplished twelve research projects, published nearly twenty research papers, and has given speeches in international or domestic conferences nine times, including one keynote speech.

## Assessment literacy and language teachers? A case study with teachers of Portuguese as Foreign Language

There is a conceptual frame in what regards assessment, within and outside classroom, in order to provide high quality standards (relevant and accurate evaluation). Assessment cannot be separated from its object: what is assessed and those who are involved in assessment process. What we propose is the analysis on ideas and representations that teachers of Portuguese as a Foreign Language (PFL) have about assessment. Bearing in mind the existence of differences in the curricula, contexts, learners and teachers' previous knowledge, we shall present the results of one survey applied to teachers of PFL that focuses on major topics such as: understanding what they consider of major importance in assessment; what do they evaluate and how do they do it when they are running a PFL course. The ideas and representations of language teachers determine how the assessment has an impact on classes. Moreover, we will find also impact in learners, especially, by the transfer process from teachers to them, affecting the way they engage in assessment process of their linguistic competences. Our study is based on different contexts – China (Macau) and Portugal, so we can see if the cultural and educative contexts are important in this relation. We shall also discuss the need for developing flexible attitudes and confidence regarding assessment in the teaching and learning process as well as in its articulation with assessment systems that lead to international certificates, such as the case of CAPLE examinations (ALTE). To reach this, the teachers' training in language testing is critical. Besides, the assessment awareness could contribute significantly to foster analysis of the most visible problems of students. So, discussion in assessment literacy of teachers is fundamental.

**Catarina Isabel Sousa Gaspar**, University of Lisboa, Portugal

Catarina Gaspar is Assistant Professor in the University of Lisbon, where she has taught since 2000. She holds a PhD in Latin Linguistics (2009). Her research interests include: romance languages, disclosure policy and language policy. Between 2009 and 2013, she worked as a researcher at the CAPLE. She is Director of the Master in Portuguese as FL/L2 and member of the Scientific Commission of the Program in Portuguese as Foreign Language/Second Language.

**Maria José dos Reis Grosso**, University of Macau/University of Lisboa, Portugal

Maria José dos Reis Grosso is Associate Professor in the University of Lisbon, and since 2012, she is Associate Professor in the University of Macau. She was Director of the CAPLE (2009-2012). She holds a PhD in Applied Linguistics (2000). Her areas of study and teaching focus on Applied Linguistics to the Teaching of Portuguese (FL/L2) and benchmarks in the areas connected to it, including teaching, learning, assessment (FL/L2), disclosure policy and language teaching and intercultural studies.

## Assessment literacy in college learners of EFL Writing

This study tries to sketch a learners' language assessment literacy (LAL) profile through the examination of learners' assessment feedback. In a college English as a Foreign Language (EFL) writing course, with the aim of enabling learners to assess and improve their own writing, fifteen students were instructed to examine writing samples through observing assessment and

feedback modeled by the teacher. They then practiced the same thing with peers on their own work.

Data came from written feedback students provided to each other on Moodle discussion forums, follow-up author revision reports, essays before and after each round of assessment, and end-of-term individual stimulated recall interviews. Unlike previous studies focusing mostly on feedback from the perspective of a language writing teacher, this research examined learner feedback data for its impact on and characteristics relevant to learning. Feedback data were analysed using a grounded-theory approach, with essays and other data serving the purpose of triangulation.

Results were synthesised in a learner LAL framework involving both cognitive and affective dimensions. The cognitive dimension included four aspects: 1) identifying the gap between performance and target; 2) elaborating on the identified gap with learned knowledge in the subject matter; 3) bridging the gap or attempts to do so; and 4) being resourceful while trying to bridge the gap. Data in the third aspect encompassed a spectrum of knowledge and competence, representing different levels of LAL in moving from current performance to the learning goal. In the affective dimension, two aspects included the learner assessor being encouraging and skeptical. The former was a way to show acknowledgement on achievement, despite imperfection in performance. The latter exemplified an educated awareness that assessment and feedback may be subject to more scrutiny. In conclusion, learners' LAL may be very distinctive from other stakeholders' LAL. A comprehensive LAL model should take classroom learners into consideration.

**Shu-Chen Huang**, National Chengchi University, Taiwan

Shu-Chen Huang is a professor and director at the Foreign Language Center at the National Chengchi University in Taiwan. She has taught English as a foreign language at the tertiary level in Taiwan for nearly 20 years. Her research interests lie in learning-oriented assessment and EFL learning motivation. She has published in Language Assessment Quarterly, Language Testing, and other professional journals.

## Can a test of structural reasoning help to predict language outcomes?

Language learners are usually guided to language courses according to their level. However, in Finland different kinds of study paths described in national curriculum are also offered for adult migrants. Along a basic study path there are paths for faster and slower learners as well as for illiterate ones. To select the best study path, study skills are tested in initial testing. One of these tests is a test of structural reasoning, created in Finland in 2008. The test of structural reasoning consists of three tasks measuring the speed and accuracy of recognising a given word form, the ability to analyse structural similarities and differences, and the ability to use analogies. The test does not require any Finnish skill. It aims to produce information about test takers' competences, so that a suitable study path can be selected, thus speeding up their entry into the labour market. The test of a structural reasoning has been used in initial testing in Testipiste, an assessment center for adult migrants' language skills, since 2010 with some 3900

test takers per year. This study aims to find out how useful and reliable the test of structural reasoning really is. Can it help to predict language outcomes in the end of one year language training? Do language outcomes correlate with the results of structural reasoning? This research combines data from both initial testing before language training and proficiency tests after training. It is already known from experience that language outcomes vary according to the type of study path. If a connection between test results and language outcomes is found, then students with weaker test results should perhaps be provided with longer language training despite of limited resources which are due to the current amount of asylum seekers.

**Elina Stordell**, Testipiste, Finland

Elina Stordell is a language tester in Testipiste, an assessment centre for adult migrants' language skills, since 2010. She is interested in listening comprehension, study skills and statistical analyses.

## **Certification of Proficiency in Polish as a foreign language and its influence over the Polish labour market**

The system of certification in Poland is quite new if we compare it to other European countries. Till last year only three levels of exams were available; now there are all levels mentioned in the Common European Framework of Reference for Languages (CEFR) from A1 to C2, including exams for children. This upward trend shows that the position of certificates of Polish as a foreign language is growing and this field needs to be explored. On the other hand, it should be checked how employers are perceiving that kind of language skills verification. The main aim of my speech will be presenting how it works in companies, which are employing non-polish speakers. Is it really important in the recruitment process? Is there any encouragement for foreigners to improve their language skills? Is it necessary to think about certificates with specialisation (business or medical Polish)? I will try to answer at length to those questions mentioning the research which I did during my master degree in different companies. The challenge is to popularise the certificate as the objective way to verify the level of language knowledge and skills. This type of language assesment will definitely support learning in the 21st century, because the world is becoming more and more multilingual.

**Dominika Bartosik**, Jagiellonian University Kraków, Poland

Dominika Bartosik obtained her Master's degree in 'Advertising Management' and is about to finish her second Master's degree in 'Teaching Polish as a foreign language'. Her thesis is focused on Certification of Proficiency in Polish as a foreign language and its influence over the Polish labour market.

## **Certifications: Tools for a policy of educational cooperation and to accompany language learning. The example of French in Italy**

The CIEP, together with the French diplomatic network, accesses key sociological, geographical and economic data in order better to understand the integration of certifications in local educational systems. Initiatives carried out in Italy by our colleagues in the French Embassy in Rome as well as by the CIEP are an emblematic example. The DELF qualifications in French,

for young candidates at school, have met with growing success in a certain number of countries of the world, especially in Europe. Adolescents make up over 63% of the total of some 410,000 candidates a year who take the DELF and DALF exams, which was not the case ten years ago, before the harmonisation of exams worked out by ALTE members based on the Common European Framework of Reference for Languages (CEFR). There are about 40,000 of these candidates in Italy. The DELF qualification has been integrated as a tool for promotion of the French language in the policies of educational cooperation pursued by French embassies with local educational authorities. We have in this way been able to note that numerous ministries of education have become interested in foreign language qualifications issued by certifying bodies not directly under their own authority but from other countries. The case of the introduction of DELF into the Italian educational system (just as in Germany, Spain, Greece or Mexico) shows particularly the wish of the countries, even in a sometimes difficult economic climate, to integrate foreign language certification into their educational policy. The process of validation of test content in the development of our exams also appears to reinforce the logic of such an introduction into educational systems. In fact the quality standards and exam development processes sanctioned by ALTE seem, according to the evidence, to contribute to this success.

This presentation will be drawn up in two languages: English and Italian (English speaker with Italian slides and Italian speaker with English slides).

**Virginie Salles**, Institut français d'Italie/Ambassade de France en Italie, Italy

Virginie Salles è incaricata della gestione centrale DELF-DALF in Italia presso l'Institut français Italia, ente culturale dell'ambasciata di Francia. In quanto tale, cura la promozione e la diffusione delle certificazioni ufficiali DELF-DALF sul territorio italiano, coordina e gestisce la rete dei centri accreditati. E' anche formatrice di esaminatori-correttori. Di conseguenza, si interessa alle questioni di valutazione delle competenze delle lingue. Ha condiviso le sue riflessioni sulle strategie per gli insegnanti e i candidati che si avvicinano alla certificazione DELF, attraverso diversi manuali per la preparazione di questi diplomi.

**Lucile Chapiro**, CIEP, France

Lucile Chapiro has been working at the Department of Evaluation and Certifications at the CIEP since 2010. She is responsible for the management of the DELF for school exam. She also trains raters, examiners and item writers for the DELF and DALF exams. She is co-author of the French textbook Soda (Clé International, 2012) as well as two training books for the DELF for school exams: ABC DELF Junior Scolaire A1 and A2 (Clé International, 2012-2013).

## **Certificazione PLIDA. Alcune riflessioni su valutazione e testing per gli apprendenti provenienti da lingue distanti e in particolare sinofoni**

Recentemente, sulla base dei dati raccolti durante gli ultimi sei anni di attività del PLIDA e dei feedback provenienti dai Centri certificatori è stata avviata una revisione del programma della certificazione attraverso la riscrittura del syllabo, la riformulazione dei formati d'esame e dei relativi criteri di valutazione. Il processo di revisione del syllabo e dei formati ha lo scopo di avvicinarsi a una migliore applicazione delle linee guida indicate nel QCER e di mettere a frutto le più recenti riflessioni sul testing linguistico. Tale processo ha fatto emergere interessanti temi di ricerca legati ai risultati conseguiti da studenti parlanti lingue distanti, in particolare sinofoni.

Sulla base dei feedback raccolti, i risultati di questi apprendenti nei test linguistici evidenziano uno svantaggio rispetto ad apprendenti provenienti da lingue più vicine. Sulla base dell'analisi delle prove e degli studi disponibili (Rastelli 2015) tale svantaggio non è sempre imputabile alla distanza tra lingua madre e lingua target ma piuttosto ad abitudini cognitive, stili di apprendimento, mancanza di familiarità con schemi di lettura (reading pattern) e ascolto presenti nei test. L'ipotesi prende in considerazione, in particolare, la necessità di revisionare il formato d'esame sulla base della distinzione tra grammatica e processing. È necessario infatti, guardare oltre la superficie della performance e dei prodotti linguistici poiché lo studente sinofono non è in grado di rendere in lingua le sue intuizioni prima di aver accordato all'italiano i suoi strumenti cognitivi (espressivi, relazionali, comunicativi) e le sue capacità di processare suoni e forme grafiche molto diverse da quelle della L1. Il contributo intende presentare piste di riflessione e ricerca per la formulazione di sillabi e test utili ad accrescere la validità e l'affidabilità degli esami e l'equità degli stessi presso un pubblico di sinofoni.

**Silvia Giugni**, Società Dante Alighieri, Italy

Silvia Guigni is the supervisor of the Progetto Lingua Italiana Dante Alighieri (PLIDA). She has gained long-standing and diversified experience in language teaching as L2, in language teachers' training and in curriculum development. She taught Italian Language to foreigners in many contexts in Italy and abroad, for instance through the Bundessprachenamt in Germany, the Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI) in Italy, and the University of Roma Tre in Italy). She was co-author and supervisor of the Piano dei corsi ADA (the teaching syllabus of SDA).

**Barbara D'Annunzio**, Società Dante Alighieri, Italy

Barbara D'Annunzio is in charge of the ADA project at the Società Dante Alighieri in Rome. In 2010 she received her PhD in Linguistics and Modern Philology (course of studies: didactic of languages) from University Ca' Foscari in Venice. She has been teaching professor at University Ca' Foscari in Venice, Ca' Foscari-Harvard Summer School, University of Catania, where she taught didactics of foreign modern languages.

## Combining the reliability of judgement with the validity of external alignment to create a powerful tool for teacher led assessment

Comparative judgement (CJ) is a method of judging pairs of student responses to select one as being better than another. Using multiple judgements for the responses from a group of students, it is possible to establish a reliable rank of the responses from the strongest to the weakest. This is a well-established method of rating students' work that predates the use of marking criteria but was brought into the realm of psychometrics by Thurston's law of comparative judgement in 1927. More recently, the work of Pollitt and Murray (1994), Bramley (2007), and Kimble and Pollitt (2008) have brought the power of comparative judgement back to prominence. However, whilst it is possible to produce a reliable rank order using multiple judgements, this can take a long time as judgements need to be compared in a large number of combinations. Another issue of CJ is that the rankings produced are independent of external frameworks, such as the CEFR, so a teacher using this method is unable to easily benchmark the individual pieces of work they have ranked, or establish whether their rating of work is harsher or more lenient than established norms. This presentation addresses these issues,

drawing upon the piloting by Oxford University Press of Adaptive Comparative Judgement (ACJ) with 181 students from six countries, using 35 judges from a variety of backgrounds. The talk will present the data produced in the ACJ pilot and the methodology applied to reduce the number of judgements required to obtain reliable scores, and how this was combined with the seeding of pre-calibrated responses to align ranked scripts to benchmarked scores. It will also demonstrate how classroom teachers can obtain the same reliability as trained assessors using the ACJ system, making it the ideal tool for valid and reliable classroom assessment of productive skills.

**Ed Hackett**, Oxford University Press, United Kingdom

Ed Hackett is the Head of Assessment Research in the English Language Teaching division at Oxford University Press (OUP). At OUP he is responsible for the design and validation of the OUP ELT English language tests. Prior to joining OUP, Ed worked at Cambridge English Language Assessment for eleven years, following a fourteen year career as an EFL teacher and teacher trainer in the UK and abroad. Ed holds an MSc in TESOL/Educational Management from Aston University.

## Comparing L2-speech perception across different L1-backgrounds: Intelligibility and the Common European Framework of Reference for Languages (CEFR)

The CEFR was developed to make L2 assessment in Europe comparable. However, little research so far has focused on two key elements of language learning, intelligibility and comprehension, within this framework. This is particularly relevant for L2 speech; however, speech-related CEFR sections and relevant level descriptions are vague and none of the assumptions made therein (such as the supposed linear progression between levels) have been sufficiently tested yet. In addition, many of the perceptive skills mentioned in the CEFR are not linked to proficiency levels. I will present a study which examines intelligibility and comprehension as key CEFR elements to investigate: 1) whether listeners' levels of L2 proficiency influence how they evaluate the intelligibility of L2 accents; 2) what the differences in actual comprehension are between learners at different CEFR levels; 3) what the differences in actual comprehension are between learners with different L1s. In this study, participants are Spanish and German learners of English from each of the six CEFR proficiency levels: A1 beginner to C2 mastery level. We tested their 1) perception of intelligibility and 2) actual comprehension of Spanish and German L2-accented English speech. The participants evaluated the intelligibility on Likert scales; comprehension was tested through transcriptions of the speech stimuli. The results reveal that proficiency levels significantly influence evaluation of L2 speech, though not as anticipated. The transcriptions indicate very specific structures in L2 accents that inhibit intelligibility; some of these structures (e.g. variation in vowel length) affect comprehensibility for learners from both L1 backgrounds while others (e.g. variations stress patterns) seem to affect Spanish learners more than German learners of English. Overall the results partially confirm a linear progression as proposed in the CEFR; some structures inhibit intelligibility for learners at lower levels of proficiency only, while some structures are relevant across all levels of proficiency.

**Bettina Beinhoff**, Anglia Ruskin University, Cambridge, United Kingdom

Dr Bettina Beinhoff is a Senior Lecturer in Applied Linguistics at Anglia Ruskin University, Cambridge (UK). She obtained her PhD in Applied Linguistics from the University of Cambridge. Her main research interests are in the areas of L2 speech acquisition and sociolinguistic aspects of L2 accents. She has published widely on L2 speech and has authored the book "Perceiving Identity through Accent – Attitudes towards Non-Native Speakers and their Accents in English" which was published in 2013 by Peter Lang, Oxford.

## Comparing speaking performances across tests and languages: Evaluating the success of an institutional rater training programme

In a large-scale project of proficiency assessment at a Midwestern US University, standardised, computer-based oral proficiency interviews (OPI-C) produced by Language Testing International (LTI) for the American Council on the Teaching of Foreign Languages (ACTFL) were administered to a large number of foreign language students. A subset of these students also took the Simulated Oral Proficiency Interview (SOPI) which is used as an institutional test informing students whether they meet the language requirement of their study programme or not. From 23 students of Chinese, 27 students of French, 26 students of Spanish and 11 students of German, two ratings (SOPI and OPI-C) were available. The official ACTFL rating was performed by LTI raters. The scoring of the SOPIs used the same ACTFL proficiency scale and was conducted by faculty members who were not trained ACTFL raters and had no previous experience in rating speaking performances following the ACTFL proficiency scale. These novice raters were trained in a two-day workshop following a classic rater training design of familiarisation with the ACTFL proficiency guidelines, standardisation or norming, and qualification. Attending the rater training in late fall 2016, these faculty members rated the speaking performances over the winter break. Each rater was asked to justify his or her score on each of the 15 speaking tasks in the SOPI. In this work-in-progress the presenters evaluate the success of the in-house rater training in different languages by comparing the ratings the students received on both speaking tests. Differences in rater agreement will be examined by analysing the justifications provided by the raters as well as the discussions during the standardisation phase of the rater training, and by linking these findings to the ACTFL guidelines and performance descriptors. The findings should allow for an adjustment of the in-house rater training.

**Koen Van Gorp**, Michigan State University, United States

Koen Van Gorp received his PhD in 2010 from the KU Leuven and is currently Head of Foreign Language Assessment at the Center for Language Teaching Advancement (CeLTA) at Michigan State University. From 2010 till 2015 he was Director of the Certificate Dutch as a Foreign Language (CNaVT), the international test of Dutch as a foreign language. His research interests are task-based language teaching and assessment.

**Daniel Reed**, Michigan State University, United States

Daniel Reed is the Testing Director for Michigan State University's English Language Center. He oversees test administration as well as research and test development in support of MSU's English exams. He has played major roles in the development of proficiency tests for more than half-a-dozen languages and has conducted research on language aptitude and the assessment of speaking proficiency. In addition, he has worked with English Language Centre Director, Dr Susan Gass, to develop a large-scale English language testing program in Greece.

**Susan Gass**, Michigan State University, United States

Susan Gass received her PhD from Indiana University. She has published widely in the field of second language acquisition (more than 30 books and 100 articles) and is the Co-Principal Investigator (with Paula Winke) on a 3-year Proficiency grant awarded by the Flagship programme of the US Department of Defense. She has served as President of the American Association for Applied Linguistics and of the International Association of Applied Linguistics (AILA). She has won many local, national, and international awards.

## Constraining issues in face-to-face and Internet-based language testing

Despite the increasing importance of language testing (Chapelle and Voss 2016), technology design matters related to testing have hardly been addressed in academic research. Recent work in this area has looked at the use of devices such as mobile phones (García Laborda, Magal Royo, Litzler and Giménez López 2015) or tablet PCs (García Laborda, Magal Royo and Bakieva 2016), yet little has been written on test-taking anxiety related to computerised speaking tests (Lai, 2010). Although students nowadays are digital natives, in our experience they seem to experience a significant level of anxiety when facing Internet-based oral tests. This research addresses the contrast between face-to-face oral tests and an online oral test. Forty-six students at Universidad de Alcalá responded to a questionnaire about their feelings after taking these two types of tests with similar tasks and prompts. Although positive responses were expected with regard to the online version, the students still reported that they were more at ease during the face-to-face type of interviews. The results of this research can shed light on what needs to be revised in computer or internet speaking tasks in high-stakes language tests.

**Jesús García Laborda**, Universidad de Alcala, Spain

Jesús García Laborda obtained his PhD in English Philology (2000) and a second EdD (2010) from the same university. He has specialised in computer-assisted language learning and technology-based applications for language testing. He currently lectures at Universidad de Alcalá where he is also the head of the Modern Philology department. He has published in Language learning & Technology, Educational Technology & Society, British Journal of Educational Technology, the book series Studies in Language Testing, and others.

**Mary Frances Litzler**, Universidad de Alcala, Spain

Mary Frances Litzler, PhD, has thirty years of professional experience as a teacher of English language, linguistics and translation and as a translator of Spanish, French and Catalan into English. In recent years, she has worked for the British Council and Universidad de Alcalá but she has also worked for the University of Southern California and Temple University Japan. Her major research is in the areas of learning logs and out-of-classroom learning as well as technology for language learning and language testing, but she has also been involved with medieval text editing. She has published papers in the book series Studies in Language Testing, and a number of journals including Onomázein, Educational Technology and Society, among others.

## Construction of CDA-based dynamic intervention for EFL listening test

Cognitive Diagnostic Assessment (CDA) is superior to traditional language testing in its ability of uncovering knowledge structures and processing skills underlying learner performances to provide diagnostic feedback about learners' cognitive strengths and weaknesses. Dynamic Assessment (DA), by offering prompts as a form of mediation when individuals encounter difficulties and tracking their responses, can uncover abilities fully developed and those in the process of developing, to simultaneously assessing and promoting development. However,

unlike CDA, the test items in DA don't examine certain attributes, thus the mediation offered doesn't target the cognitive attributes as well. Consequently, the remedial instructions provided may not be fine-grained enough to pinpoint the specific weak areas to be worked on. Therefore, how CDA and DA can be integrated so as to achieve a fine-grained diagnosis as well as dynamic intervention is a challenging issue in language assessment. This study aims to construct attribute-based mediation based on the modified cognitive diagnostic listening test developed by the research team in Xi'an Jiaotong University. Both qualitative and quantitative approaches are employed. Bilog, Fleiss Kappa, and R software are used to ensure the quality of the test items, meanwhile determine the attributes to be intervened. Then the attribute-based mediation is designed through meta-analysis, expert judgment, and students' think-aloud protocols. The test is conducted in a DA format and a fine-grained learner profile generated through Cognitive Diagnostic Model with Learning Potential Score and Mediated Score are provided. Finally, interviews are conducted to examine the effectiveness of the designed mediation. This study can open up a new perspective for researchers to design mediating prompts. And its ultimate purpose is to provide a preliminary theoretical framework for the development of an online Cognitive Diagnostic language Assessment with dynamic intervention.

**Yihe Yan**, Xi'an Jiaotong University, China

Yihe Yan is a postgraduate student of Applied Linguistics in School of Foreign Studies at Xi'an Jiaotong University. Her research directions are language assessment and online diagnostic assessment. She has attended one domestic meeting and published two research papers. With a great passion for linguistics, she dedicated herself to doing research and she also has been honoured as an "Excellent graduate student" and won professional scholarships.

**Xiaomei Ma**, Xi'an Jiaotong University, China

Xiaomei Ma is a professor and PhD supervisor of Applied Linguistics in School of Foreign Studies at Xi'an Jiaotong University. She also works as Executive Director of the China Institute of Educational Linguistics. She devotes herself to the field of Personalized English Learning, Language assessment, and Online diagnostic assessment. She has accomplished 12 research projects, published nearly 20 research papers, and has given speeches in International or Domestic Conferences nine times, including one keynote speech.

## Cut scores for combined constructs

For many examinations it is accepted practice to report scores for a combination of skills (eg., reading and listening, or speaking and listening). However, the performance descriptors against which the standard is to be set – in many cases, descriptors from the CEFR, or descriptors derived from the CEFR usually refer to each skill on its own, or if they address interaction, limit themselves to the receptive skills relevant for interaction, and do not cover other aspects of the receptive skill in question. Therefore, a way to define the relationship and relative importance of each skill for the combined result has to be found. Determining a cut score for a combination of skills requires a special standard setting strategy (Kaftandjieva 2010). The presentation will demonstrate how this was operationalised for a new examination in German where cut scores were to be set for listening-plus-speaking and reading-plus-writing, on the basis of the work of c. 200 candidates from pretesting. Typical problems will be outlined and discussed.

**Beate Zeidler**, telc gGmbH, Germany

Beate Zeidler is senior validation officer at telc gGmbH, a not-for-profit company offering language examinations, owned by Deutscher Volkshochschul-Verband (DVV), the largest adult education association in Germany. She has been working full-time for DVV since 1991, after being an English teacher and doing freelance work. She has taught courses on language learning and testing at various educational institutions in Germany, and has been involved in developing examinations such as the Deutsch-Test für Zuwanderer (German for immigrants).

## Designing a principled approach for rater training and norming protocols: Integrating theory and practice

It is widely recognised that both initial rater training and subsequent norming before actual rating occasions are crucially important to the validity and success of operational testing programmes and research programmes that employ judges. However, the theoretical and empirical underpinnings of particular rater-training designs and the details of related activities (e.g., norming or recalibration) have not been clearly established in the fields of language testing and second language studies. With the goal of developing a principled approach to the design of rater training and norming protocols (materials and procedures), the testing team in an English programme for international students at a major mid-western university undertook a series of steps. The first step was to conduct a literature review in order to survey what current practices have been documented. The investigators were interested in both theory and practices and so searched both scholarly journals as well as test manuals and different codes of practice. Next we conducted a survey and follow-up interviews with key individuals involved in rater training for major test providers. In addition, the investigators examined the details of the rater training and norming protocols for the English exams at their own institution as well as data from the administration of English exams administered in Europe. Research questions included whether current training practices resulted in acceptable levels of agreement, whether agreement was stronger at particular scale levels or for particular test levels (e.g., a B2 versus C2 essay exam). The investigation culminated in a framework for guiding research and practice, with indications of parts of the framework that needed additional work from a principled or theoretical basis as well as additional backing data. The results of this work will be presented along with a recommendation for allotting an increased role for e-learning systems in training examiners and raters.

**Daniel J Reed**, Michigan State University, United States

Daniel J Reed is the Testing Director for Michigan State University's English Language Center. He oversees test administration as well as research and test development in support of the university's English exams. He has played major roles in the development of proficiency tests for more than half-a-dozen languages and has conducted research on language aptitude and the assessment of speaking proficiency. In addition, he has worked with ELC Director, Dr Susan Gass, to develop a large-scale English language testing programme in Greece.

**Heekyoung Kim**, Michigan State University, United States

Heekyoung Kim received her Ph.D. from Texas A&M University and currently works as Assessment Specialist in the Testing Office of the English Language Center at Michigan State University. She co-ordinates development of test materials, manages test item banking and planning, oversees test administrations and scoring/reporting test

results. Her interests include statistical analysis and validation of tests. Recently, she extended her interest to applying e-learning software to online rater training/norming to make it more efficient, interactive, and engaging.

**Aaron Ohlrogge**, Michigan State University, United States

Aaron Ohlrogge has worked as a test developer for 14 years for Michigan State University and the University of Michigan. He holds a BA in English & Linguistics and an MA-TESOL. When not developing tests, he also teaches ESL and language assessment courses.

## Developing a Japanese language test for a multilingual online assessment system: Towards an action-oriented approach to Japanese instruction in Europe

This paper presents the different steps involved in developing an online placement test of Japanese based on the Common European Framework of Reference for Languages (CEFR). This is a part of a project at Université Grenoble-Alpes called Innovalangues, where one aim is to create a multilingual online formative assessment system covering six languages: Italian, English, Chinese, Japanese, Spanish, and French as a Foreign Language. The Japanese test is scheduled to become operational in September 2017 as a placement test for students from levels A1 to B1. In developing this assessment, our project contributes to transforming Japanese instruction in Europe and making it more action-oriented. Unlike European languages, Japanese does not have a completed CEFR-based referential on which to organise language education and assessment. Therefore, we have had to develop our own syllabus, referring to the CEFR descriptors and a few Japanese syllabi associated with the CEFR (e.g. JF Standard), as well as taking into account the specific characteristics of the Japanese language. Our syllabus has been designed around the communicative situations which Japanese learners in France or in Europe will encounter. We see Japanese learners as “social agents”(Council of Europe 2001:18) performing acts in society, even if these are very simple acts and even if they are beginners (A1). In designing and validating the tasks, we had to make adjustments to existing descriptors, due to characteristics of Japanese for European learners, such as difficulties related to the difference in graphic systems or the honourific language. In this presentation, we will present our approach and highlight how we managed to construct action-oriented tasks to ensure the authenticity of situations, despite the difficulty of using authentic resources.

**Tomoko Higashi**, Université Grenoble Alpes, France

Tomoko Higashi (PhD in linguistics) is a Senior lecturer of Japanese and a member of the Laboratoire de Linguistique et Didactique des Langues Etrangères et Maternelles (LIDILEM) research group at Grenoble-Alpes University (UGA), France. Currently, she is Head of the Japanese department and is responsible for the validity of the Japanese test development team of the Innovalangues project at UGA. Her research interests focus on intercultural pragmatics and interaction analysis. She co-led the Project on Language Activities and Competences of the CEFR B1 level. She has published many articles and textbooks.

**Chieko Shiota**, Université Grenoble Alpes, France

Chieko Shiota (PhD in linguistics) is currently a Pedagogical Engineer on the research project Innovalangues, Université Grenoble-Alpes, France. She has been an associate professor at Saga University, Japan, and a lecturer at Bordeaux Montaigne University, France. She was a member of the Japanese Language Proficiency Test Committee. Her research interests focus on teaching methods for prosody and the analysis of syntax and prosody in different oral communication styles. She has published many articles in Japan and in Europe.

## Developing operational framework and descriptors of pragmatic effectiveness for China's standard of English

The scale of pragmatic effectiveness is an important part of China's Standard of English, but pragmatic effectiveness is also an intricate subject in the field of language proficiency and pragmatics research, not to mention developing a scale of it. Some researchers have defined sociolinguistic competence and pragmatic competence (e.g. Hymes, Bachman and Palmer) which shed some lights on the development of language proficiency scales, for example, the scale of "sociolinguistic appropriateness" is one of CEFR scales based on the above mentioned theory of competence. The researcher tries to develop an operational framework of pragmatic effectiveness based on the theory of language use, namely, pragmatic effectiveness means that a language user can express himself appropriately in the linguistic, situational, and sociocultural context to achieve a certain communicative function and the ideal communication goal. Based on this operational framework, the researcher determines some key words to collect descriptors from the existing proficiency scales and academic papers. A pool of descriptors is set up. Besides, on the basis of the relative researches on writing and speaking, she finds the factors that determine the levels of descriptors. Then the descriptors are edited according to these factors. Finally the descriptors are made into questionnaires, and a nation-wide survey among teachers and students from primary school to universities is conducted. Rasch Analysis is conducted to develop the scale of pragmatic effectiveness.

**Shuhua Wang**, NEEA, China

Shuha Wang is currently the professor in the foreign language department of Beijing Wuzi University. She obtained her PhD in applied linguistics from Beijing Foreign Studies university. Her dissertation is on the development and validation of descriptors of language comprehension ability of Chinese learners of English.

## Developing productive writing tasks that test young learner A1 and A2 level communicative writing abilities

This presentation describes the development approach and choices made to create writing tasks for the Cambridge English: Young Learner (YL) tests that support teachers and schools in teaching integrated and practical reading and writing skills. The YL tests are intended to support strong teaching and learning among young learners, and one way they do so is by including tasks that elicit authentic language usage. For these reasons, during the tests' regular 10-year review, the development of productive writing tasks was identified as a priority for the Movers and Flyers reading/writing papers. Specifically, adding a productive writing task promotes teaching extended productive writing skills, as well as more fully assessing the intended construct of young learner writing. The writing tasks developed are designed to support children to produce extended writing that might be expected at the level assessed (A1 for Movers and low A2 for Flyers). The tasks are the final task on each reading/writing paper. The Movers task asks candidates to produce progressively more independent writing by asking candidates to complete two sentences about a picture, then to write responses to two questions about the same picture,

then write two sentences about the picture. The Flyers task asks candidates to produce a short written narrative by presenting candidates with three pictures forming a sequence of events, and asking children to write 20 to 30 words about the three pictures. The tasks were trialled in five countries and revised, and initial mark schemes developed. The mark schemes were piloted, underwent trialling with 175 samples and 41 markers in three countries, and were revised according to each round of results and marker feedback. Since the purpose of the tasks is to assess ability to write for communicative purposes, the focus of mark scheme development was on successful communication.

**Maggie Dunlop**, Cambridge English Language Assessment, United Kingdom

Maggie Dunlop is a Senior Research Manager at Cambridge English Language Assessment. She researches the Cambridge English: Young Learners tests and learning-oriented language assessment, including integration of assessment, instruction and digital learning. Maggie also specialises in educational program evaluation. Maggie holds an MA in International Education from the George Washington University, DC, USA and a PhD in Language and Literacies Education from the University of Toronto, Canada.

**Kathryn Davies**, Cambridge English Language Assessment, United Kingdom

Kathryn Davies is Assessment Manager for Cambridge English: Young Learners (YL) Reading and Writing at Cambridge English Language Assessment. She has taught young learners in the UK, Spain and Japan, and has been involved in the development of YL material, both as a writer and editor at Cambridge University Press. Kathryn holds an MA in Linguistics and a Cambridge Diploma in Teaching English to Speakers of Other Languages (DELTA).

## Development of consecutive interpreting strategic competence scale

In the multilingual world, interpreting has become an increasingly commonly-used language activity, especially in China, a country where international business and communication are developing rapidly. However, there are few scales describing interpreting competence as a language activity. This study intended to develop a scale for consecutive interpreting strategic competence as part of the China Standards of English under the project of the National Education & Examination Authority (NEEA) in China. By bibliometric analysis of documents, quasi-experiments and immediate retrospective interviews, descriptors for consecutive interpreting strategic competence were collected. The descriptors were revised based on the Pearson principles and results from quasi-experiments and immediate retrospective interviews, in order to compose a preliminary scale. The study then used exploratory factor analysis (EFA) to analyse the items in the scale. It was found that strategic competence of consecutive interpreting is composed of six factors in two dimensions. The scale was validated through structural equation model (SEM) based on the data from 206 MTI interpreting students using the scale and interpreting tasks.

**Yi Xu**, Guangdong University of Foreign Studies, China

Yi Xu obtained her PhD from the Guangdong University of Foreign Studies in 2016. She received her BA in 2010 and an MA in Language Testing in 2013. Xu's research interests include interpreter's competence and interpreting testing and assessment. She is now a key member of several national research projects such as the China Standards of English, Strategic Planning and Policy Management of National Language Service and Test for Translation and Interpreting Major (Chinese/English)-Band 8.

## Diagnostic assessment: Its use in teaching and learning foreign languages

Assessment in many parts of the world is focused on high-stakes standardised examinations which are either based on paper and pencil or computerised. Due to its high-stake nature, validity and reliability of the assessment are one of the major concerns in assessment development and implementation. Despite the original purpose of the assessments, traditional types of assessments are often treated separately from instructions. As such, the necessity to connect the assessment results with instructions is more emphasised. This presentation introduces diagnostic assessment practices adopted at the Defense Language Institute (DLI), USA. Diagnostic assessment at DLI is used to better gauge students' proficiency levels and to provide tailored instructions to foreign language learners. The concept of diagnostic assessment was influenced by dynamic assessment, which is grounded on the mental development theory developed by Russian philosopher, psychologist, and educator Vygotsky. Dynamic assessment is an interactive assessment with the intent of identifying learners' actual development and potential development level. DLI's diagnostic assessment includes personality and motivation level analysis and learning style identification in addition to interactive dynamic assessment between teacher and students. Based on the analysis of the available information and teachers' one-on-one assessment of the learners, teachers diagnose students' strengths and weaknesses and design a tailored learning plan to assist learners' individual needs. Diagnostic assessment can be on-going and is directly related to classroom activity designs and providing individual learner support. The presenter will share her experiences of being a teacher educator training teachers of diagnostic assessment as well as her experience of implementing diagnostic assessment with learners of advanced level Korean.

**Hyunsoo Hur**, Defense Language Institute Foreign Language Center, United States

Hyunsoo Hur is Associate Professor at the Defense Language Institute, USA. She has taught foreign languages and trained foreign language teachers working for the US Government. Her areas of interest include applied linguistics, linguistic anthropology, language acquisition, teacher education and intercultural communication. She has been a presenter at numerous conferences and also published in various journals. She has earned her PhD from the Pennsylvania State University, USA.

## Empowering learners for a demanding labour market: the “Groups for the Experimentation of Plurilingualism” program in Catalonia

Since the second half of the 20th century and as a result of globalisation, our societies have strongly changed. Moreover, the economic crisis caused the loss of thousands of jobs in Europe and revealed a growing gap between the actual skills of workers and the skills the labor market needed. Mastering the technical skills of a job is no longer enough. Nowadays professionals must be able to critically analyse reality, solve problems in a creative way and be used to teamwork. And one of the most valued skills is communicative competence in general and foreign languages proficiency in particular. Education systems across Europe have echoed this reality and Catalonia hasn't been an exception. For this reason, the Department of Education has launched several programmes in order to bring about deep changes in the educational

system. The presentation will explain one of these programmes: the Group for Experimentation of Plurilingualism programme (GEP). Launched in 2013, the programme aims at activating the students' plurilingualism in at least three languages by developing interdisciplinary school projects. Besides the communicative competence, it also reinforces professional skills and promotes lifelong learning. Since its beginning, about 400 Catalan schools have already participated. There are primary and secondary schools and some vocational training centres involved. The programme lasts for two years. During this time schools' principals are guided towards good management and deployment of the programme within the school and teachers are trained. The teachers, who are not language specialists, are trained to teach their subject in a foreign language and to use multimedia and multilingual texts within their ordinary classes. They are also introduced to the project-based learning method to foster teamwork among students, develop their critical thinking and promote their academic and professional growth.

**Montserrat Montagut Montagut**, Government of Catalonia, Department of Education, Spain

Montserrat Montagut Montagut is currently working in the Ministry of Education at the Generalitat de Catalunya. She holds a degree in Catalan Philology and a postgraduate diploma in Terminology and Lexicography. As the head of a Foreign Languages Service, she manages programmes of educational innovation aimed at increasing students competence in foreign languages and at fostering cooperation between teachers. She also supports international schools projects and organises events for the dissemination of plurilingualism. She participated in the Catalan version of the Common European Framework of Reference for Languages (CEFR). She also teaches the Master's in Secondary Education at the University of Barcelona.

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

The influence of standardised English language tests has never been greater, due to increased internationalisation as well as the ever-growing number of English as a Medium of Instruction (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.' This suggests that there are other linguistic skills and competencies beyond language proficiency that need to be gauged, especially for PG entry. So, what are these skills? And how can we best assess them?' 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.

**Dr Karen Ottewell**, University of Cambridge, UK

Karen is the Director of Academic Development & Training for International Students section at the University of Cambridge, which provides training to assist international students in further developing and honing the skills required to succeed in an English-speaking academic context. She did her MA (Hons.) at St. Andrews, and her MPhil and PhD at Cambridge. She has recently started an EdD at the Faculty of Education at Cambridge, investigating the influence of rhetorical transfer on L2 discipline-specific written academic literacy.

## English language learning and assessment needs of economic migrants in the UK

As of the first quarter of 2016, approximately 2.2 million EU born were employed in the UK labour market. Cambridge English Language Assessment investigated what it feels like to be an economic migrant in the UK, and explored their aspirations as well as their English language learning and assessment needs in depth. The research took place from September to December 2016 with economic migrants from Spain, Poland and Romania. As this was shortly after the referendum decision was made for the UK to leave the EU, this research also explored the impact this development was having on these economic migrants' decisions to remain in the UK. The research used an ethnographic approach to gain in-depth, qualitative insights. This approach enabled researchers to speak to people in their natural contexts, thus gaining an understanding of their social spaces and networks. Fieldwork was conducted in two cities in the UK: London and Reading. Many people with limited English language skills felt that their English levels were holding them back both in getting a job, and in progressing in that job. For new arrivals especially, limited English was a big barrier, causing fear at the interview stage. Many economic migrants talked about needing encouragement and support for first interviews.

Whilst many of the people that participated in this study were keen to improve their English, UK education could be difficult to navigate for new economic migrants.

**Dr Sanjana Mehta**, Cambridge English Language Assessment, United Kingdom

Sanjana Mehta is Head of Market & Customer Insight at Cambridge English. She is responsible for conducting commercial research to inform the sales and marketing strategies of the organisation. Sanjana has worked and studied in India, Belgium and the UK. She has a PhD in Education and Psychology from the University of Cambridge. Sanjana is interested in how technology can individualise learning and create more equal opportunities for educational success.

## Exploring teachers' language assessment literacy: A social constructivist approach to understanding effective practice

Jones and Saville (2016) assert that the two key purposes of assessment are to promote learning and to measure and interpret what has been learned. In terms of classroom assessment, this implies that teachers have a central role to play in planning and/or implementing appropriate assessment procedures to monitor and evaluate student progress in their classrooms. But teachers' attitudes and beliefs, based on their own experiences of assessment, exert a powerful role in shaping their decisions, judgements and behaviour (Borg

2006). Exploring teachers' levels of assessment literacy in terms of their own assessment experiences may help teacher educators to better understand the factors which promote or prevent effective assessment, thus contributing to more targeted and empowering teacher education. The research to be discussed in this presentation adopts a social constructivist model of learning and meaning-making, with the language classroom representing the community of practice. Drawing on Davies' (2008) components of assessment literacy Skills, Knowledge and Principles, Phase 1 of the project consisted of interviews with teachers in which they were invited to estimate their understanding of the individual components of the assessment process and indicate how much they would like to learn about each. Classroom observations then took place which focused on teachers' actual assessment practices, followed by post-observation, reflective interviews. Finally, focus group discussions were conducted with further groups of experienced teachers, which confirmed the findings from Phase 1. Four key findings will be presented relating to teachers': 1) previous training in assessment; 2) attitudes to language testing and associated theory; 3) understanding of assessment in its broader sense; 4) understanding of 'language assessment literacy'. It would seem that there are considerable differences in understanding between teachers and those who research and write about language testing/assessment, which may have important implications for the development of future teacher education courses.

**Vivien Berry**, British Council, United Kingdom

Vivien Berry is Senior Researcher in English Language Assessment at the British Council, where she leads an assessment literacy project to promote understanding of basic issues in language assessment, including the development of a series of video animations, with accompanying materials. With external colleagues, she is also investigating the speaking construct in face-to-face and video-conferencing delivered modes of an international speaking test. She has published extensively on many aspects of language assessment and regularly presents research findings at international conferences.

**Susan Sheehan**, University of Huddersfield, United Kingdom

Susan Sheehan is a Senior Lecturer in Teaching English to Speakers of Other Languages (TESOL) at the University of Huddersfield. She teaches both undergraduate and postgraduate students. She researches issues related to English language assessment and teacher education. Susan is a committee member of the British Association for Applied Linguistics (BAAL) special interest group for testing, evaluation and assessment. She gave presentations at both the BAAL and European Association for Language Testing and Assessment (EALTA) conferences this year. Susan previously worked for the British Council where she undertook several research projects related to testing and assessment.

## **FREPA descriptors and their role and contribution to integration of students from mixed linguistic backgrounds in a multilingual world.**

The integration of descriptors that will enhance better teaching and learning as well as the integration of students from mixed linguistic backgrounds is at the centre of the Framework of Reference for Pluralistic Approaches (FREPA), based on the Common European Framework of Reference for Languages (CEFR), where four models of language and culture teaching approaches are selected alternatively in the building of guidelines for teachers having to elaborate pedagogical activities in multicultural, multilingual environments. Targeting the development of skills related to knowledge, self-knowledge and expertise of learners who are native, migrant, in mobility, or surrounded by a multilingual reality, the resources proposed by the

FREPA foster the continuity of interaction between different linguistic realities with the same goal: sharing knowledge discovery and respect for the culture of the other. This framework is a practical example of conceptualised and operationalised knowledge that can be integrated in a teacher preparation and professional development context. In this presentation, we will look at the role of language constructs and contexts in language learning and propose a connection between FREPA guidelines and scopes and challenges associated with a model chosen by practitioners who have integrated a learning component, acquired and evaluated through experiential learning projects with community service, into their courses. This links the broader theme of assessment tasks, similar to those that examinees might encounter in real-life multilingual contexts and content, with the recognised representation and value of experiential language learning in the community, as well as the revisited value of assessment based on students' journals used as an essential reflection and evaluation component.

**Laura Ambrosio**, University of Ottawa, Canada

A French Second Language (FSL) professor at the University of Ottawa (Canada), Laura Ambrosio is interested in practical and theoretical aspects of language acquisition, teaching and evaluation, including immersion and international accreditation. She publishes and contributes to the development of FSL pedagogical tools and manuals. Her current research focuses on the impact of various components of experiential learning in languages at the university level, as well as the needs and resources available to International Languages teachers in Canadian school contexts.

## From online diagnostic language assessment to tailored EFL learning --- CDA-based EFL listening diagnostic model construction

Personalised English as a Foreign Language (EFL) learning emphasises “tailored instruction”, aiming at meeting the learners’ diverse requirements. The challenge is, however, how to precisely identify learners’ individual differences and their gaps in language learning so as to provide “tailored” instructions. Our research team from Xi’an Jiaotong University (XJTU) has developed a web system of diagnostic assessment and strategy guidance (PELDiaG) which integrates psychometric methods and web technology in order to address the above challenge. This study focuses on one of the primary components of PELDiaG, addressing how a EFL listening diagnostic model is constructed by using Cognitive Diagnostic Approach (CDA). It aims to accurately measure learners’ strengths and weaknesses in listening comprehension. The construction starts with listening cognitive attribute identification, diagnostic test development and Q-matrix model design, followed by the G-DINA model analysis using OX software, processing the data to obtain the best model-data fit. The validated model can generate the following diagnostic results: 1. The learners’ group mastery pattern of listening attributes; 2. the individual listening attribute mastery levels; 3. the individual learners’s mastery pattern, that is, learners who share the same score but have different attribute mastery levels; 4. diagnostic feedback both in text and graph. The findings are significant in achieving the goal of personalized assessment so as to fulfill the purpose of “tailored” EFL learning. Once the test is validated, it is added to the system database designed for online diagnostic purposes. Through the system,

these learners' diagnostic reports and the guidance can be quickly made available in order to achieve the aim of "tailored" learning.

**Xiaomei Ma**, School of Foreign Studies, Xi'an Jiaotong University, China

Xiaomei Ma, Professor and PhD supervisor in applied linguistics, works at the School of Foreign Studies at the Xi'an Jiaotong University. Her research area covers L2 teaching, language assessment and technology. Her current research focuses on web-based EFL personalised learning diagnosis and strategy guidance, including L2 learning strategies, L2 reading, L2 listening and cognitive diagnostic assessment. She leads about 20 research projects, has published over 20 research papers and has won over 10 Awards and Honours involving teaching, research and whole-person education.

## Goethe's professional development of raters: Live-test data analysis as assessment

Human resource development at Goethe-Institut, the cultural institute of the Federal Republic of Germany with a global reach, includes a policy of continuous development for its 3,500 employees' skills and conduct. Development measures comprise training, coaching and systematic feedback. Goals and contents focus on competencies which are essential to realise the institution's strategic goals. In terms of the German language courses of Goethe-Institut and the area of language test development and test administration, one of the main objectives is to comply with ALTE's minimum standards. Goethe-Institut's challenge currently lies in creating a development scheme for its raters that is relevant and efficient enough in order to meet the ALTE minimum standard 12. That means that at more than 500 test centers and partners worldwide, the fair and reliable rating of candidates' of writing and speaking performances have to be a decentralised test administration with more than 435,000 exams documented. Data regarding achievement of raters has to be collected and analysed. The presentation summarises the results of a practical research study. Thereby, a great deal of importance is placed on the implementation of a feasible and resource-saving process for decentralised organisations and test providers which meets the minimum requirements and demands of the institution, stakeholders and test takers. The presentation will focus on examples of the administration of the writing subtest of the Goethe-Zertifikat B2 in Athens and Budapest. The individual ratings of several teams of two raters were collected, analysed with a commonly used software (SPSS, R) and Excel to permit a statistical evaluation in the form of correlations regarding the raters' performance.

**Michaela Perlmann-Balme**, Goethe-Institut, Germany

Michaela Perlmann-Balme holds a PhD from LMU Munich University and has taught German as a foreign language at the University of Otago, New Zealand. She develops and revises German tests for Goethe-Institut. Currently she is the chair of the Standing Committee of ALTE. She was responsible for the German tests in the European Survey of language competence (ESLC, Surveylang) investigating foreign language achievements of learners in schools and has coordinated the development of the German test for migrants.

**Jane Lloyd**, ALTE Validation Unit, United Kingdom

Jane Lloyd has over twenty years' experience in EFL and EAP and has worked as a teacher, trainer and educational manager specialising in language corpora, materials design and assessment. She has worked as a teacher and in the fields of curriculum and testing in language schools, state schools and higher education in the

Far East, Turkey and the Middle East. She is interested in teacher training, curriculum innovation, large-scale testing, language policy and data analysis.

## How big should the carrot be? An investigation into effects of differential incentivisation on students' standardised proficiency test scores

In both teaching and testing for proficiency, an important issue in understanding outcomes is the motivation that students have for achieving any particular level of ability. In most university settings, the only requirement for students of language is seat time (i.e., X number of credit hours completed). We report the findings of a study conducted on 257 students completing second year Spanish courses at a large US university. The programme in which they are enrolled has a stated outcome goal of Intermediate-Mid in speaking, listening and reading on the American Council on the Teaching of Foreign Languages (ACTFL) proficiency scale. However, the program does not require that students meet this goal. In our study, we incentivised two groups of students: for one group (n = 96) 10% of their final grade was based on meeting pre-established program goals: for a second group (n = 73), 5% of their grade was based on the same criteria. With a control group (n = 88), there was no incentivisation: They only received 5% for taking the three proficiency tests. We compared the proficiency measures on standardised listening, reading and speaking tests for the three groups taking into account the effect of gender. Our findings show that the 10% group performed significantly higher than the two other groups. In addition, girls outperformed boys in the 10% group. These results seem to suggest that programmes should only enact proficiency-based exit-tests if there are requirements tied to them that articulate the programs' stated goals (see Chalhoub-Deville 1997). We suggest that exit-test requirements be put in place, but that such tests will have to be integrated within the program and count as part of the students' programme requirements.

**Susan Gass**, Michigan State University, United States

Susan Gass received her PhD from Indiana University. She has published widely in the field of second language acquisition (more than 30 books and 100 articles) and is the Co-Principal Investigator (with Paula Winke) on a 3-year Proficiency grant awarded by the Flagship programme of the US Department of Defense. She has served as President of the American Association for Applied Linguistics and of the International Association of Applied Linguistics (AILA). She has won many local, national, and international awards.

**Koen Van Gorp**, Michigan State University, United States

Koen Van Gorp received his PhD in 2010 from the KU Leuven and is currently Head of Foreign Language Assessment at the Center for Language Teaching Advancement (CeLTA) at Michigan State University. From 2010 till 2015 he was Director of the Certificate Dutch as a Foreign Language (CNaVT), the international test of Dutch as a foreign language. His research interests are task-based language teaching and assessment.

## How far can digitalised language assessment assist in the teaching and learning of languages within the Italian university system?

The Italian university system provides courses at low cost for high numbers of students, on the basis of selection or open access, and on the principle of allowing students to take and retake exams numerous times at regular intervals. With university fees low and with few

significant alternative sources of income, the system is heavily reliant on paltry public funding (1.0% of GDP, against an average of 1.5% in the OECD as a whole: OECD 2011). Student numbers are high: in 2010, 48% of 19 year olds were enrolled at university (see Turri 2016). Exam sessions involve written and spoken assessment of scores of students throughout the academic year. Communicative language assessment requires subjective rather than objective testing (Morrow 1979). Reliability-validity tension (Davies 1978) means that the former are more valid but less reliable and the latter, the opposite. With subjective tests (STs), come concerns of rater bias and consistency (Schaefer 2008) that are aggravated by time constraints and the volume of scripts; with objective tests (OTs), which are better suited to automatic assessment and digitalisation, the testee is often purely responsive rather than productive. In this paper, we will discuss experiences of administering examinations at universities in Italy analysing both the issues and the practicalities involved in running valid and reliable tests within such an ecosystem. In particular, we will contrast traditional paper-based tests principally with computer-based multiple-choice tests (paper and digital) and with online platforms like Moodle as well as some custom-made apps for use with tablets. We will look at how such diverse tools can be used in conjunction with each other, discussing how far such tools effectively assist in the teaching and learning of languages at university and suggesting realistic and practical ways in which progress may be made.

**Thomas Wulstan Christiansen**, Università del Salento, Italy

Thomas Christiansen is an associate professor (English Language) and Vice Director of his University Language Centre and has been involved in teacher training since 2000 in southern Italy and Albania. He has worked as an external consultant on speaking tests for Cambridge English Language Assessment. He has also been language co-ordinator and director on a series of university master's run by universities in Italy and Poland. He has a special interest in tools for assessment and new technology in teaching.

## How politics influences the reception of a test: The case of an English C1-test for lecturers in Flemish universities

Apart from a testing purpose, language tests may have a considerable social or political impact. As tests select and inherently discriminate, passing a high-stakes test may yield important benefits, while failing a test can jeopardise one's future. Whereas tests can be employed as a means to create and maintain social order (Shohamy 2006), resulting in 'the politicisation of assessment' (McNamara and Roever 2006:213), they can also serve as a means to empower people (e.g. to gain access to education or employment). The presentation discusses the implementation and perception of the Interuniversity Test of Academic English (ITACE) for lecturers. In 2013, the Flemish government decided that all lecturers in higher education needed to prove their language skills if they wanted to teach in a language different from their mother tongue. This meant all lecturers teaching in English curricula needed to prove their C1 level. The Interuniversity Testing Consortium developed a test that fitted the purpose. Even though it was highly contested, the ITACE showed to have a considerable advantage when compared to other standardised tests such as TOEFL and IELTS. The case study reveals how

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

the implementation of the ITACE determined its reception. Implemented top-down, the test sparked a media storm in which the test became the scapegoat. Its very purpose – quality assurance in higher education – was largely neglected in the media. In the presentation we will discuss the press coverage, and we will argue where and why it was inaccurate. We will show how the top-down implementation of the test was counterproductive to the acceptance of the test as a means to ensure the quality of teaching and improve employability.

**Frank Van Splunder**, Interuniversity Testing Consortium, Belgium

Frank van Splunder is a language lecturer and tester at the University of Antwerp. His main research area is in language policy regarding English-Medium Instruction in Higher Education, both in Flanders and in other regions. Currently he is involved in the ITACE project as well as several other projects related to language policy and testing (e.g. in Cuba and in the DR Congo).

**Catherine Verguts**, Internuversity Testing Consortium, Belgium

Catherine Verguts coordinates the Germanic Languages department at the Ghent University Language Centre. She also has experience giving coaching sessions on test development and has been part of expert groups giving feedback on high-stakes tests. Moreover, she is part of the test development teams of ITNA (accredited with an ALTE Q-mark) and Interuniversity Test of Academic English (ITACE).

## How to assess mediation?

This presentation relates to a project of the Council of Europe's Language Policy Unit, currently in progress, aiming to develop a descriptive scheme and scales of illustrative descriptors for Mediation, the fourth (after Reception, Production, and Interaction) kind of language activity presented in the Common European Framework of Reference for Languages (CEFR). The first results of the project work indicate that the approach taken to Mediation is far broader than the one applied originally in the CEFR in that the relevance of the new descriptors is not restricted to the foreign/second language classroom. The new definition of Mediation competences seems to be equally relevant to all other contexts of language use. This striking interpretation may provide the 'missing link' between the set of language competences traditionally attributed to the learning, teaching, and assessment of foreign/second languages and those highlighted as required to get educated in all other subjects – under the umbrella of a really coherent plurilingual ability. In my presentation I am going to examine the new Mediation descriptors from the assessment perspective looking at ways to operationalise them as testing tasks and items within the Learning Oriented Assessment framework. A report on experimenting with assessment of Mediation skills in Polish – both as language of education and foreign/second language – will be included.

**Waldemar Martyniuk**, Jagiellonian University Kraków, Poland

Waldemar Martniuk is Assistant Professor and Executive Director of the Centre for Polish Language and Culture at the Jagiellonian University in Krakow, Poland. He is the author of textbooks, curricula, and testing materials for Polish as a foreign language, and chair of the State Certification Commission for Polish as a Foreign Language. He has worked as Executive Director of the European Centre for Modern Languages in Graz (2008-13), project coordinator at the Language Policy Division of the Council of Europe (2005-2007), and is a Visiting Professor and lecturer at universities in Germany (Bochum, Giessen, Goettingen, Mainz, Muenster), Switzerland (Basel), and in the USA (Stanford University).

## Implications of employing performance-based testing in a university context

The issue that this paper addresses is the employment of performance-based testing in a university context, in particular, at the Sapienza University of Rome, as well as the financial and practical implications of adopting such an approach, as opposed to the more common selected response approach to testing. The matter is addressed by aiming at designing a performance-based test of the English language at a CEFR B2 level. In particular, the issues addressed are: 1) the feasibility of employing a theoretical model of English language knowledge, (the Bachman and Palmer one, Bachman and Palmer 2013) to design a performance-based test and analytic and holistic scales that would adequately assess the written and spoken competence in English of first-year university students of the Sapienza University of Rome; 2) test validation within the university context for the test takers in question; 3) the feasibility of employing such a test at the Sapienza University of Rome considering the financial and other practical implications (cost-effectiveness, rater training, etc). To gather data, a questionnaire on personal data as well as two written and spoken tasks were designed accompanied by holistic and analytic rating scales based on the aforementioned model and the CEFR B2 illustrative descriptors. In the pilot test phase, 54 students of the Sapienza University of Rome completed the written test and the questionnaire, which was then rated by two raters. The pilot sample analyses have demonstrated a high variance of scores as well as a high internal consistency of the test ( $\alpha = .950$ ), which encourages the idea of employing the test at the University level. Cost-effectiveness and lack of trained raters, however, have proven to be a challenge.

**Snezana Mitrovic**, Sapienza University of Rome, Italy

Snezana Mitrovic, GI - Pearson LCCI Academic Manager, has been working in the area of assessment for 10 years, helping teachers prepare students for LCCI exams, running seminars and workshops recognised by the Italian MoE and carrying out a variety of projects. Her main areas of interest are the methodology of TEFL and assessment. Her wish to reconcile the commercial and academic in assessment has led her to take up a PhD programme in assessment at the Sapienza University of Rome.

## Integrating corpus linguistics & classroom-based assessment: Evidence from young learners' written corpora

Analysing writing development as a function of grade level is important in elementary school children because the developmental patterns are strongest at a young age when successful interventions are needed (Haswell 2000, McNamara et al 2010). At the same time, many researchers have highlighted the central role that assessment plays in second language classrooms and have expressed the need for research into classroom-based language assessment (CBLA), an area that is gradually coming into its own in the field of language testing and assessment (Hasselgreen 2008, Leung 2014, Turner 2012). Motivated by the above literature, this presentation reports on a one-year longitudinal study that aimed at assessing the writing proficiency of young learners of English as Foreign Language (EFL) as a function of grade level. A total of 50 EFL students aged 9 to 11 years old took part in the study from whom 500 written essays were collected. Using a range of advanced computational linguistics and

automated machine learning systems, the study aimed to examine the degree to which essays written at different grade levels can be distinguished from one another using a number of linguistic features related to propositional density, lexical sophistication, syntactic complexity and cohesion. The results showed statistically significant differences between the linguistic features identified in the essays produced by EFL young learners at different levels of language competence. These findings support the notion that linguistic development occurs in the later stages of writing development and is primarily related to producing more elaborate texts with more sophisticated words, more complex sentence structure and fewer cohesive features as a function of grade level. The presentation concludes by providing practical guidance to EFL teachers as to the kind of linguistic strategies young learners develop as a function of their grade level and suggestions to consider when designing EFL classroom curricula and classroom-based writing assessment activities.

**Trisevgeni Liontou**, Ministry of Education, Greece

Jenny Liontou works as an Adjunct Instructor at the Hellenic Open University and at the Faculty of English Studies, University of Athens. She holds a PhD in English Language and Linguistics with specialisation in EFL Testing & Assessment. She holds a BA in English Language & Literature, a MA in Lexicography: Theory and Applications and a M.Sc. in Information Technology in Education. She has worked as an EFL teacher in primary and secondary education, a freelance item writer, oral examiner and script rater for various international EFL examination boards and has published papers in the aforementioned fields.

**Dina Tsagari**, University of Cyprus, Cyprus

Dina Tsagari is an Assistant Professor in Applied Linguistics/Teaching English as a Foreign Language (TEFL) with specialisation in the area of Language Testing and Assessment at the Department of English Studies, University of Cyprus, Cyprus. Dina is the director of the Language Testing and Assessment Lab of the University of Cyprus and the coordinator of the Classroom-based language assessment (CBLA) Special Interest Group – EALTA. She has published widely and presented in numerous local and international conferences.

## Integrating technology with language assessment: Automated speaking assessment

With the advancement of natural language processing, machine learning, and speech recognition technologies, automated evaluation of non-native English speech is no longer a dream. An auto-marker, once accurately calibrated, is expected to enhance the reliability of a test because it eliminates idiosyncratic behaviours of human raters (Brown 2012, Williamson 2013). In addition, implementing an auto-marker for a speaking test of English would significantly shorten the turnaround time of score reporting and reduce the cost spent on examiner training and hiring. This paper evaluates a computer-delivered, fully automated speaking test of English that is used for selection, placement, or end-of-course evaluation of students enrolled at post-secondary education institutions. It begins with a review of the speaking test, including its structure and task types, and then summarises the validity evidence that has been obtained to support the use of the auto-marker. Approximately 1,200 candidates of various first languages and English proficiency levels sat the test and completed a post-test survey. Their oral responses were recorded by an online test delivery system and evaluated by both the auto-marker and certified human examiners. Based on Weir's (2005) socio-cognitive framework, some validity

enquiries were made about the construct coverage of the auto-marker (i.e., the language features it is capable of assessing), the agreement and relationship between automated scores and human scores, the auto-marker's level of severity when scoring, and candidates' perceptions of a non-human examiner. The research reported in this paper will shed light on the future development of automated speaking assessment.

**Jing Xu**, Cambridge English Language Assessment, United Kingdom

Jing Xu is a Senior Research Manager at the Research and Thought Leadership Group of Cambridge English Language Assessment. His research areas include automated language assessment, computer-assisted language learning, and test validity theory. Jing received his doctoral degree in Applied Linguistics and Technology from Iowa State University (ISU). Before joining Cambridge English, he worked as the assistant coordinator of ISU's Oral English Certification Test and a research intern at Educational Testing Service (ETS).

**Graham Seed**, Cambridge English Language Assessment, United Kingdom

Graham Seed is a Senior Assessment Manager in the Assessment department of Cambridge English Language Assessment. His primary responsibilities and interests are the management of the production of online speaking tests, examiner management, as well as teaching awards and curriculum development. He holds a master's degree in Language Testing from Lancaster University, UK, and previously worked as an EFL teacher and examiner in Bosnia, Russia and the UK.

## Investigating scoring procedures in language testing

One of the main concerns of a language tester, in the design and implementation of a test, is the method of scoring that needs to be selected for the tool used to perform the evaluation. Method of scoring is the precise quality in a test which a) remains in the hands of the language tester and b) determines test impact. As a result, it may be possible to argue that this attribute indirectly reveals the tester's ethical beliefs and personal stance in testing pedagogy. This study attempts to challenge the typical 1-0 method of scoring in multiple choice test items and implements a polychotomous partial-credit scoring system in official tests administered for the Greek State Certificate of Language Proficiency (GSCLP). The multiple choice tests chosen to act as the subjects of the study were completed by a total of 1917 subjects in different levels of the GSCLP test. The purpose of this investigation is to attempt to provide an alternative to the standard 1-0 method of scoring, typically selected in such situations, and thus challenge this ethical decision. The proposed polychotomous partial credit scoring system was tested and compared to the standard scoring system in different levels of the GSCLP test. Results clearly indicate that this scoring procedure provides refined insights to students' interlanguage level and enhances sensitivity in scoring procedures without jeopardising test reliability. Impact to fairness and ethics in language testing, resulting from the analysis of the data, is considered in the discussion section of the paper.

**Anna Mouti**, University of Thessaly, Greece

Anna Mouti holds a BA in English and a BA in Italian Studies, an MA in Language Didactics, and a PhD in Applied Linguistics from Aristotle University of Thessaloniki, Greece. She has a long teaching experience as a visiting lecturer in English for Academic Purposes (EAP), English for Specific Purposes (ESP) and Greek as a second language. She is currently a member of the Special Teaching Staff at University of Thessaly. Her research interests include Second Language Acquisition (SLA), Language Testing and Individual Differences in Language Learning and Assessment.

## Investigating the necessary elements to design and implement a communicative test for engineering students: A backwash effect

This presentation has as a main purpose to evaluate engineering students' L2 communicative competence via the implementation of a localised, suitable, and specific test for them. The elements and characteristics of this test will be taken from the sociolinguistic aspects inherent to the engineering field as well as from needs analysis data. Several frameworks to design communicative tests already exist in the market, but the ones proposed by Salvisberg (2010) and Granville (2011) represent appropriate models and guides to design a purposeful test for engineering students. Because of the need for using the English language as a communicative tool, the implementation of this test can be seen as a possible solution to meet stakeholders' perceptions and interest in finding ways to promote the communicative competence in this type of learners. Assessment of English for Specific Purposes (AESP) has a major role in this project, as well as the field of communicative competence. The backwash effect of this test implementation will be observed and analyzed to get results and conclusions. According to Watkins, Dahlin, and Ekholm (2005), "an effective way to change student learning is to change the form of assessment." This statement portrays the main purpose of this study.

**Ada Luisa Arellano Méndez**, Mextesol, Mexico

Ada Arellano is the English Coordinator at the Polytechnic University of Altamira, Mexico. She holds a BA in English and a MA in ELT, granted by the University of Southampton, UK. She is also a certified English teacher, with In-Service Certificate in English Language Teaching (ICELT), Teaching Knowledge Test (TKT), and other Cambridge English certifications. She has been a speaker in several congresses in her own country with several topics related to assessment. She is currently studying a PhD in ELT in a distance programme at the University of Southampton, UK. Her main research field is assessment. Her research is mainly based on studying the backwash effect after the implementation of a communicative test for engineering students.

## Investigations into the on-screen marking function in a tablet-based English reading test

While computers have been prevailingly used for assessing English reading comprehension, test takers have been documented to experience a number of difficulties caused from the on-screen reading conditions. In particular, one of such reported difficulties is the inability to apply marking (e.g., highlighting and annotating) on the given reading passages or test items, as it would be allowed in printed materials (Khalifa and Weir 2009). Acknowledging that tests should be biased for the best (Swain 1984), the present study created a reading test operated through a tablet PC (iPad) with an on-screen marking function for letting test takers make markings on a test in a manner similar to the paper-based reading/testing contexts. Participants in this study were 89 Korean college students, who were divided into three groups: A paper-based group, a marking-enabled tablet PC group (+ marking function), and a marking-disabled tablet PC group (- marking function). All participants took pre- and post-reading tests, with the former delivered in paper and the latter presented through the assigned formats (paper, + marking tablet PC, - marking tablet PC) according to participants' testing groups. After data collection, a series of quantitative (e.g., the number of words and items marked by participants)

and qualitative (whether the markings on texts corresponded to information relevant to the answers) analysis on participants' marking was performed. In addition, a correlation statistics was conducted between the quantity of marking and participants' reading test scores. Results indicated that marking strategies are frequently employed across test formats and that the amount of marking (on both the texts and on the test items) significantly correlated with test performance in all testing groups. Complimented by the interview responses of participants, the present study discusses the facilitative role of the on-screen marking feature in the environment of on-screen reading assessment.

**Shinhye Lee**, Michigan State University, United States

Shinhye Lee is a PhD candidate in Second Language Studies at Michigan State University. Throughout her doctoral studies, she has been working as a research assistant for Drs. Paula Winke and Susan Gass for the Language Proficiency Flagship Initiative and taught undergraduate language and American culture courses. She also worked as an item writer/reviewer for ACTFL Reading/Listening Proficiency Tests on Korean. Her research interests primarily include language testing practices and speaking performance.

## Italiano L2: Nuovi pubblici, nuove tendenze, nuove forme di valutazione

Le forme di accertamento delle competenze linguistico-comunicative delle lingue seconde o straniere si sono negli ultimi decenni gradualmente modificate per affrontare in maniera adeguata i numerosi cambiamenti e le innovazioni didattiche risultanti da studi e ricerche sull'apprendimento delle lingue e sulla valutazione delle competenze. L'enfasi da tempo posta sull'uso delle nuove tecnologie per l'apprendimento e per l'insegnamento, la preminenza delle competenze audio-orali, l'adozione di approcci lessicali, l'utilizzo di attività comunicative fondate su compiti non linguistici sono solo alcune delle innovazioni che sollecitano modifiche anche nelle forme di valutazione adottate. Nel caso dell'italiano L2, sia in ambito formativo sia in ambito certificatorio, oltre alle innovazioni didattiche, si è dovuto tenere conto anche dei mutamenti dei profili delle diverse popolazioni migranti e delle loro differenti aggregazioni presenti in Italia. La presenza della popolazione immigrata, caratterizzata sempre più spesso anche da una categoria particolarmente vulnerabile quale quella del rifugiato e richiedente asilo politico, con profili linguistici e con conoscenze e abilità estremamente diversificate – dall'analfabetismo alle competenze informatiche –, nonché la richiesta di offrire esami certificatori che tengano conto di questi nuovi profili e ne rilevino le competenze, hanno determinato la necessità di interrogarsi sulle forme di accertamento più idonee a fare emergere le capacità linguistico-comunicative di questo nuovo 'pubblico' certificatorio. Questa presentazione offre i primi risultati di una ricerca di nuove forme e modalità di accertamento e di valutazione delle competenze linguistico-comunicative in italiano a livello A2, di gruppi di migranti e di rifugiati nell'ottica del quadro delle competenze chiave per l'apprendimento permanente.

**Lucilla Lopriore**, CERT IT Università Roma Tre, Italy

Lucilla has a PhD in Spoken Assessment in Italian Certifications from the University for Foreigners of Siena and an MA TEFL from Reading University. She is a scientific committee member and responsible for CERT IT (University of Roma Tre, CLIQ). She is also an Associate Professor, Foreign Languages Department at Roma Tre University, a teacher and teacher educator for pre- and in-service second language teachers and for CLIL teachers. She has published extensively in teaching methodology, and is interested in Italian L2, assessment, CLIL and ELF.

**Giuseppina Vitale**, CERT IT Università Roma Tre, Italy

Giuseppina Vitale, as Lecturer on contract, teaches the course on "Language Learning" at Naples University "L'Orientale". She has also collaborated - as test developer and test rater - with the Certification of Italian as a Second Language – CERT.IT – at Roma Tre University. Her main research interests are: second language teaching, particularly Italian as a second language, language assessment, and e-learning.

## La verifica come occasione di apprendimento e aggiornamento attraverso l'esperienza della certificazione glottodidattica DILS-PG di II livello

Obiettivo del presente contributo è quello di evidenziare, attraverso l'esperienza della certificazione in "Didattica dell'italiano come lingua straniera" DILS-PG di II livello, elaborata e prodotta dal CVCL (Centro per la Valutazione e le Certificazioni Linguistiche) dell'Università per Stranieri di Perugia, come la verifica delle conoscenze e competenze richieste a un docente di italiano a stranieri per poter svolgere al meglio il proprio lavoro non sia finalizzata solo alla valutazione, ma possa anche costituire un'occasione di apprendimento, di sistematizzazione, di riflessione sulla propria esperienza di insegnamento e uno stimolo per un aggiornamento continuo. Verranno pertanto prima illustrate le conoscenze e le competenze testate nell'esame di certificazione glottodidattica DILS-PG di II livello (Santeusano 2013a, 2013b, 2014a, 2014b) e poi presentati i risultati preliminari di una ricerca più vasta, basata su una sistematica raccolta ed analisi dei dati che emergono dalla somministrazione della DILS-PG di II livello (Marasco/Santeusano in corso di stampa, Santeusano in fase di preparazione). In modo particolare, dopo aver descritto il profilo degli iscritti all'esame, relativamente al campione di riferimento, verranno analizzate le risposte dei candidati nel questionario somministrato al termine della prova soffermandosi sui singoli casi e proponendo elementi e spunti di riflessione in base alle risposte stesse.

**Nicoletta Santeusano**, Università per Stranieri di Perugia - CVCL, Italy

Nicoletta Santeusano obtained her PhD in Linguistics and didactics of Italian as a Second Language (ISL) at the University for Foreigners of Siena. She taught ISL at the University of Cologne in Germany (1998-2004). Currently she is working at the University for Foreigners of Perugia as a teacher of ISL and as a teacher trainer. As an expert in didactics, she works at the II Centro per la Valutazione e le Certificazioni Linguistiche (CVCL) in the elaboration of the Certificate in Didactics of ISL (DILS-PG). She is an author of publications about learning and acquisition of ISL, DILS-PG and of textbooks for learning ISL.

## Language learning and assessment transformation: An opportunity for educative innovation

Being able to speak and to communicate in a complex and globalised world in different languages enhances personal development, employability and academic and professional success. In our opinion, plurilingualism should therefore be a social and personal goal. From this perspective, the aim of language education is profoundly modified to develop a linguistic repertoire, in which all linguistic abilities have a place (the languages offered in educational institutions should be diversified and students given the opportunity to develop a plurilingual competence). The aim is no longer simply to achieve 'mastery' of one or two, or even three languages, each taken in isolation, with the 'ideal native speaker' as the ultimate model. Language learning is a lifelong task and the development of a young person's motivation, skill and confidence in facing new language experience out of school comes to be of central importance. To this end, promoting structural changes in the school system are imperative. Firstly, changes in the educational relationship and the teachers' and students' roles: The teacher is no longer a knowledge transmitter, they must foster a love of learning in their students. Secondly, methodological changes: Learning through all languages and subjects and co-ordination between language teachers and subject teachers. Thirdly, changes in the classrooms: The classroom must be a flexible environment that allows students to take the lead and participate in collaborative team work. Last but not least, new training strategies and professional teachers development: The responsibilities of educational authorities, qualifying examining bodies and teachers cannot simply be confined to the attainment of a given level of proficiency in a particular language at a particular moment in time, important though that undoubtedly is. Let's see it in the Catalan Framework for Plurilingualism.

**Mònica Pereña**, Generalitat de Catalunya, Spain

Mònica Pereña has a degree in Catalan Philology at the University of Barcelona and a Diploma of French Studies at the University of Toulouse-Le Mirail, France. She is currently Deputy Director of Language and Multilingualism at the Ministry of Education of Catalonia from where she runs the linguistic policy of the Catalan education system. From that position she develops linguistic programmes oriented to improve teachers' performance, on the methodological field and in school organisation, and has responsibilities on in-service teacher training.

## Learners' goals and the impact of assessment for and as learning: Examples from computerised diagnostic and dynamic assessment

The current understanding of the language assessment underscores the participatory nature of assessment, where both learners and teachers contribute to design and implementation of assessment, and bridging the gap between assessment and instruction. In the present paper, we will synthesise the findings of several studies in order to argue for the importance of taking into account learners' goals and perceptions of and beliefs about assessment when interpreting their performance on computerised assessments for and as learning. We will also discuss how teachers' understanding and perceptions of assessment can impact their use of such assessments. We will draw upon examples from validating DIALANG, a

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

computerised diagnostic assessment system in 14 languages (e.g., Alderson 2005) and research conducted by the authors and elsewhere (e.g., Thouësny 2011) on computerised dynamic assessment to illustrate how learners' goals in, perceptions of and beliefs about assessment impacted on the way they interpreted and used the feedback they received and, in fact, their performance. We will also touch upon some relevant research on teachers' classroom assessment practices to illustrate how teachers employ language assessments in the classroom. Finally, we will illustrate how experience of assessment for and as learning can facilitate changes in learners' beliefs about and perceptions of (and consequently, perhaps, goals in) assessments.

**Dmitri Leontjev**, University of Jyväskylä, Finland

Dmitri Leontjev is a postdoctoral researcher at the Centre for Applied Language Studies of the University of Jyväskylä, where he also received his PhD in May 2016. His research interests and most publications relate to corrective feedback and L2 dynamic assessment. He is particularly interested in studying impact of dynamic assessment on learners.

## Learning from assessment: Teachers' relationship with data driven learning.

In an era of pedagogy which values data-driven learning, the wealth of data offered by a corpus of speaking assessment texts is potentially gold dust to teachers and course developers. But the process of converting valuable spoken corpus data into classroom materials is not necessarily straightforward, as a recent project conducted by Trinity College London reveals. This presentation outlines a project to create classroom materials from assessment data and the stumbling blocks which occurred along the way. Firstly, the project involved analysing data from the corpus which appears to be missing from current coursebooks and which could significantly enhance learners' communication skills across cultures. The data was then contextualised and presented to teachers and trainers in order to give them an opportunity to use their expertise in guiding how this data could be developed for and utilised in the classroom. The outcome was disappointing with teachers retreating into what they knew and resisting the data in favour of traditional and prescriptive approaches. Finally, the presentation describes how the classroom materials were eventually created but raises questions about how far teachers can absorb new approaches to input without further training.

**Elaine Boyd**, Trinity College London, United Kingdom

Elaine Boyd has been involved in language teaching, teacher training and assessment design for a range of international organisations. She develops courses for teacher education and assessment literacy and has published articles in the fields of assessment and intercultural communication as well as authoring several coursebooks. Elaine is a postgraduate tutor at University College London and her PhD investigated spoken language and pragmatics. She is currently working with Lancaster University on the Trinity Lancaster Corpus of Spoken Language.

## Learning oriented assessment: Making the connections between learning, assessment and technology

In recent years Learning Oriented Assessment (LOA) has emerged as the dominant theoretical trend for conceptualising the relationship between learning and assessment. This is mainly because LOA does away with the binary terms traditionally associated with assessment (formative vs. summative, classroom vs. large-scale, assessment for learning vs. assessment of learning) and views all forms of assessment as supporting learning in a systematic and integrated fashion. In this presentation we will briefly present the concept of LOA and will discuss its key features and the synergies between learning and assessment tasks, learner involvement and targeted feedback, which have been shown to increase learner autonomy and learning gains. We will then illustrate the practical application of key LOA principles through technology – more specifically, through a demonstration of the on-line tutoring system ‘Write and Improve!’ and the ELT course ‘Empower’ (Cambridge University Press and Cambridge English Language Assessment), which provides a platform for learner development through integration of learning and assessment. Through this demonstration of practical applications, we aim to show how technology can play a fundamental role in helping teachers and learners in the collection of learning data, in structured record keeping, and in providing individualised and pedagogically useful feedback through automated scoring and feedback systems.

**Angeliki Salamoura**, Cambridge English Language Assessment, United Kingdom

Angeliki Salamoura holds a PhD in English and Applied Linguistics from the University of Cambridge. She is Principal Research Manager at Cambridge English where she leads on research related to Learning Oriented Assessment, language learning and CEFR issues. Angeliki has also worked on English Profile. She has published widely in academic and teaching forums, and she presents regularly at international conferences.

**Sarah Unsworth**, Cambridge English Language Assessment, United Kingdom

Sarah Unsworth is Assessment Group Manager at Cambridge English Language Assessment, where she has worked on developing and producing tests for General English and for English for Specific Purposes. Sarah is currently managing a project which includes the provision of a Cambridge English blended course that incorporates online assessment and appropriate follow-up activities. Sarah has substantial experience in assessment development and delivery, especially in the area of Learning Oriented Assessment. She also has many years' experience as an English language teacher and teacher trainer in Germany and the UK. She has a London Chamber of Commerce and Industry diploma in Teaching English for Business and a postgraduate certificate in Education in Modern Languages and TESOL.

## LT123: meeting the challenges of providing quality outsourced test materials for a range of clients

LT123 is an independent company providing expertise that supports exam boards and other awarding bodies as well as publishers producing exam materials. The company has access to a bank of experienced and successful test writers, editors and other relevant experts enabling us to produce large quantities of quality test materials at relative speed. Outsourcing to us can, obviously, be a major advantage for our clients whose own staff may be too few or too involved in other projects for them to be able to cope in-house. For publishers there are additional benefits

in that they can be confident that exam support materials that we write are being produced by assessment experts and not just by a teacher or writer who happens to be familiar with the exam in question. We go through the processes of exam production outlined in the ALTE Manual for Language Test Development - determining the specifications, planning and organising the writing teams, dealing with editing and setting up a thorough system of quality control. The ways in which we manage each of these stages may differ somewhat from the methods used by large exam boards but we take pride in the fact that our approach is as rigorous as the best of those. The challenges for us are that we have to meet the needs of our clients. We, therefore, have to spend time making sure that we understand precisely what those needs are, and if necessary negotiating to adjust their focus. We have recently joined forces with NILE in order to further extend our capacities and consultancy services we can offer around the world. We have developed an LT123 Manifesto outlining the principles that we believe form the basis of our approach to quality assessment.

## **Felicity O'Dell, LT123, United Kingdom**

Felicity O'Dell is a freelance writer based in Cambridge, UK. She has many years of experience of teaching, mainly English as a Foreign Language (EFL) but also modern languages. She regularly prepares materials for exam boards and has also written or co-written over 50 books for English Language Teaching (ELT) publishers. She has been a consultant for LT123 since it was established. Originally from Scotland, she has a degree in Russian, a doctorate on education in the USSR and an MA in Teaching English to Speakers of Other Languages (TESOL).

## **Russell Whitehead, LT123, United Kingdom**

Russell Whitehead runs LT123 Ltd, a Cambridge-based consultancy specialising in assessment and materials for paper and digital whose clients include exam providers, publishers and universities. He has degrees from Oxford and London universities. He was a teacher for over ten years before becoming an item writer and author. He established LT123 in 2011.

## **Measuring (linguistic) integration? German tests for migrants**

Language tests within the context of migration are not merely designed to assess individual learning success but have implications far beyond that. Certificates have legal implications as requirements for residency and citizenship and they are even supposed to indicate the degree of integration. Also, various stakeholders make very diverse use of these certificates – test takers, course providers, teaching staff, public authorities, politics and society. The language test provider is therefore challenged to fulfil all the requirements in an appropriate manner without compromising test fairness. The talk will focus on the German test for migrants (Deutsch-Test für Zuwanderer) and related tests for German as a second language. In contrast to a foreign language context (in schools etc.) it is clear that not only must the topics be tailored to the target group but that linguistic competence takes on a very different dimension. Migrants have to organise their lives in the new linguistic environment. Thus, they cannot wait until they have reached a certain competence level before they start using the language in real life. Finally, the talk will demonstrate that the German test for migrants does in fact require willingness to integrate in the sense that someone who studies hard but never leaves the

classroom to familiarise him- or herself with the practical use of the language will find it difficult to pass the test. In this respect it is, to a certain extent, a measure of integration or the willingness to integrate. Providing a special test for migrants should enhance validity, not endanger it. The test provider has to find a valid balance between “integration” as part of the test construct and keeping within the boundaries of a language test.

**Sibylle Plassmann**, telc, Germany

Sibylle Plassmann is Head of Test Development at telc gGmbH. Working in a multi-national team providing tests in 10 different languages, her own focus lies on German as a second language. She is an expert on language tests for migrants, (medical) Language for Specific Purposes (LSP) tests and academic tests. She was part of the development team of the Deutsch-Test für Zuwanderer (German test for migrants), for its Austrian version as well as for the recently developed Swiss language test for migrants (fide Sprachnachweis).

## Measuring the washback of a learning-oriented assessment

Tests have an undeniable impact on learning and teaching. In contexts where the use of tests is necessary, how to use tests for learning-oriented assessment to generate the intended positive washback on learning and teaching is a challenge for course developers and language testers. To achieve critical alignment of assessment with learning and teaching, the Hong Kong University of Science and Technology (HKUST) has designed a customised curriculum-driven assessment tool called the English Language Proficiency Assessment (ELPA) for the English Core Curriculum for foundation year students. The ELPA assessment model builds on principles of assessment for and as learning, and the hypothesised relationship between teachers' assessment literacy and involvement and their teaching effectiveness. In this presentation, we will introduce the ELPA assessment model and how it provides timely diagnostic, formative feedback for students on the one hand, and engages course developers and teachers in a dialogic reflection on pedagogy on the other. Findings from a research study into the washback of ELPA will be presented. The study investigates the directions and intensity of the impact ELPA has on learning and teaching, via comparisons of students' scores on criterion abilities in testing and non-testing situations, students and teachers' beliefs, motivation, assessment literacy and ownership to examine the test washback. The study aims to show how a test operates within an educational context for an ecologically grounded understanding of test washback. Implications for assessment and curriculum integration for enhanced learning and teaching will also be discussed.

**Edward Li**, Center for Language Education, The Hong Kong University of Science and Technology, China

Edward Li is a Senior Instructor in the Center for Language Education at the Hong Kong University of Science and Technology. He is Head of Assessment and Evaluation and Director of the English Language Proficiency Assessment (ELPA) project. He has served as Chair of English Language Assessment Review Committees for primary and secondary education in Hong Kong. His publications and research interest are in the fields of language testing and assessment, curriculum alignment, and programme evaluation.

**Keith Tong**, Center for Language Education, The Hong Kong University of Science and Technology, China

Keith Tong has 38 years' experience in language education in Hong Kong, teaching at secondary and tertiary levels, and being involved in teacher education for language teachers. He is currently Acting Director of the Center for Language Education at the Hong Kong University of Science and Technology, and is responsible for the development and implementation of language courses for both undergraduates and postgraduates at the university. His areas of interest include English for Academic Purposes (EAP), Content Language Integrated Learning (CLIL) and language assessment.

## Modeling oral proficiency development across four languages with the ACTFL OPIc

In a multilingual world, assessing speaking ability in second/foreign languages is useful for many reasons, including language teacher certification, employment decisions, monitoring language learning progress, or evaluating language programmes (American Council on Teaching Foreign Language (ACTFL) 2014). In the United States, the ACTFL Proficiency Guidelines (2012) are used to describe varying levels of foreign language proficiency, and the Guidelines are parallel across foreign languages. Accompanying the Guidelines are a suite of ACTFL-sanctioned language proficiency tests for adults. This presentation investigates the construct validity and comparability across four languages of one ACTFL speaking test, the Oral Proficiency Interview-Computer (OPIc), complementing recent investigations of listening and reading abilities as measured by ACTFL tests (e.g., Tschirner 2016). One key assumption of construct validity is that changes in the underlying attribute, in this case speaking proficiency as defined by the Guidelines, should result in commensurate changes in the measurement outcome, in this case proficiency ratings on the ACTFL scale (Borsboom et al 2004). In ACTFL's Guidelines, it is also implied that successive proficiency classifications entail increasingly large gains in ability; this too should be reflected in measurements. To this end, longitudinal OPIc score data from 840 students studying Chinese ( $n = 149$ ), French ( $n = 255$ ), Spanish ( $n = 386$ ), and Russian ( $n = 50$ ) at a large Midwestern US university were collected to a) investigate proficiency development over time and b) to compare proficiency development across languages. Each participant has at least two OPIc scores, with some participants' scores spanning two years. Background data and information on language learning experience were also collected. The relationship between language study and proficiency growth in terms of OPIc scores will be discussed in light of target language and relevant background variables. Findings will have implications for using the OPIc and for tertiary foreign language learning.

**Daniel R Isbell**, Michigan State University, United States

Daniel R Isbell is a PhD student in Second Language Studies at Michigan State University, where he also works in the English Language Center testing office. He earned an MA TESL from Northern Arizona university and has over five years of English teaching experience. His primary research interest is language assessment. He has previously presented work at the Language Testing Research Colloquium and the Midwest Association of Language Testers, and he has published in the journal *Language Testing*.

## Modelli di validazione qualitativa in contesti di large-scale assessment per le competenze linguistiche

Costruire l'argomentazione di validità di un test è un processo sempre in fieri che procede parallelamente al ciclo del testing (ALTE, 2011), dalla concezione del costrutto alle tappe di analisi e verifica post-test. In questo contributo ci soffermeremo sulla descrizione delle misure finalizzate ad accrescere la trasparenza sulle caratteristiche del sistema di valutazione delle competenze linguistiche SELF, test online di posizionamento, con orientamento diagnostico-formativo, di tipo criteria-reference e autocorrettivo, sviluppato per sei lingue (italiano, inglese, cinese, giapponese, spagnolo e francese), già somministrato in italiano, inglese e cinese a oltre 15.000 studenti francofoni iscritti in 14 diverse istituzioni. Ci focalizzeremo in particolare sulla descrizione dei risultati ottenuti in occasione delle tappe di validazione qualitativa post-somministrazione ad ampia scala. La validazione quantitativa, le analisi psicometriche e lo standard setting ('bookmark method') hanno determinato la costituzione di una banca di item affidabili, la definizione dei punti di cesura e l'attribuzione dei livelli (nel nostro caso sulla base della scala del QCER). Il modello di validazione qualitativa si è però rivelato indispensabile per effettuare osservazioni ragionate sui risultati del test, correggere leggere tendenze alla sopravvalutazione o sottovalutazione legate alle caratteristiche del campione e conoscere la percezione della comunità composta da insegnanti e studenti, principalmente francofoni ma con biografie linguistiche e socioculturali estremamente diversificate. Nello specifico abbiamo confrontato i risultati di SELF con la performance rilevata in classe dall'insegnante in occasione di una verifica iniziale (contôle continu) e abbiamo indagato la percezione degli studenti osservando la correlazione con i giudizi espressi in autovalutazione. L'osservazione di questi dati procede di pari passo con l'adozione di alcune misure per la 'concurrent validity': osservazione della correlazione con i risultati di un esame di certificazione linguistica di tipo comunicativo-azionale (CLES) e della correlazione tra i risultati in 'espressione scritta breve' di SELF e le produzioni aperte di tipo guidato, grazie ad un nuovo standard setting basato sul metodo del corpus di produzioni di apprendenti.

**Cristiana Cervini**, Université Grenoble Alpes, France

Cristiana Cervini is Adjunct Professor in foreign language teaching and Research Fellow at the University of Bologna. After six years of different research grants on projects about language learning, applied linguistics and CALL, since 2012 she has been responsible for the testing and evaluation digital and multilingual device called SELF (Système d'évaluation en langues à visée formative) within the project IDEFI Innovalangues, coordinated by the University of Grenoble Alpes. She also belongs to the scientific committee of the CLES certification.

**Monica Masperi**, Université Grenoble Alpes, France

Monica Masperi is a Senior Lecturer in Linguistics and Didactics at Grenoble Alpes University (UGA, France). At the head of the LANSAD department (Languages for non-specialists) from 2004 to 2016, she is presently Head of the "Pôle recherche & actions" related to UGA Language Learning Centre and Scientific Director of the Innovalangues project (ANR-IDEFI). Her current research as a member of the LIDILEM research lab focuses on Italian didactics, plurilingualism and the use of technology in language teaching and learning.

## Monitoring languages in a three language policy setting: Experiences in Kazakhstan

This presentation will show the implementation and results of an educational monitoring system based on computer-based tests (CBTs) in Kazakhstan. Kazakhstan has a language policy with three official languages: Kazakh, Russian and English. The vast majority of the population either speak the Kazakh or Russian language at home and the English language is actively promoted by the government. To improve proficiency levels in all three languages at a young age, these languages are given a prominent position in the Kazakh educational system. Three years ago, a system monitoring the progress in abilities in these three languages was introduced in Kazakhstan by Cito and the Kazakh counterpart CPM. For all three languages, the four skills (reading, writing, listening and speaking) are measured. The monitoring system aims to diagnose the strengths and weaknesses of students. Results are delivered in such a way that they can be used to support language teaching effectively. Student growth in performance over the years can be assessed meaningfully among others because the results of the monitoring system are analysed using an item response theory model. A next step that is being considered in the monitoring system and that will be discussed during the presentation, is the linking of the monitoring system to the Common European Framework of Reference for Languages (CEFR) using the data-driven direct consensus (3DC) method.

**Remco Feskens**, Cito, National Institute for Educational Measurement, Netherlands

Remco Feskens works as a senior research scientist at Cito in the Netherlands. He obtained his PhD in methods and statistics from Utrecht University. His main research interests are survey methodology and educational measurement. He (co)-authored several papers which have been published among others in the Review of Educational Research, the Medical Teacher and the Journal of Official Statistics.

**Anneke de Graaf**, Cito, National Institute for Educational Measurement, Netherlands

Anneke de Graaf works as an international senior consultant in the domain of language assessment at Cito in the Netherlands. She is also involved in the development of national language exams in the Netherlands.

## Nonverbal delivery in speaking assessment: An intercultural case study

Nonverbal delivery, with its interactiveness and interdependence with verbal language, plays a crucial role in communication effectiveness. However, how nonverbal delivery in English speaking is intervened by L2 learners' native culture remains underexplored. This study, adopting a Multi-Trait, Multi-Method approach, first validated a four-dimensional (pronunciation and intonation, grammar and vocabulary, discourse management and nonverbal delivery) speaking rating scale for group discussion in formative assessment (Pan 2016). It was found that, considering the interpretability and consistency with previous studies regarding speaking ability taxonomy, a second-order correlated-trait, uncorrelated-method model not only provides sound goodness-of-fit indices ( $\chi^2(28) = 462.796$ ,  $p=.818$ ; CFI = 1.000; NNFI = 1.024; SRMR = .015; RMSEA = .000; 90% C.I. = .000, .060), but also presents divergent validity ( $\Delta\chi^2(9) = 403.08$ ,  $p < .001$ ,  $\Delta CFI = .472$ ) and discriminant validity ( $\Delta\chi^2(17) = 425.68$ ,  $p < .001$ ,  $\Delta CFI = .146$ ).

From a multimodal discourse analysis perspective, an intercultural case study was carried out to see whether test takers of the same proficiency level but different cultural backgrounds (Chinese vs. Malaysian) may perform differently in nonverbal delivery. It was found that although the participants are of the same proficiency level as measured by their International English Language Testing System (IELTS) score and other academic achievements, their performance in nonverbal delivery seemed different. The Malaysian participant would not only instantiate a diversity of nonverbal channels along with the accompanying verbiage, but also shift the function-specific channels whenever necessary. By comparison, although the Chinese participant performed quite satisfactorily in terms of formal nonverbal channels, a certain passiveness and hesitation could be detected. This study extends the literature regarding the construct of speaking ability and yields significance in the sense that if nonverbal delivery is incorporated into speaking assessment, cultural factors should be accordingly taken into account.

**Mingwei Pan**, Guangdong University of Foreign Studies, China

Mingwei Pan is Associate Professor of Applied Linguistics at the Faculty of English Language and Culture, Guangdong University of Foreign Studies. He obtained his PhD from Shanghai International Studies University and worked at the Hong Kong Polytechnic University. He was also a Fulbright scholar at University of Illinois at Urbana-Champaign. His main research interests include language assessment literacy and interface between language assessment and corpus linguistics.

## Open Badges: A new way to prove skills

Verified, secure and transportable, Open Badges are a twenty-first century way to collect and display evidence of learning, skills, qualifications or interests. They can be displayed on social networks, online CVs, portfolios or jobsites with tamper-proof digital evidence 'baked in' to the badge. Employers, educational institutions and others can see displayed skills, interests and achievements – all at the click of a button.

This talk will introduce Open Badges and show how they are being used in industry and education by organisations including Microsoft, O2, and the BBC. We will also present and discuss findings from Cambridge English's pilot studies using badges to support teachers' Continuing Professional Development (CPD).

We will explore questions such as whether Open Badges work as an incentive for CPD and the relative benefits and risks of these alternative credentials to teachers, learners and employers. We will share ideas on how teachers and institutions might use badges as a marketing tool, to recognise staff development or to support learners. Badges can be used to acknowledge achievement and gamify learning– encouraging learners to move from awareness, through to knowledge, skill and mastery of a topic area, subject or skill.

**Sarah Ellis**, Cambridge English Language Assessment, United Kingdom

Sarah Ellis is Head of Assessment Services, for Cambridge English Language Assessment, Southern Europe and is based in Bologna, Italy. She has an extensive background in teaching and teacher training in Italy and the UK, including CELTA and DELTA courses and has been involved with Cambridge English examinations for over fifteen years. Her work as a trainer has also involved working in Romania, Greece, Serbia, Bulgaria, Slovenia and Croatia. She is currently working on projects related to teacher education, assessment and Continuing Professional Development (CPD) for Teachers.

## Outcomes of the introduction of an external English language assessment in Portugal

This presentation reports on the findings of a two year study into the impact on high school classroom practice in Portugal of the introduction of an external benchmarking test. The study collected test data and attitudinal data from all 15 year olds in high school classrooms in Portugal, and surveyed high school teachers and parents. The data provided a detailed picture of the engagement of pupils with English inside and outside the classroom, their linguistic aspirations, and their proficiency. It also identified attitudinal factors which had an impact on their performance on the test, and looked at the views of parents, teachers, pupils and school administrators towards external measurement and certification of English.

**Jane Lloyd**, Cambridge English Language Assessment, United Kingdom

Jane Lloyd has over twenty years' experience in English as a Foreign Language (EFL) and English for Academic Purposes (EAP) and has worked as a teacher, trainer and educational manager specialising in language corpora, materials design and assessment. She has worked as a teacher and in the fields of curriculum and testing in language schools, state schools and higher education in the Far East, Turkey and the Middle East. She is interested in teacher training, curriculum innovation, large-scale testing, language policy and data analysis.

## Pensare CLIL con Cambridge

Il termine CLIL, introdotto da David Marsh e Anne Maljers nel 1994, è l'acronimo di Content and Language Integrated Learning, apprendimento integrato di contenuti disciplinari in lingua straniera veicolare. Nel 2014, in Italia, sono diventate operative le norme inserite nei Regolamenti di riordino (DPR 88 e 89/2010) che prevedono l'obbligo di insegnare una disciplina non linguistica (DNL) in lingua straniera secondo la metodologia CLIL. La Direttiva introdotta dal Ministero della Pubblica Istruzione riguarda in particolare il riordino della scuola secondaria di secondo grado. La presentazione vuole offrire uno spunto di riflessione su come l'introduzione della metodologia CLIL nelle scuole assicuri un cambiamento in termini di qualità e arricchimento dei curriculum nazionali. Va da sé che, come tutte le nuove sperimentazioni, l'introduzione della CLIL favorirà nuovi metodi di apprendimento e nuove forme di valutazione degli studenti e degli insegnanti. Nelle scuole italiane, dal 1996, i programmi Cambridge International Examinations hanno favorito e facilitato il processo di come pensare CLIL e di guardare ad un innovativo sistema educativo. Le scuole statali in Italia hanno trovato che l'inserimento curricolare dei programmi Cambridge traducono l'essenza stessa del CLIL, con l'adozione di una metodologia

nuova e creativa chiamata “team teaching”. Il “team teaching” è stato ormai adottato da anni nelle scuole italiane per ovviare alle carenze linguistiche di alcuni insegnanti di materie non linguistiche. In questo modo le lezioni sono impartite da due insegnanti: da quello madre lingua, ed anche esperto della materia, nonché dall’insegnante titolare di cattedra. L’innovativo approccio metodologico di questi programmi ha interessato un numero sempre crescente di insegnanti nelle scuole Italiane e ha trovato larga corrispondenza negli studenti che testimoniano di guardare alle materie in modo più pratico e curioso.

**Alessandra Varriale**, Cambridge International Examinations, United Kingdom

Alessandra Varriale has been working with Cambridge International Examinations since 2014. She is in charge of building relationships with schools, key educational institutions and government authorities in various countries throughout Europe, as well as strengthening ties with our long-term partners in the region. She has worked for Bell Educational Services in the UK, as a Regional Sales Executive covering Europe and Latin America, where she managed and developed relationships with key stakeholders. Previously, she has also worked in Italy for a number of companies, including Deutsche Bank. Alessandra also spent some time as a freelance teacher of French and Italian in the UK and is fluent in French, Spanish and Italian.

**Gisella Langé**, Italian Ministry of Education, Italy

Gisella Langé is a Foreign Languages Inspector with the Italian Ministry of Education, advising on internationalisation and foreign languages. A researcher, writer and consultant on issues relating to Foreign Language teaching, curriculum development and intercultural education, she has vast experience of working on culture and language learning solutions and web-based teacher training. An expert within the European Commission and the Council of Europe, thematic groups, studies and projects she has been involved with include language policies, early language learning and bilingual education. She is currently involved in National Groups and Committees organised by the Ministry of Education on Primary teacher training, National Guidelines and CLIL.

## Penser l’organisation dématérialisée de tests de langue à grande échelle

À l’ère numérique, les tests au format papier/crayon apparaissent désuets, voire d’un autre âge. Les candidats et la société civile s’interrogent sur l’empreinte environnementale, les coûts et la perte de temps de l’impression papier et des envois postaux entre les centres de passation et l’organisme assurant l’édition et/ou la correction des épreuves. Le numérique offre également de nouvelles possibilités qui, tout intéressantes soient-elles, posent des défis concernant la validité de la mesure, comme la possibilité d’utiliser de nouveaux types d’exercice, la passation adaptative des tests, la saisie sur ordinateur des productions écrites, voire l’automatisation de leur correction... Le passage au numérique ne va par ailleurs pas de soi dans un contexte où un test est distribué à grande échelle via un réseau de centres de passation. Au-delà des problématiques de la conception, de la passation et de la correction des épreuves, c’est toute la chaîne administrative qui est concernée, de l’inscription des candidats à la délivrance des attestations de résultats. Il ne faut pas s’arrêter au milieu du chemin de la transformation numérique. De surcroît, les contraintes ne sont pas les mêmes selon qu’il s’agit de lancer un nouveau dispositif d’évaluation ou de digitaliser un dispositif existant, et si on souhaite faire cohabiter une version électronique et une version papier/crayon d’un même test en garantissant leur équivalence. Ces différentes contraintes doivent être prises en considération dans le design du système d’information. Dans cette communication, nous présenterons

l'expérience du Centre de langue française, qui offre depuis 2005 une version numérique du Test d'évaluation de français et a procédé ces dernières années à une refonte intégrale de son système d'information, pour disposer d'une solution évolutive et poussant la dématérialisation des opérations jusqu'au niveau des centres de passation.

**Dominique Casanov**, Centre de langue française, Chambre de commerce et d'industrie de région Paris Ile-de-France, France

Dominique Casanova is Head of the Scientific and Digital Unit of the Centre de langue française. He has been working on psychometric and IT issues regarding assessments since 2005 and conducted the complete renewal of the IT system in recent years. Dominique Casanova is an engineer in computer science and applied mathematics and he has managed a French Language Centre in Morocco for six years. He has published various papers on psychometric issues regarding the Test d'évaluation de français.

## Post-entry language assessment in higher education: The interaction between home and school language in relation to academic language proficiency

From 2010 onwards, all students of several KU Leuven faculties and colleges participate in a large-scale, low-stakes language test of academic Dutch at the start of the first year (De Wachter and Heeren 2013). The test consists of 25 to 28 selected-response vocabulary and reading items ( $\alpha=0.72$ ), lasts maximum 30 minutes and was constructed to detect possible at-risk students at an early stage (De Wachter and Heeren 2013). A predictive validity study on 2660 university students shows a significant positive correlation with study success (De Wachter et al 2013). This study uses the data of 3574 college and 6268 university students who were tested between 2011 and 2014. We have analysed the interaction between the language(s) students speak at home and their language of pre-university instruction in relation to their academic language test scores. A linear regression was run with the test score as response variable and as predictors: 1. home language, 2. whether or not there was a difference in home and schooling language, 3. gender, 4. whether they enrolled in a university study or a college programme and 5. the year the test was taken. The analysis shows that when Dutch-speaking students had a language of schooling different from their home language, their language test results were lower. However, when multilingual students received their education in Dutch, they had higher language test scores. The results of native speakers seem to follow the linguistic interdependence hypothesis (Cummins 1979). The reversed pattern of multilingual students appears not to follow Cummins' model but their results actually reflect different subgroups within the larger multilingual group. Although more research is needed, this is an interesting finding because these subgroups seem to have different levels of academic language proficiency, indicating a need to diversify teaching and support.

**Lieve De Wachter**, KU Leuven, Belgium

Lieve De Wachter teaches academic writing and presentation skills at the faculties of Social Sciences and Arts at the KU Leuven. She is promoter of several projects such as a digital platform for academic Dutch and a Dutch writing aid. She is also the head of the KU Leuven writing centre and her research topics are language testing and academic writing.

**Jordi Heeren**, KU Leuven, Belgium

Jordi Heeren teaches academic language skills at the KU Leuven. He develops several academic language supporting initiatives such as a digital learning environment and language workshops. His research topics are language testing and academic writing.

## Predicting readability of texts for Italian L2 students: A preliminary study

Text selection and comparability both for teaching and assessment purposes are central issues when addressing L2 students with texts to read and comprehend. In comparison to subjective selection, quantitative approaches supported by automated analysis can provide more objective information, analysing texts at language and discourse level (Khalifa and Weir 2009). Readability formula widely used in the past, such as the Flesch Reading Ease, the Flesch-Kincaid Grade Level and, for Italian, the GulpEase index (Lucisano and Piemontese 1988), have been criticised because they do not fully address the issue of text complexity. In the past decade, a new readability formula called Coh-Metrix was proposed (Crossley, Greenfield and McNamara 2008), taking into account a wider set of language and discourse features. More recently, a similar approach was proposed to assess readability of Italian texts through a tool called READ-IT (Dell'Orletta et al. 2011). While READ-IT was only tested on randomly selected newspaper texts, this contribution is focused on the development of a similar, computational tool to be applied on texts specifically selected in the context of assessing Italian as L2, identifying a set of linguistic features to be used in order to establish text difficulty. Two corpora of texts have been collected from the CELI item bank at B2 and C2 level. The statistical differences in the occurrence of the selected features have been analysed, belonging to four different categories: length features (at word and sentence level), lexical features (the percentage of basic vocabulary words, and the measure of lexical diversity), morpho-syntactic features (verbal tenses and moods, passive constructions, nominalisations, and the measure of lexical density), and discursive features (mechanisms of cohesion, intensification, mitigation and topicalisation). This contribution will show the study's preliminary results and how they can be applied both for teaching and assessment purposes.

**Giuliana Grego Bolli**, CVCL - University for Foreigners - Perugia, Italy

Prof Giuliana Grego Bolli is Full Professor in Linguistics at the University for Foreigners of Perugia. She is currently Director of CVCL (Centre for Language Assessment and Certification). Her research interests are related to the concept of quality in Language Testing and its application to the testing process and to the use of language tests within the migration context. She is author and co-author of more than forty publications in the area of language assessment.

**Stefania Spina**, University for Foreigners - Perugia, Italy

Stefania Spina is Associate Professor of Linguistics at the University for Foreigners Perugia (Italy). Her main research interests are Corpus Linguistics, collocations acquisition in Italian as a second language, and the language of media. Her research relies on corpus-based methodologies (she has created a reference corpus of Italian, the Perugia corpus) and on computational techniques, in the field of POS tagging and automatic extraction of collocations. She is also interested in linguistic changes produced by new media, such as social networks.

**Danilo Rini**, CVCL - University for Foreigners - Perugia, Italy

Danilo Rini joined the CVCL (Centro per la Valutazione e le Certificazioni Linguistiche) of the University for Foreigners of Perugia in 2005. He currently works on the development of a computer based test, as a test developer and reviewer, and on the quality assurance of CELI (Certificato di Lingua Italiana) exams.

## Principled development of a score reporting for young language learners tests based on research in psychology of teaching and learning

When considering the impact of assessment on teaching and learning, one critical but often neglected aspect of a test is the score reporting system. Within the context of Weir's (2005) socio-cognitive framework for language test validation, score reporting forms an essential facet of scoring validity; candidates' performance on a test should be reported in a manner which balances a variety of needs: consistency with the purpose of the test and the intended use of test scores, supporting the test design in terms of construct and measurement properties, and appropriacy of the reporting system for the stakeholders including, but not necessarily limited to, candidates and end users. This presentation first surveys general issues around the reporting of scores for young learners, for whom score reporting can be a particularly sensitive issue. The test scores for this group need to positively impact learning, as well as be useful and readily interpretable for a range of stakeholders including parents and schools, and the young learners themselves. Building on the theoretical discussion, the presentation discusses a case study of how these considerations informed and shaped the development of a revised score reporting system for Cambridge English: Young Learners tests. The presentation also describes the collaborative systematic methods used across 10 countries to develop a score reporting system that is both grounded in research on pedagogy and psychology of learning, and is comprehensible and useful to diverse learners globally.

**Maggie Dunlop**, Cambridge English Language Assessment, UK

Maggie Dunlop is a Senior Research Manager at Cambridge English Language Assessment. She researches the Cambridge English: Young Learners tests and learning-oriented language assessment, including integration of assessment, instruction and digital learning. Maggie also specialises in educational programme evaluation. Maggie holds an MA in International Education from the George Washington University, DC, USA and a PhD in Language and Literacies Education from the University of Toronto, Canada.

**Mark Elliott**, Cambridge English Language Assessment, UK

Mark Elliott is a Senior Validation & Data Services Manager at Cambridge English Language Assessment. He leads data analysis for validation and research of the Cambridge English: Young Learners tests and Cambridge Assessment Admissions Testing. He also specialises in test equating, pretesting, and grading procedures. He holds an MA in Mathematics from the University of Cambridge and an MA in ELT and Applied Linguistics from King's College London.

## Public examinations in Hong Kong: Stakeholder recognition and understanding

Hong Kong has a standards-referenced school examinations system and a single qualification, the Diploma of Secondary Education (HKDSE), serves as both a school-leaving qualification and as an examination for college entry. The Hong Kong Examinations and Assessment Authority, which runs the HKDSE, communicates the prevailing standards to stakeholders (parents, teachers, employers etc.) in various ways, such as by publishing Level Descriptors which describe the performance to be demonstrated to achieve each level (grade). It is important that the standards are understandable to stakeholders since high-stakes educational and employment decisions are based on them. This talk presents the findings of research into the understandability of the Level Descriptors for HKDSE English Language.

**Neil Drave**, Hong Kong Examinations and Assessment Authority, China

Neil Drave is the Senior Manager responsible for English subjects at the Hong Kong Examinations and Assessment Authority.

## QualiCEFR: A Quality Assurance template to achieve innovation and reform in language education through CEFR implementation

The CEFR has informed teaching, assessment and testing practices worldwide. Yet, its implementation is largely uninformed by Quality Assurance (QA) or impact studies. This has resulted in inconsistencies in CEFR implementation, which in turn have reduced the benefits in terms of innovation in language education policies and pedagogy. This presentation will report on QualiCEFR, a two-year international comparative research study funded by the Social Science and Humanities Research Council of Canada, integrating qualitative and quantitative research methods with a QA approach to inform and improve CEFR implementation. QA generally involves a systematic study of design and delivery processes, with each step in each process being defined with appropriate standards. QualiCEFR promotes a culture of QA in relation to course planning, teaching, and assessment by identifying promising implementation practices, and by providing guidance in the form of questions at each stage of the process. The project consists of two phases: firstly, a comparison between Switzerland and Canada, two multilingual countries with decentralised education systems, the former having been at the forefront of CEFR implementation, the latter in the earlier stages. The focus is on transparency and coherence in language curriculum reform and teacher development, appropriateness of QA procedures, and identification of successes and challenges in the CEFR implementation process. Over 40 interviews with key players in the two countries, together with a follow-up survey, have been conducted and thematically analysed using NVivo software. Informed by the results of the analysis, CEFR-related initiatives, promising practices and implementation outcomes that can be replicated and upscaled, have been identified. Subsequently they have been mapped to generic principles of quality to form the basis for a template, with procedures and indicators, to be used

by language educators and stakeholders to evaluate and profile strengths and weaknesses of practice in their contexts.

**Enrica Piccardo**, OISE – University of Toronto, Canada

Enrica Piccardo is Associate Professor at OISE – University of Toronto and at the Université Grenoble-Alpes France. She has extensive experience in language teaching, teacher education and second/foreign language education research. A CEFR specialist, she has coordinated two ECML CEFR-related projects and collaborated in the Council of Europe project to develop descriptors for mediation, online interaction and plurilingual/-cultural competences. Her other research focuses on complexity theory, emotions and creativity in language education, assessment and its role in the curriculum.

**Brian North**, Eaqals/Eurocentres, Switzerland

Brian North has extensively collaborated with the Council of Europe during his long career at Eurocentres. He developed the CEFR levels and illustrative descriptors, co-authored the CEFR itself, the prototype European Language Portfolio, the Manual for relating examinations to the CEFR, Eaqals' CEFR Core Inventories for English and French, and the European Profiling Grid for language teacher competences (EPG). His most recent project is the 2014-16 update and extension of the CEFR descriptors.

**Eleonora Maldina**, OISE – University of Toronto, Canada

Eleonora Maldina holds a Master of Arts degree in Second Language Education. She works at the University of Toronto as a research assistant on CEFR-related projects in the Department of Language Education, and as an English for Academic Purposes instructor in undergraduate and graduate programmes. She is also the co-founder of ScuolaScuola, an Italian language school in Toronto, where she develops and teaches Italian language and culture courses to adult learners.

## Reconsidering the impact of language assessment on language learning and teaching: A survey on an Italian examination for young learners

This presentation focuses on the impact of language assessment on language learning and teaching. It starts from the premise that in recent decades the presence of foreign children in the Italian educational system has caused an emergency situation in which the basic need is represented by learning Italian and developing suitable materials for the teaching and assessment of communicative language competence in Italian (ISTAT, 2014). As stated by Saville (2000), tests provided by examination boards may have an impact on educational processes and on society in general. As asserted by Shohamy (2011), the assessment policies and practices are mostly based on monolingual constructs whereby test-takers are expected to demonstrate their language proficiency in one language at a time. Starting from the last claims, we conducted a survey to reflect on the multilingual competencies of foreign young learners of Italian as a second language (ISL) and on their resulting linguistic, communicative and educational needs. We developed a detailed information questionnaire for students involved in the survey. It included traits focusing on standard demographic data and linguistic background. Then, two examination booklets were administered, respectively of Common European Framework of Reference for Languages (CEFR) levels A1 and A2, to a sample of young learners made up of ISL learners and Italian native speakers. This impact research intended to investigate the effect of the multilingual competencies and the maintenance of L1 of young learners of ISL in their performances. Secondly, the aim of the research is to reflect on the test construct validity and its appropriateness for its purpose and context of use (Messick 1996, Saville 2012).

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

**Paola Masillo**, University for Foreigners of Siena, Italy

Paola Masillo holds a PhD in Linguistics and Teaching of Italian as a foreign language from the University for Foreigners of Siena (Italy). Her PhD project focuses on validation analyses on language tests for migrants. She has an MA in Linguistics and Intercultural Communication, with a specialisation in language testing and assessment. Her main research focuses on language policies in education and migration contexts, educational linguistics, language testing and assessment. Currently, she is a Research Fellow at the University for Foreigners of Siena.

**Carla Bagna**, University for Foreigners of Siena, Italy

Prof Carla Bagna is Associate Professor in Educational Linguistics at the University for Foreigners of Siena. She is currently Director of CLUSS (Linguistic Centre for Italian Language). Her research interests are related to applied linguistics, language testing, sociolinguistics, and linguistic mediation. She is author and co-author of many articles in the area of language testing, language policy, languages in contact.

**Sabrina Machetti**, University for Foreigners of Siena, Italy

Sabrina Machetti is Associate Professor in Educational Linguistics at the University for Foreigners of Siena. She holds a PhD in Linguistics and Teaching of Italian to Foreigners from the University for Foreigners of Siena and is currently Director of CILS (Centre for the Certification of Italian as a Foreign Language). Her research interests are related to applied linguistics, language testing and certification. She is author and co-author of many publications in the area of language assessment.

## Reflective practice and professional development qualifications for teachers of bilingual learners

In recent years there has been a significant increase in the number of state and independent schools in Europe and beyond offering bilingual programmes and there is growing recognition of the benefits of studying school subjects through a second or additional language. However, in addition to the benefits of bilingual education, there are also challenges. Teachers may feel anxious about introducing an innovation such as bilingual education and may have a lot of questions about how best to proceed. They may lack confidence about introducing a more complex curriculum and feel they lack sufficient resources and support. Cambridge International Examinations fully understands these concerns and has long experience of producing bilingual education resources and professional development programmes for teachers. This presentation will demonstrate how the Cambridge Professional Development Qualifications (PDQs) provide a strong framework to support the effective continuing professional development of teachers and school leaders. In particular, the PDQ in Teaching Bilingual Learners invites teachers to reflect more deeply and critically on their teaching and learning practice, as well as analyse the language demands of their curricular subject. It demonstrates how to engage with a model of integrating content and language learning and how to apply their developing knowledge and skills in designing and managing a sequence of learning for bilingual learners.

**Martin Nuttall**, Cambridge International Examinations, United Kingdom

Martin Nuttall is Regional Director, Europe at Cambridge International Examinations. He has extensive experience in education, first as a teacher and lecturer in English Language and Linguistics at schools and universities in Europe and Japan, and more recently in a variety of roles for Cambridge Assessment, including within the ALTE Secretariat. Martin has an MA in Applied Linguistics from the University of Essex and a degree in Modern Languages from the University of Cambridge.

## Spanish for business in language accreditation

Although we cannot point out an exact date, we must go back to the eighties of the last century to find some demand for teaching Spanish for specific purposes (EFE). From business spheres, academic and commercial institutions (University Chambers of Commerce) began to develop their own certifications to accredit knowledge, especially in the business world. It created the need to prepare candidates for such tests and, therefore, it was also in those years when specific materials in this field started being published. Furthermore, if we add the incorporation of Spain to the European Economic Community in 1986 and the liberalisation of energy and media companies, as well as of banks, in Latin America during the nineties, we witnessed a growth in the demand for EFE that will no longer decline over the years. There is no doubt that, given the immense competition in which the labor market moves, with a fierce fight for every job, the knowledge of languages for professional purposes is not only necessary for the very essence of the work that will be done, but also for the need to be distinguished, to excel in a market saturated with professionals in which companies and institutions have to select candidates and, to this end, need solid arguments that support recruitment. That is, they need knowledge and accreditation. To meet the current demand of this type of certification for the Spanish, we have some models that will be discussed in the presentation.

**Marta García**, Universidad de Salamanca, Spain

Marta García holds a PhD in Assessment and a Master's Degree in teaching Spanish as Foreign Language, both from the University of Salamanca. She has extensive experience of teaching Spanish as a foreign language. Aside from teaching, she has also contributed her expertise to the production, development, revision and assessment of the Diploma de Español como Lengua Extranjera (DELE) tests and is currently the Academic Coordinator of the Business Language Testing (BULATS) Spanish tests. She has participated in ALTE projects and is currently working in the ALTE Language for Specific Purposes (LSP) group. Her main research interests include non-verbal communication in the teaching of Spanish and Assessment for LSP.

## Standard valutativi e didattici per l'italiano nel mondo: le attività della Società Dante Alighieri

La Società Dante Alighieri, che ha come obiettivo la valorizzazione della lingua e della cultura italiane nel mondo, è composta da una rete di quasi 500 Comitati disseminati in più di 80 Paesi; nel 2015 all'interno della rete Dante Alighieri sono stati organizzati quasi 10.000 corsi di lingua italiana per 199.105 studenti. Attraverso la certificazione PLIDA, inoltre, la SDA offre ogni anno a circa 10.000 persone la possibilità di attestare il proprio livello di competenza in lingua italiana grazie a 320 Centri certificatori PLIDA autorizzati, oltre che fra i Comitati SDA, anche in Università italiane ed estere, Istituti Italiani di Cultura, CTP/CPIA e scuole di lingua. All'interno di una realtà così estesa e complessa è nata negli ultimi anni l'esigenza di condividere standard e buone pratiche legati alle attività di insegnamento e valutazione linguistica. Attraverso il Progetto Lingua, la SDA ha deciso di dare una risposta a questa esigenza progettando un piano di interventi mirati negli ambiti della didattica, della certificazione e della formazione dei docenti. Nel corso della comunicazione verranno presentati i progetti realizzati e quelli in via di sviluppo;

ci si concentrerà in particolare su ADA, il sillabo per la programmazione didattica della SDA, sulle attività di formazione in presenza e a distanza dei docenti e dei commissari d'esame e sul piano di revisione e di aggiornamento della certificazione PLIDA. Verranno quindi proposte delle riflessioni su come l'attività certificatoria influisca sulla formazione continua dei docenti e su quali siano gli strumenti più efficaci per condividere uno standard di valutazione all'interno della rete dei Centri certificatori.

**Giammarco Cardillo**, Società Dante Alighieri, Italy

Giammarco Cardillo graduated in Linguistics at the "Sapienza" University of Rome. He is the supervisor of the Progetto Lingua Italiana Dante Alighieri (PLIDA) examining plan within Società Dante Alighieri (SDA) and he is also involved in training projects for teachers, item writers, examiners and raters. He co-wrote the Piano dei Corsi ADA (the teaching syllabus of SDA) and is co-author of Italian grammar manuals for high school.

**Paola Vecchio**, Società Dante Alighieri, Italy

Paola Vecchio obtained her PhD in Linguistics and History of Italian Language from the "Sapienza" University of Rome researching the syntax of Italian dialects. She has gained long-standing experience in teaching Italian as a second language, and working in Italy and worldwide as teacher, trainer, author and test maker. At Società Dante Alighieri (SDA), she is the supervisor of item writers' training and currently manages the Progetto Lingua Italiana Dante Alighieri (PLIDA) syllabus and exam formats upgrading process.

## Student perceptions of the CEFR levels and their ability to rate their speaking in English

Students learning English in Spain often refer to their levels in the target language as A2, B1 or B2, etc., but when they are asked to indicate what these levels of the Common European Framework of Reference for Languages (CEFR) actually mean, many of them are at a loss as to what is involved in one level compared to the others. This lack of knowledge can lead to insecurity and a lack of direction when they are required to take a test in order to certify a specific level for academic or professional reasons, a situation which is increasingly common these days. Nevertheless, little research appears to have been done to determine student understanding of the CEFR levels. Hence, this exploratory study seeks to examine the perceptions and prior instruction on the CEFR levels of a group of 70 university students who were enrolled in programs in English, Translation or Education. As part of a larger study on the Aptis Test funded by the British Council, the students were interviewed in pairs and then answered a short questionnaire asking for their impressions of what is entailed in having a B1 and a B2 level of English. Several weeks later they observed a video recording of their initial interview and rated themselves on various aspects of speaking in comparison with some sample Cambridge exam videos available online. Finally, after taking the Aptis Test, they were asked if the process had helped them to better understand their levels on the CEFR. This talk will present the results, which can serve to inform teachers so that they can aid students in better understanding the CEFR levels. As a result, the test takers can begin to visualise what they need to do to perform better on exams.

**Mary Frances Litzler**, Universidad de Alcalá, Spain

Mary Frances Litzler, PhD, is an Assistant Professor of English at Universidad de Alcalá in Spain, where she teaches English language and linguistics. Her research interests and publications are related to English language testing, language learning activities undertaken by students after their classes, and learner autonomy. She has published in *Studies in Language Testing (SILT)*, *Journal of Language Teaching and Research*, *Educational Technology and Society*, and *Theory and Practice in Language Studies*, among others.

## Students and instructors' perceptions of the construct-(ir)relevance of language to literacy competence in testing literature: A work in progress

Unlike in the first language, one of the biggest dilemmas in testing literature in undergraduate English literature programmes in English as a Foreign Language (EFL) contexts is the extent to which language should constitute part of the target construct intended to be measured. There are some who argue that language is construct-irrelevant and should be minimised in favour of a focus on literary competence. In practice, however, this does not seem to be the case as language proficiency considerations seem to creep into assessment and colour professors' judgment of student work. The current study seeks to examine instructors' and students' perceptions of the degree to which language proficiency forms part of the construct of testing literature in two departments of English and Persian literature. To this end, a total of fifty students in two poetry courses, one in the English department and the other in the Persian, were invited to respond to a questionnaire designed to gauge their perceptions of the extent to which language forms, or should form, part of the construct of literary competence. The views of six literature professors on the same issue were also elicited via semi-structured interviews. We are in the process of analysing quantitative and qualitative data via inferential statistics and inductive coding. We believe that given the fact that testing literature is far behind the current state of the art in language assessment, this comparative study can expand our understanding of the construct definition in literature testing in both first and foreign language. The study carries implications for high-stakes external literature exams as well as for achievement testing in undergraduate English literature programmes in EFL contexts.

**Seyed Rahim Moosavinia**, Shahid Chamran University of Ahvaz, Iran

Seyed Rahim Moosavinia is an associate professor of English literature in the department of English language and literature, Shahid Chamran University of Ahvaz, Iran.

**Kioumars Razavipour**, Shahid Chamran University of Ahvaz, Iran

Kioumars Razavipour is an assistant professor of language testing and assessment in the department of English language and literature, Shahid Chamran University of Ahvaz, Iran.

## Teaching Italian as a second language to migrants. Mixed competence levels and linguistic backgrounds in the same classroom: A challenge

Language competence is no doubt a tool for migrants which can contribute to their better integration, provide better job opportunities as well as more satisfactory relationships in their life environment. That's why many NGOs who deal with migrants and refugees try to contact them to offer language courses in the early period of their arrival to the host country. In Italy, Germany, Belgium and Spain, the four European countries where the Community of Sant'Egidio has developed its services for migrants and refugees, they can obtain free language courses in the first months after they have reached the country. The classes are formed by people with various linguistic and cultural backgrounds, different competence levels in their native language, even no competence in written production or illiteracy. As their first need is to communicate in the target language, this is taught through a learner-oriented communicative approach, in which the everyday life situations are vehicles to introduce and practice language items. No need of grammar structures in the first year of course, which provide up to 100 hours of teaching, and allows the majority of students to reach level A1+ at least in oral interaction. It is only after they are able to communicate at basic level that support for poor or illiterate readers is provided with extra parallel courses, in which little groups of learners with more uniform linguistic background or competence levels are taught specific issues, such as reading or writing in the target language. Thus, having reached a satisfactory ability of listening comprehension, students are able to understand messages about history or culture of European countries, through class conferences, frontal lessons, and class activities or discussions. It is mainly important to the role of the personal relationship between teacher and students and among the students in class, a real educational community.

**Cecilia Pani**, Comunità di Sant'Egidio, Italy

Cecilia Pani was born in Rome in 1959. Graduated in Modern Languages and Literatures, she teaches English in lower secondary school in Rome. She has co-ordinated the Italian Language and Culture free courses of Sant'Egidio Community in Rome since 1991 and supports integration of foreign students in Italian state schools. She published as co-author two volumes of Italian L2 manual "L'italiano per amico". She has participated in seminars and international conferences on language teaching and on integration of migrants in European countries. .

## Test takers' attitudes to different online speaking assessment formats

This presentation draws on the ongoing research about test takers' (TT) attitudes to different online speaking assessment formats. The paper gives a brief account of the state of the art online speaking assessment and proceeds to raise the issues like: What are test takers' attitudes towards online speaking formats? What interlocutors allow better TT performance? Although still an ongoing research, the paper provides the findings from the pilot study and attempts to put forward directions on what interlocutors (human/non-human) would better interact with TT.

**José F L Pascoal**, CAPLE - University of Lisbon (UL), School of Arts and Humanities (SAH), Portugal

José F L Pascoal is CAPLE Deputy Director and a Senior Lecturer in the Department of General and Romance Linguistics of the UL SAH. He obtained his master's degree in Portuguese Descriptive Linguistics/Applied Linguistics from the UL SAH with a dissertation on test construction and language frameworks and is currently developing his PhD research. He authored a version of the Vantage Level for Portuguese and co-authored the Threshold Level for Portuguese, national syllabi for Portuguese for school years 10-11-12. He also co-authored text books, books on exams to access Portuguese in education systems and assessment organisations worldwide. He is also a teacher trainer.

## Test takers' views and feedback on the L2 Sami language and the language test

This paper presents a study of test-takers' views and written feedback on learning North Sami and the L2 North Sámi language test with implications to language teaching and test development. With 1957 native speakers in Finland (Statistics Finland 2015), North Sami has a national language status in the Sami Homeland. In Finland, also in the Sami Homeland, Finnish and Swedish have the official language status. Testing North Sami as L2 evolved after the Sámi Language Act (2003) which stated that the Sami have the right to use their own language in official interactions. Test taker profiles indicate that there is a need to show North Sámi skills particularly in the candidates' professional fields such as teaching.

In addition to the feedback received for test development, the purpose of the study was to find out the language learner's views: their beliefs on and attitudes towards the language and its use and meaning in the candidates' everyday life. During one test administration, the candidates (N=20) were asked to fill in a questionnaire which focused on language use, difficulties in language learning and the definition of and feedback on language knowledge and skills in the four subskills used in the test. The candidates were further asked for their views on the future of the language. The data was analysed using content analysis for categorising meanings expressed in the questionnaire. The results of the questionnaire indicate e.g. that the language is regarded as significant in representing identity or language of working life and that the use of the language in everyday life was according to some candidates restricted to working life alone. Test feedback suggests that language learners were familiar with the overall test structure and that problems during the test mainly concerned timing in listening.

**Sari Ahola**, Centre for Applied Language Studies, University of Jyväskylä, Finland

Sari Ahola has a Licentiate in Applied Linguistics and an MA in German, Russian and Finnish. She is a research coordinator of the National Certificates of Language Proficiency (since 2014) at the University of Jyväskylä. She worked as a researcher in the test system between 2005–2013 and furthermore has over 10 years of experience in language teaching in the Finnish education system. Her research interests include test development in general, with a special focus on the item writing process, rater training, raters' views on language and candidate participation in the Finnish intermediate level test for acquiring Finnish citizenship.

**Henna Tossavainen**, Centre for Applied Language Studies, University of Jyväskylä, Finland

Henna Tossavainen has an MA in English and Finnish languages. She has approx. 20 years of language teaching and assessment experience on various levels in the Finnish education system, mainly at the university

of Jyväskylä. She is currently involved in the test development of the L2 English, Finnish and Northern Sámi language tests in the Finnish National Certificates. Her research interests relate to the language testing processes, the CEFR and particularly to the development of L2 writing at various levels.

## Testing pre-service teachers' spoken English proficiency: Design, washback and impact

Assessment plays a crucial role in a globalised economy and it has an effect on teacher recruitment, training and professional development. In some contexts, a high proficiency in English is considered fundamental in order for teachers to be guaranteed employment. This paper discusses the recent introduction of the Spoken English Proficiency Test for Teachers (SEPTT) in Malta. By means of this test, the regulator of the English Language Teaching industry in the country is seeking to ensure high levels of spoken English proficiency amongst pre-service teachers. SEPTT tests candidates' ability to use spoken English for a variety of functions, including conversing, explaining, presenting information, and giving feedback in a context specific to teaching. The paper shows how the test is based on the idea that people require different proficiencies that are always situated in particular contexts and bounded by a particular social practice (Freeman et al 2015). Just as general English proficiency cannot address all the linguistic needs of students in the world beyond the classroom, it cannot fully specify the demands on teachers' use of language inside the classroom when teaching the language. Hence, an English for Specific Purposes (ESP)-derived approach to language proficiency is required whereby teacher education focuses on the specific linguistic needs of teachers when enacting their role. This paper describes the research that went into SEPTT's design and implementation, and evaluates the test's washback and impact. Due to SEPTT's introduction as part of a set of legal requirements for pre-service teachers, the research presented in this paper will resonate with language testers and educators working in contexts where assessment has a pivotal role within national policy development.

**Daniel Xerri**, University of Malta, Malta

Daniel Xerri is a lecturer in TESOL at the University of Malta, the joint co-ordinator of the IATEFL Research SIG, and the chairperson of the ELT Council within the Ministry for Education and Employment in Malta. He holds postgraduate degrees in English and Applied Linguistics, as well as a PhD in Education from the University of York. Further details about his talks and publications can be found at: [www.danielxerri.com](http://www.danielxerri.com)

**Odette Vassallo**, University of Malta, Malta

Odette Vassallo is the director of the Centre for English Language Proficiency and a senior lecturer in English language and linguistics at the Department of English, University of Malta. She has a teaching and research interest in applied linguistics and holds a PhD in Linguistics from the University of Nottingham.

**Sarah Grech**, University of Malta, Malta

Sarah Grech lectures on the MA TESOL programme at the University of Malta, as well as phonetics, phonology and language variation on a range of undergraduate and postgraduate programmes. She is a linguist with research interests in language variation and English language proficiency.

## The Cambridge English Global Analysis: Understanding English proficiency worldwide

The Cambridge English Global Analysis, recently published online, summarises candidate performance for Cambridge English exams in 71 countries or territories. Based on the results of more than a million candidates taking high quality exams, the Global Analysis offers a tool to help understand the nature of language competence within countries, territories and regions, furnish comparisons between them, and track changes over time. The Global Analysis has already attracted interest around the world (Brasser 2016), and, as this interest suggests, information about relative English language proficiency is in much demand, probably because it is seen as an important indicator of competitiveness in our globalised economy, and something which should inform education policy. There is, however, a temptation to overinterpret such information, downplaying the importance of test quality, representativeness of the sampling, or other methodological aspects which underpin it. Tendencies to do just this are evidenced in a range of respected news titles (New York Times: Schuetze 2013; The Economist 2012; The Financial Times: Skapinker 2015) and elsewhere. This presentation will describe the approach behind the construction of the Cambridge English Global Analysis, and consequent implications for its use, such as what the appellation of Common European Framework of Reference for Languages (CEFR) levels to countries should be taken to mean and to what extent conclusions may be drawn about wider populations. In addition, Cambridge English's approach to accurately benchmarking country performance will be outlined and it will be explained how this is used to supplement the core Cambridge English Global Analysis.

**Michael Corrigan**, Cambridge English Language Assessment, United Kingdom

Michael currently works as Senior Validation Manager for Cambridge English Language Assessment, working on a range of exams and projects. Between 2005 and 2013, his role was ALTE Validation Manager, collaborating with members of ALTE on the statistical analysis of their test data. Michael holds a PhD in language testing from the University of Bedfordshire.

**Andrew Coombe**, Cambridge English Language Assessment, United Kingdom

Andrew Coombe is Global Head of Marketing at Cambridge English. He read Modern Languages at Oxford University and gained an MBA from INSEAD, Fontainebleau, France. He has worked in Marketing, Sales, Finance and Strategy roles in business for UBS, Ford and Volkswagen; he has lived and worked in France, Switzerland and Germany. He is also the Chairman of the Alliance Française de Cambridge.

## The construction and validation of China Standards of English-Speaking: Principles, procedures and progress

As one component of the China Standards of English (CSE), the CSE-Speaking (CSES) was constructed recently with a view to promoting the teaching, learning and assessment of oral English across all educational levels in China. This presentation will introduce the principles and procedures followed in developing the CSES and report the progress of the project. Based on an analysis of the status quo of oral English teaching, learning and assessment in China, the project team proposed three fundamental principles of scale construction. First, the primary function of

the CSES would be to serve the needs of English language education in China, and therefore, voices of teachers and learners, the major groups of stakeholders, should be heard and their needs met. Second, Can Do descriptors would be used to illustrate the development of oral English proficiency. These descriptors should be informative, user-friendly and user-oriented. Third, special attention should be paid to the salient features that learners at each stage exhibit in performing oral English communicative activities. The development of the CSES has gone through three stages. First, theoretical and empirical studies of oral English proficiency were reviewed and an operational definition of oral English abilities was proposed. Second, an initial descriptor pool was established based on a thorough review of documents, textbooks and tests, which contained about 4,263 descriptors of oral English proficiency. Third, the descriptors were selected, revised and translated. Teachers' views were elicited through workshops on the clarity, relevance and applicability of the descriptors to oral English teaching and learning. Descriptors were also produced for those aspects of proficiency which were under-represented. A pool of 568 descriptors in 22 sub-categories at 9 provisional levels has been established and validated qualitatively. Large-scale surveys have also been conducted among stakeholders and the data will be analysed quantitatively to scale the descriptors.

**Yan Jin**, Shanghai Jiao Tong University, China

Yan Jin is a professor of Applied Linguistics at the School of Foreign Languages at the Shanghai Jiao Tong University.

**Wei Jie**, Shanghai Jiao Tong University, China

Wei Jie is a PhD candidate in language testing and assessment at the Shanghai Jiao Tong University.

## The development of diagnostic assessment system for senior high schools in China

The Diagnostic assessment system for Senior High (UZX) is a web-based English diagnostic testing and learning system, following the process of “diagnose—discover problem/give advice—provide focused exercise”. Namely, teachers and students can know the strengths and weaknesses of the subskills with reference to their learning targets, then receiving the recommended exercises automatically to improve the weaknesses. Using Rasch measurement techniques, the system allows students to track their English language development as each time they take the Assessment they receive a combination of texts and items that are targeted to their proficiency level. This paper reports on the development of the UZX as an instrument that assesses Chinese Senior High students' English learning and on the results of the initial administration in four high schools (more than 1200 students in about 20 classes) in 2016.

**Liping Liu**, Foreign Language Teaching and Research Press, Beijing, China

Liping Liu works as the Vice Director of Validation and Assessment, R & D Center, Foreign Language Teaching and Research Press. She is the core member of the Diagnostic Assessment projects team in the Press.

**Zunmin Wu**, Beijing Normal University, China

Zunmin Wu is a professor in the School of Foreign Languages and Literature, Beijing Normal University where she founded the Centre for Foreign Language Testing and Assessment. Until now, the Center has undertaken many large assessment projects from the Ministry of Education, the School of Foreign Languages and Literature, Foreign Language Teaching and Research Press, etc.

## The evaluation of Chinese students of Italian L2: Practices at the Universities for Foreigners of Perugia and Siena

The assessment of Chinese students' proficiency in Italian, an issue concerning a lot of institutions, be they universities or not, in order to let the students enroll in undergraduate or high specialisation courses, should not be unique or anomalous, whereas it should be framed within the tensions which have characterised the scientific debate on language testing. These tensions arise from the positive and negative consequences of language testing, and in particular of the standardised assessment of language certification exams, on Italian L2 teaching and learning, with specific reference to Chinese students, who belong to a culture that emphasises the outcomes of school performance and takes into great account the results of the exams for their potential positive impact on the educational context and on society itself. As a result, the institutions dealing with language certification need to provide evidence of the exam's evaluation process, to guarantee that the results and their consequent use are valid and reliable. From this perspective, the paper shows the results of the administration of A2 level (CELI and CILS) certification tests to Chinese students at two Chinese Universities where Italian is taught: the University of Nanjing, where the CELI 1 was administered, and the University of Hebei, where the CILS A2 was administered. The paper provides a comparison between the results obtained in the two exams in China and those obtained by the same students in Italy after an eight-month period, highlighting strengths and weaknesses of a path that should allow Chinese students to enroll in a university career.

**Giuliana Grego Bolli**, University for Foreigners of Perugia, Italy

Giuliana Grego Bolli is Full Professor in Linguistics at the University for Foreigners of Perugia. She is currently Director of CVCL (Centre for Language Assessment and Certification). Her research interests are related to the concept of quality in language testing and its application to the testing process and to the use of language tests within the migration context. She is author and co-author of more than forty publications in the area of language assessment.

**Sabrina Machetti**, University for Foreigners of Siena, Italy

Sabrina Machetti is Associate Professor in Educational Linguistics at the University for Foreigners of Siena. She has a PhD in Linguistics and Teaching of Italian to Foreigners at the University for Foreigners of Siena and is currently Director of CILS (Centre for the Certification of Italian as a Foreign Language). Her research interests are related to the applied linguistics, language testing and certification. She is author and co-author of many publications in the area of language assessment.

## The generation of an individualised Cognitive Diagnostic Report for College English writing

Cognitive Diagnostic Assessment (CDA) is specifically designed to measure a student's knowledge structures and processing skills by integrating cognitive psychology and educational measurement. Compared with conventional educational tests, the advantage of CDA lies in its ability to provide personalised diagnosis and guidance. The results of a CDA yield a profile with specific information about a student's cognitive strengths and weaknesses, which has the potential to guide instructors, parents, and students in their teaching and learning processes. However, current research in CDA is extremely limited and focused on listening and reading. Researchers have focused less on cognitive diagnostic writing assessment, not to mention the cognitive diagnostic feedback for English writing. Among the studies in English writing feedback, teacher feedback and peer feedback have been dominant, while the finer grained cognitive diagnostic feedback hasn't been generated and evaluated yet. Based on the merits and shortcomings of the previous research in the field of cognitive diagnostic assessment and English writing feedback, this study is intended to probe how to generate an individualised cognitive diagnostic report of English writing. The Reduced RUM available in the software R is adopted as a major research tool to get the students' mastery profile of each attribute by inputting data of the Q matrix and test-takers' performance. The study follows the framework presented by Mary R. Roberts and Mark J. Gierl (2010) with the purpose of developing score reports for cognitive diagnostic writing assessments. The individualised cognitive diagnostic report will provide learners with a total score, a specific attribute mastery profile, strategies for improving the weaknesses and the wrong items rated by the checklist, which would be beneficial to both students and teachers in English writing. Therefore, the study will make a certain contribution in the field of CDA and English writing.

**Tan Yandan**, Xi'an Jiaotong University, China

Tan Yandan is a graduate student in the Schools of Foreign Studies at Xi'an Jiaotong University, majoring in SLA and Language Teaching. She is devoted to applying the Cognitive Diagnostic Assessment to teaching and learning in terms of English writing. She was a volunteer for the PROMS 2016. She has published two articles in journals: A Survey on Learning Motivations of Non-English Majors in Agricultural and Forestry Universities; A Review of Chinese Studies on Personalized Learning in the 21st Century.

**Ma Xiaomei**, Xi'an Jiaotong University, China

Ma Xiaomei is a professor and PhD supervisor of Applied Linguistics in School of Foreign Studies at Xi'an Jiaotong University. She also works as Executive director of the China Institute of Educational Linguistics. She devotes herself to the field of personalised English learning, Language assessment, and online diagnostic assessment. She has accomplished twelve research projects, published nearly twenty research papers, and has given speeches in International or Domestic Conferences nine times, including one keynote speech.

## The impact of an integrated teaching, learning and assessment framework on students' performance and perceptions

In order to overcome the limitations of summative assessment, this paper attempts to propose an integrated teaching, learning and assessment framework and further explore its impact on students' performance and perceptions. A mixed-method case study research design in a longitudinal manner was employed to study how twenty-five students learnt to write in English in a naturalistic classroom setting over a semester. The study was conducted with a triangulation of data collection and analysis: multiple sources of data included pre-and post-course writing tests, student writing samples, reflective journals, questionnaires, interviews, and classroom observation; multiple data analysis included content analysis, frequency analysis, textual analysis, t-test analysis and ANOVA. The major findings are: through the analysis of students' writing performance, the participating students wrote an average of 6 drafts for one task, revised more by themselves with the help of AWE software and peers, paid more attention to language use after receiving multiple-sourced feedback, and improved the quality of their writing to different degrees between, within and across tasks/drafts as well as the pre-test and post-test. Through the analysis of students' perceptions of writing learning, results show that due to more interactions, their attitudes towards writing and revising changed, they self-reported writing progress, and writing confidence was enhanced; The contribution of the current research can be summarised as follows: (1) the integrated framework was put forward to guide teaching and assessment; (2) the research results contributed to the classroom assessment literature theoretically and practically; (3) the integration of written feedback and recorded oral feedback from teacher, peer feedback and AWE feedback was successful carried out in the study; (4) this study enriched assessment research on writing in the EFL context of Mainland China.

**Huang Jing**, China West Normal University, China

Huang Jing is currently a lecturer in China West Normal University with 18 years of teaching and research experience. She got her PhD degree in language testing from Tsinghua University in China in January 2014. On a scholarship from Chinese Scholarship Committee, she has worked as a visiting scholar under the supervision of Prof. Paul Nation and Dr. Peter Gu in Victoria University of Wellington in New Zealand from September 2014 to February 2015.

## The impact of online teaching practices on Greek EFL learners' reading perceptions & exam performance

This paper reports on a one-year longitudinal study that adopted a blended teaching approach based on designing and implementing an online EFL course to be used by Greek students aged 13-14 years old along their more traditional face-to-face lessons. The reason for creating a more dynamic learning environment aligned with the rest of the curriculum was to increase EFL learners' engagement and motivation through their exposure to authentic online material and participation in a variety of reading, writing, speaking and listening tasks. To this end, a number of online activities were designed including: a) an online classroom with hand-

outs, extra activities, resources and discussion groups for students to further develop their digital literacy along with their English language competence, b) a wiki for students to make a contribution and post their own messages on a specific topic, c) a series of Skype group discussions with invited external guest speakers, d) a private YouTube space for students to upload their videos and watch relevant EFL material. Data analysis of pre- and post-achievement tests on English language reading comprehension performance along with students' Computer-Assisted Language Learning (CALL) Attitude questionnaire showed that, in general, participants in this study had a positive attitude toward CALL while, at the same time, open online access technologies gave them the opportunity to further develop their EFL reading comprehension skills. The paper concludes by highlighting the fact that online class components were not designed around the tools, attempting to fit the online tools into a task-based EFL lesson, but rather served the learning objectives of the actual lesson based on a blended teaching approach, in which face-to-face and online learning activities were relevant to and complemented one another.

**Trisevgeni Liontou**, Ministry of Education, Greece

Jenny Liontou works as an Adjunct Instructor at the Hellenic Open University and at the Faculty of English Studies, University of Athens. She holds a PhD in English Language and Linguistics with specialisation in EFL Testing & Assessment. She holds a BA in English Language & Literature, an MA in Lexicography: Theory and Applications and an MSc in Information Technology in Education. She has worked as an EFL teacher in primary and secondary education, a freelance item writer, oral examiner and script rater for various international EFL examination boards and has published papers in the aforementioned fields.

## **The implementation of a French language certification: Positive washback and wider resulting effects**

Since the generalisation of language teaching after the licence-master-doctorat (LMD) reform which underlined the necessity to improve the students' language level and allow their mobility, language assessment has become an issue in French universities and raised questions which had hardly ever been crucial before: how can we train our future European citizens to become multilingual, and later, productive in a globalised economy? Who will train them? How do we certify the students' levels of competency? Which level of the Common European Framework of Reference for Languages (CEFR) should be targeted? Which tools could help them improve their level? How much will it cost? Our presentation will show how the implementation of the Higher Education Language Certification (CLES) at the University of Corsica resulted in improvements for the teachers, for the students and for the institution, and allowed for the elaboration of a language policy for students belonging to what we call the LANSAD (Language for students of other disciplines) sector. Our approach is that of complexity, as presented by Morin (1988), in which change arises from the actors involved in the certification, following "enovation" dynamics, a term introduced by G  linas & Fortin (1996), in opposition to the "innovation" concept. Our theory is that assessment via the CLES positively impacted teaching, since it introduced the CEFR and the action-oriented approach, as well as language learning,

since the needs of the individual were taken into account in a now massified landscape with the building of a language centre. It also empowered the teachers, belonging now to teams working together during interlanguage meetings, all involved in the conception process and for some of them, integrated in the national team, as well as the institution, which gained in quality (language policy implemented) and readability (results-oriented evaluation). We will thus describe the CLES system and explain what we identified as being a lever of change and good practice.

**Stéphanie McGaw**, University of Corsica, France

Stéphanie McGaw teaches linguistics at the University of Corsica. She is completing her PhD which deals with language policies in higher education, and which is entitled: "Language certification, the enlightening prism of a multicompetence serving political, economic and social stakes". She is in charge of her institution's language policy and has been a Vice-president for International Relations for nine years (2007-2016).

## The use of test taker productions in redesigning writing assessment grids: A corpus based study

If assessment is to be seen as an organic part of the learning cycle, resulting not only in reported scores, but also in useful feedback and washback, the teachers need to be skilled and thorough assessors, benefitting of expert constructed evaluation instruments. Starting from these premises, the research team at the Department of Romanian language, culture and civilisation in Babeş-Bolyai University, Cluj-Napoca, designed new writing assessment grids, in close connection with the CEFR. After successive evaluation sessions, feedback on the adequacy of the grids was collected from the assessors. In addition, input from test takers was used for improving the grids. We built a corpus of test written productions and, based on analysis of data from the corpus, we modified our assessment grid, submitting the initial measurement-driven instrument to a performance-driven "test". We present here the results of our study and the consequent modifications of the grid focusing on the criterion of accuracy at level A1. The changes are connected to: remodelling of the assessment process from grading together both productions of the test takers to discrete assessment of their texts; clearer band descriptions, based on findings from corpus analysis (e.g. what exactly 'long' in 'long sentences' means or how you understand 'small' in 'uses adequate vocabulary with small inadvertencies'); improved possibility of discrimination between the grid bands; and better calibration in grading between the criterion of accuracy and the other criteria in the grid.

**Dina Vilcu**, Babeş-Bolyai University Cluj-Napoca, Romania

Dina Vilcu is a lecturer at the Department of Romanian language, culture and civilisation, Babeş-Bolyai University, Cluj-Napoca, the institution where she also obtained her PhD in philology. She teaches and assesses Romanian as a foreign language, focusing, in her research, on theoretic linguistics, on the one hand, and on language assessment, on the other. Her main research interests are related to the validation of assessment for less tested languages and the use of the CEFR in different evaluation contexts.

**Lavinia-Iunia VasIU**, Babeş-Bolyai University Cluj-Napoca, Romania

Lavinia-Iunia VasIU is junior lecturer at the Department of Romanian language, culture and civilization at the Faculty of Letters (Babes-Bolyai University of Cluj-Napoca) where she teaches and assesses Romanian as a foreign language. She is interested in and has published in the fields of second language acquisition, pedagogy

and assessment, Romanian as a second and a foreign language, teacher training, quality management and corpus linguistics.

**Antonela Arieșan**, Babeș-Bolyai University Cluj-Napoca, Romania

Antonela Arieșan is a junior lecturer at the Department of Romanian language, culture and civilisation, Babeș-Bolyai University in Cluj-Napoca, Romania, where she teaches and assesses Romanian as a foreign language. Her research is centred mainly to the field of Romanian as a foreign language, concerning matters of teaching, and evaluating the vocabulary.

## Towards a scale of academic language proficiency

With increasing use of the Common European Framework of Reference for Languages (CEFR) in international school contexts and with the blurring of first, second or foreign language distinctions, we believe that there is a potential need for a supplementary scale of academic language proficiency. Many educators often tend to focus on the listening, speaking, reading and writing scales of the CEFR or on the global scale, which only touches on academic contexts as it has to encapsulate other contexts such as a social or foreign language. The CEFR provides a wealth of specialised scales – for example, the text processing scale, the pragmatic scales (thematic development, propositional precision, coherence and cohesion), the strategic scale of compensating, as well as tables that combine scales which draw together ‘relevant qualitative factors for production’ (Council of Europe 2009:149). However, aspects of academic language are found across various scales, which can make it hard to locate and apply them to school contexts. Furthermore, it is often assumed that academic language proficiency features in the upper parts (higher proficiency) of the CEFR scale, where there are high expectations for foreign language learners that would not always be met even by learners for whom English is a first language. Conversely, the lower end of the CEFR scales might not always capture the academic language that learners may be developing in their early stages. In this presentation we propose an academic language proficiency scale that would draw together aspects of academic language ability found in other scales and, if needed, could add new skills not currently covered by the CEFR. It might even be possible to provide a memorable quality name for each level, with descriptors of academic language proficiency either for individual subjects or in general.

**Stuart Duncan Shaw**, Cambridge International Examinations, United Kingdom

Stuart Duncan Shaw is interested in demonstrating how Cambridge Assessment meets the demands of validity in its assessments. Assessment books include: *Examining Writing: Research and practice in assessing second language writing* (Shaw and Weir 2007); *IELTS Writing Assessment Revision Project: towards a revised rating scale* (Shaw and Falvey 2008); *Validity in Educational and Psychological Assessment* (Newton and Shaw 2014); *Language Rich: Insights from Multilingual Schools* (Shaw, Imam & Hughes 2015). He is a Fellow of the Association for Educational Assessment in Europe.

## Using Cambridge Write & Improve for ESL learners

There are quite a few challenges in the development of an automated writing placement model for non-native English learners, among them the fact that exams that encompass the full range of language proficiency exhibited at different stages of learning are hard to design. However, acquisition of appropriate training data that are absolutely relevant to the task at hand is essential in the development of the model. The Cambridge English exams were used to sample a representative and balanced training set consisting of essays and their Common European Framework Reference for Languages (CEFR) scores. Using the CEFR-based Cambridge Learner Corpus dataset, the task was conceptualized as a supervised machine learning problem, and primarily focused on developing a generic writing model. Such an approach facilitates the modelling of truly consistent (internal) marking criteria regardless of the prompt/task delivered, while has the additional advantages of requiring smaller dataset sizes and not necessarily requiring re-training or tuning for new prompts/assessment tasks. The system is developed to predict someone's proficiency level on the CEFR scale, which allows learners to point to a specific standard of achievement. Agreement with human raters from the system is  $r = .77$ . Among a series of trials involving 3000 submissions, both numerical and qualitative feedback revealed a high user satisfaction, while in over 80% of the cases users were able to improve their overall writing score.

**Ardeshir Geranpayeh**, Cambridge English Language Assessment, United Kingdom

Ardeshir Geranpayeh is the Head of Automated Assessment & Learning at the Cambridge English Language Assessment, a department of the University of Cambridge. Ardeshir holds a PhD from the University of Edinburgh on the comparability of language proficiency testing. He has 27 years' experience of test validation and has contributed to the design, development, revision and evaluation of several internationally recognised language proficiency tests. Ardeshir has published extensively on language proficiency testing and cheating detection.

**Helen Yannakoudakis**, University of Cambridge, United Kingdom

Helen Yannakoudakis is a senior research associate at the Computer Laboratory of the University of Cambridge, working on Automated Language Teaching and Assessment (ALTA). She is also a teaching fellow at Girton college (currently teaching Unix, OOP, IR, AI), and a committee member and seminar coordinator of Women@CL. She holds a PhD in Natural Language and Information Processing from the University of Cambridge, during which she also worked on the English Profile Programme (EPP) in collaboration with Cambridge English Language Assessment.

## Using open source and open standards to create best-of-breed language learning solutions

The future of education is all about personalised learning: offering effective learning pathways using a wide range of content and learning apps curated from various sources to make it easier for teachers to teach and students to learn faster and more effectively.

The 2015 EDUCAUSE Learning Initiative (ELI) research paper introduced the concept of the Next Generation Digital Learning Environment (NGDLE). It proposes an OPEN learning ecosystem, highly inter-operable and multi-vendor by nature. It requires products and content

from multiple vendors to work well together, providing a holistic view of each student and their learning progress.

As interoperability is the linchpin of NGDLE, we need solid (open) technology standards to support this and full adherence to these standards by all parties involved. Openness has been lacking in the past, leading to closed vendor-specific solutions and siloed sets of data, creating roadblocks in students' learning pathways. Strong collaborative leadership from both vendors and end-user organizations is required to establish the OPEN learning ecosystem.

We will focus on IMS QTI (Question & Test Interoperability) in particular and how it can be used to create best-of-breed Language Learning Solutions. As an example, we will demonstrate how OAT has used the PCI (Portable Custom Interaction) specification to create a custom Audio Recording interaction type, allowing a test-taker to record speech as a response, without losing inter-operability.

The session will provide participants the necessary background information, strategies and best practices to truly own their learning ecosystem and set direction for the future, ultimately improving learning outcomes.

**Mark Molenaar**, Open Assessment Technologies S.A., Luxembourg

Mark Molenaar is currently holding the position of Chief Technology Officer at Open Assessment Technologies S.A., home of TAO: The Leading Open Source Assessment Platform for Education & Employment. He is responsible for driving the TAO product strategy, technology vision and roadmap; determining long-term product requirements and strategy by interacting with a broad range of stakeholders including EdTech partners and the Open Source community.

## Using technologies to enhance a curriculum for Content and Language Integrated Learning (CLIL)

This presentation will begin with a brief outline of the latest research in the field of technology-enhanced language learning and CLIL, providing a general overview of the international debate on this issue. After outlining the general framework and providing scientific reasons for using technologies in a CLIL environment, some examples of webtools, resources and materials will be shown, offering teachers engaged in a CLIL context useful hints and suggestions on how to take advantage of ICT in the different steps of the teaching/learning process (lesson planning, CLIL classroom activities, assessment). We will look at the implications of these technologies in developing a technology enhanced curriculum for CLIL. Technology may also play a key role in helping teachers improve their language and methodological competences in CLIL. We will consider how self-reflection and meta-cognition of the learning outcomes reached through training pathways are key issues for teachers' professional development and may lead to a more conscious understanding of their training needs.

**Letizia Cinganotto**, Italian Institute for Documentation, Innovation, Educational Research (INDIRE), Italy

Letizia Cinganotto PhD is a researcher at the Italian Institute for Documentation, Innovation, Educational Research (INDIRE). A former teacher of English, teacher trainer and author of digital content, she worked for several years at the Italian Ministry of Education, University and Research, dealing with issues relating to the upper secondary school reform, with particular focus on foreign languages and on CLIL. She has master's and postgraduate degrees in EFL, Italian L2, e-learning, digital and multimedia teaching/learning. Her main research areas are: CLIL, EFL, CALL, MALL, TELL.

**Juliet Wilson**, Cambridge English Language Assessment, United Kingdom

Juliet Wilson is the Director of Assessment at Cambridge English Language Assessment. Before working at Cambridge, she taught English and trained teachers in UK, Colombia, Hong Kong and Portugal. She has an MA in TESOL from the Institute of Education, University of London and a Certificate of Management from the Judge Institute, University of Cambridge.

## Validating university entrance test assumptions: Some inconvenient facts

In many university entrance policies, language tests are used to discriminate between L2 students who are likely to encounter linguistic problems, and those who are not. When L1 students are not required to take an entrance test, the implication is that they are expected to meet the demands that are tested in L2 students. This presentation uses a Toulmin argument structure (Kane 2013) to check two assumptions that support this use of university entrance tests in Flanders, Belgium. The first assumption is that L1 students pass the L2 language demands. The results of 160 first-year university students on an L2 entrance test were compared to those of the L2 population (N = 600). The results (logistic regression and Multi-faceted Rasch) show that 11% of the Flemish candidates did not reach the B2 level. If not all L1 students pass the L2 entrance test, the university entrance policy is not necessarily unjust. Indeed, admitting people under misguided assumptions is arguably less problematic than excluding people on false grounds. For that reason, a second assumption was investigated, the one that students who fail the entrance test will struggle linguistically. Twenty-three L2 students were tracked for one academic year, eight of which, importantly, had actually failed an entrance test. Proportionally, however, it was revealed that the group of students who had failed the language test was as successful – linguistically and academically – as the group who had passed the test. Both studies complement each other in showing that some assumptions supporting the current Flemish university entrance policy may be misguided. The implications go beyond the Flemish context, however, and as such the discussion of the paper will address two questions. Is it just to discriminate within an L2 population based on expectations that not all L1 students meet? And should passing an L2 entrance test guarantee linguistic coping at university?

**Bart Deygers**, KU Leuven, Belgium

Bart Deygers is finalising a PhD study at the Centre for Language and Education (University of Leuven, Belgium). His field of specialization is university entrance language testing, and his main research interests include justice, validity, and the Common European Framework of Reference for Languages (CEFR).

## Validation of a language test linked to a learning programme

In 2016, educational issues pose the question of how the interface between the stages of learning and those of validating skills or knowledge acquired works. From 1980 to 2010, significant progress was made with the quality of exams which validate levels of competence. The emergence of the CEFR as well as the ALTE 'Q-mark' are concrete examples of these advances. If the validation of competences (not specifically tied to a specific course of study) remains crucial, the accompaniment of the learning process by means of ad hoc evaluation tools constitute an important way of optimising the acquisition of skills and knowledge. As digital methods become common it would be wise to question the process leading to the validation of these new tools. To benchmark the specificity of digital evaluation tools used in the learning process we must question the relevance of procedures of validation which have traditionally informed testing and see how they might need to be adapted. The CIEP has recently developed an online, multilingual, multistage placement test. A series of procedures has been employed to assure an optimal take-up of the tool in the world of education and to attest the tool's validity. Among the steps that have been taken, there was first a systematic review of the literature and of professional practice, a needs analysis conducted with indirect users, a series of pilots but also a concurrent validity procedure, as well as a survey carried out with candidates about their use of the tool and lastly the comments of users. After having presented the principal results of the validation of the test, a proposal of the case and the procedure will be made to lay out a framework for allowing the validity of a test of language competence to be judged.

**Vincent Folny**, CIEP, France

Vincent Folny has worked at the CIEP for 10 years. He studied educational Evaluation and measurement in Canada (Université de Montréal). At the CIEP, he is one of the two supervisors of the Cellule qualité et expertise (CQE). He is in charge of the monitoring and the supervision of the tests' quality and validation of the CIEP. He also takes part in the establishment of language tests in Europe and worldwide. He has also participated in the ALTE association meetings and committees for almost 10 years.

**Sébastien Portelli**, CIEP, France

Sébastien Portelli has taught English as a Foreign Language for 10 years at secondary school level. He then specialised in teachers' training and worked with the educational authorities to the introduction of the CEFR in French secondary schools. He joined the CIEP's Assessment Department in 2010, and more specifically the Testing Unit which he has been running since 2012. He is in charge of the global distribution of the TCF (Test de connaissance du français) and of Ev@lang, the CIEP's latest language assessment tool.

## Washback research in the expanding circle: Insights from social psychology and linguistic imperialism

Test washback inquiry is the appropriate standpoint from which the linkages between language assessment and language learning are examined. Informed by social psychology of research and the globalised nature of English language and language tests, the current study critically reviewed washback studies carried out thus far in Iran to summarise the major findings, identify some of the main drawbacks, and put forward directions for future research. Findings

suggest that many of the studies suffer from one or some of the following problems: preconceptions concerning the validity of both local and international language tests, deterministic attitudes towards the value dimension of test washback, lack of a comparative element in research designs posing internal validity threats, the paucity of triangulation in data collection and analysis, failure to consolidate findings in light of larger psychological, educational or testing theories; ignoring the status of English as a lingua franca, and outdated views concerning the proper way of teaching English. The noted issues are examined within a combination of social psychology and linguistic imperialism. The paper concludes with directions for further inquiry into the washback of local and international tests, especially in the outer circle where research and language ideologies do not parallel those in the inner circle spheres.

**Kioumars Razavipour**, Shahid Chamran University of Ahvaz, Iran

Kioumars Razavipour is an Assistant Professor of Language Testing and Assessment in the department of English Language and Literature, Shahid Chamran University of Ahvaz. His primary areas of interest are language assessment and language policy.

## What characterises the quality of CLIL instruction?

This study builds on and contributes to the field of multilingualism. It uses qualitative and quantitative data to link the quality of instruction in Content and Language Integrated Learning (CLIL) with student experiences and achievement. CLIL is an innovative teaching approach in which a content subject is taught in a second or foreign language (Brevik and Moe 2012, Coyle, Hood and Marsh 2010). CLIL was created to address the needs of a multilingual Europe, and is present in over 30 countries (Pérez-Cañado 2012). Still, little is known about how CLIL is practiced in classrooms. As such, this study provides insight into CLIL instruction in a Norwegian context. The present study examines how the combination of content and language is addressed in mathematics and science instruction, and how this dual focus affects the quality of the CLIL teaching. In addition, the study compares the use of English as the target language in the two content subjects with the use of English in English language arts instruction in the same class. In order to analyse various perspectives on CLIL teaching, the data consist of video-observations of science, math and English instruction as well as student questionnaires in a 9th grade CLIL class. We filmed 4 consecutive lessons in each subject (N=12). This study is part of a large video project at the University of Oslo. To measure the quality of the observed CLIL instruction, we used the Protocol for Language Arts Teaching Observation (PLATO), a thoroughly validated coding manual (Grossman 2015). The results indicate that the language aspects of the CLIL instruction are of higher quality than the representation of content, and also that the language support in mathematics and science is higher than in the English language arts lessons. These findings are supported by the student perspectives expressed in the questionnaires.

**Karina Rose Mahan**, University College of Southeast Norway, Norway

Karina Rose Mahan is a PhD candidate in English Education at the Department of Languages at the University College of Southeast Norway. Her research interests include Content and Language Integrated Learning (CLIL), English and Norwegian as second languages, and vocabulary.

## What do teachers really think about using international speaking exams as a goal for students? Views from a bilingual programme

Increasingly, international language exams are inserted into national and regional language learning programmes as a means of evidencing learning, measuring the effectiveness of teaching and quantifying value for money. In general, this measurement is made by scrutinising exam results, rather than asking teachers for their views. In this research, we did not look at results, but instead went directly to the teachers for their own views of the role of the international language exam in their classrooms. The group we asked was made up of hundreds of teachers from the Comunidad de Madrid's innovative bilingual Spanish/English programme, a programme that has attracted much international attention thanks to thousands of students achieving a B1 certificate in English by the age 10-11. The research asked teachers for their views on the impact of the international language exam (either Trinity or Cambridge exams) on the classroom, on learning and on teaching. Did the teachers find that the exam motivated the students? How were the exam materials used as a training and preparation tool in the classroom? How did having an international exam as a target affect teachers' own behaviour in the classroom? Did they think the exam results were an accurate reflection of their students' communication skills? This talk will reveal some of the responses to these questions and some of the surprises. But what this talk also demonstrates is that good teacher survey research is about more than asking questions and counting percentages. Reliable research benefits from the use of up-to-date social survey methodology and practices from the field of social psychology, which can give more vivid, accurate and reliable data. This talk will be of interest to teachers, language testers, and those making important pedagogical and funding decisions around language programmes and the use of international language exams in general.

**Mark Griffiths**, Trinity College London, United Kingdom

Mark Griffiths is a teacher, trainer, author and academic researcher with over 25 years' experience in English language teaching and language testing. With a background in teaching and a PhD in Linguistics, Mark consults for publishers, writing developing English language coursebooks, and he has worked with Trinity College London as a language tester for more than 15 years. Mark designs, publishes and advises on academic research and gives training on teaching and language testing in more than 25 countries.

## What will high-stakes language testing bring to Spain through the new LOMCE exams?

After the design and reform of the Spanish Educational system, foreign languages were expected to have more relevance in the Spanish educational panorama. However, although hopes were high the importance of languages has not been reflected in high-stakes testing. This presentation addresses the policies that have led to two new high-stakes exams in Spain, addressing the rationales, tasks, rubrics and intended inferences for the Final General Education and Baccalaureate Examinations. It also analyses their advantages and drawbacks. According to these observations and analysis, we suggest the outcomes and intended washback effect according to previous experiences obtained from the former University Entrance Examination and indicate the recommendations on the reform that has been provided to the local Ministry of Education of Madrid which could easily be implemented at the national level. In examining the design of the test, we also hope to be able to foresee how foreign languages teaching will change and whether any positive effects are expected.

**Jesús García Laborda**, Universidad de Alcalá, Madrid, Spain

Jesús García Laborda holds a PhD in applied linguistics and a second PhD in Language Education, both from the Universidad Complutense de Madrid in Spain. He has given training courses in the Balkans and the USA. His current interests are computer based low-stakes assessment and English for Specific Purposes (ESP). His latest publications include in the journals *Language learning and Technology*, *Computers & Education*, and the *British Journal of Educational Technology*. He has also participated as a leader in 5 research projects in computer assisted language tests.

# WIP presentations

---

## Assessment in a globalised economy: A task-based approach to assess the proficiency of Dutch in specific occupational domains

For language tests to be in tune with the target context, a constant attention to the shifting characteristics of real-world language use is required, especially in a globalised economy where job migration is necessary to answer the national demands of labour markets. Consequently, as part of its cyclical test validation process, the Certificate of Dutch as a Foreign Language (CNaVT) is updating its business Dutch exam. The first step in this process was carrying out a needs analysis among a purposive sample in order to identify the real-life tasks and language demands that are relevant for test takers who require language certification for professional purposes. The CNaVT conducted a survey amongst teachers, recruitment agents, working professionals and test takers. The results showed that the respondents generally favoured an exam that covers language abilities for two subdomains: services and healthcare. Employers want their workers in these domains to 'make knowledge work', they want them to demonstrate functional language abilities, preferably at or above a B2 level. The new task-based occupational exam Zakelijk Professioneel (PROF) needs to meet the requirements of the test users and the needs of the test takers. For developers of a centralised, criterion-based language test this implies bringing together multiple views in a test that allows learners to demonstrate that they can use Dutch successfully in different professional contexts. During this WIP presentation we will present test tasks that have resulted from the needs analysis, and discuss which measures are taken during task development to ensure content validity, maximise rater reliability and avoid bias.

**Sarah Smirnow**, CNaVT, Centre of Language and Education, KU Leuven, Belgium

Sarah Smirnow is currently a research associate working at the Certificate Dutch as a Foreign Language (CNaVT) Project of the Centre for Language and Education at KU Leuven. She is in charge of developing and supervising the quality of the societal and occupational exam for adult learners. Sarah also was involved in the development of a test for young learners of Dutch as a foreign language. Her research interests include task-based assessment of language competences and test validity.

**Lucia Luyten**, CNaVT, Centre of Language and Education, KU Leuven, Belgium

Lucia Luyten has been working as a research associate at the Centre for Language and Education, KU Leuven since 2002. She is responsible for the societal exam suite at the Certificate of Dutch as a Foreign Language (CNaVT). The projects she has been focusing on are mainly situated in language testing and adult education. She specializes in literacy and language testing. Her research interest include authentic (task-based) language testing, reliability and scoring validity.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

## Beliefs driving the assessment of speaking: An empirical study in a Brazilian public classroom

The difficulties and challenges related to the assessment of speaking have been known for decades (Brown and Yule 1983, Luoma 2004). In the Brazilian educational system, the development of the students' speaking ability has traditionally been left aside for a number of reasons, except in some very specific school contexts. Considering the importance of speaking, the crucial role assessment plays in the teaching-learning process of all language abilities, and the relevance of language assessment literacy in teachers' professional development (Fulcher 2012, Taylor 2013, Scaramucci 2016), this on-going empirical study aims at investigating how the beliefs (Brown 2004, Barnes, Fives and Dacey 2015; among others) about oral fluency (Bailey 2005, Hughes 2011, Nation and Newton 2009) held by English language teachers of a particular public school drive how their students' speaking production is assessed. To achieve such an aim, three specific objectives were established: (1) to identify the English teachers' beliefs about oral fluency; (2) to verify, in the teachers' classroom practice, how they rely on their beliefs to teach; and (3) analyse the impact of such beliefs in way the teachers assess of their students' speaking ability. This is a two-year qualitative study started in March 2015. Three data collecting instruments are being used: questionnaires, interviews and classroom observation. In this presentation, partial data analysis will be presented and discussed.

**Eber Clayton Dutra**, University of Brasília, Brazil

Eber Clayton Dutra holds a BA in Letters (Portuguese/English Languages) from the Federal University of Goiás, Brazil. He is currently doing his MA at the Applied Linguistics Postgraduate Programme at the University of Brasília, doing research on language assessment and English teachers' beliefs about oral fluency. He has been an EFL teacher for over nineteen years, and has worked at a public language centre in the city of Brasília since 2000.

**Gladys Quevedo-Camargo**, University of Brasília, Brazil

Gladys Quevedo-Camargo has a PhD in Language Studies and is an adjunct professor at the Department of Foreign Languages and Translation, University of Brasília, Brazil. Her experiences include teaching English in language schools, being an oral examiner for the Cambridge main suite and the IELTS, and coordinating a university testing development centre. She has several articles and chapters published in Brazil and a recent publication co-authored with Matilde Scaramucci in Gitsaky & Coombe (Cambridge Scholars Publishing, 2016).

## Digital challenges in the assessment of advanced European languages students in their final undergraduate year

As E-learning Co-ordinators for the Department of European Literatures, Languages and Cultures at the University of Edinburgh, we have been facing exciting new opportunities but also digital and pedagogical challenges in the context of a vast project for the development of innovative digital approaches to language teaching and learning. We would like to propose a Work in Progress presentation on the specific challenges we are facing in the context of Higher Education and more particularly in the development of innovative teaching and assessment solutions for final year undergraduates (here in Edinburgh, final year undergraduates are expected to reach a C2 level of proficiency). The presentation will concentrate on the written

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

language component of our courses, and the development of advanced writing skills for our students. In our work both as E-learning co-ordinators and as Language Tutors (in Spanish and French respectively), we have been striving to combine precise and informed linguistic approaches with cutting-edge digital solutions for teaching and assessing students at an advanced level. We aim to present a number of digital solutions that are provided for us by our Virtual Learning Environment (VLE ) blackboard and a dedicated team of Learning technologists and support technicians at the University, and to analyse the way we are adapting them to our particular needs on French 4 and Spanish 4 Honours courses, while encountering particular challenges in the devising of new digital forms of assessment. In particular, we wish to present our findings following months of research into the suitability and effectiveness of the use of collaborative tools for online teaching and assessment (discussion boards, Wikis, etc.), with a view to starting a discussion with colleagues on the broader aspects of collaborative teaching and learning in the digital era, and the creation of trustworthy and suitable assessment tools at advanced level.

**Emmanuelle Lacore-Martin**, The University of Edinburgh, United Kingdom

Emmanuelle Lacore-Martin is Lecturer in French and E-learning Co-ordinator at the University of Edinburgh. A graduate of the Sorbonne and the Ecole Normale Supérieure, with a doctorate in French Renaissance literature, she specialises in the teaching of French language and French Renaissance literature. Her monograph on Rabelais (*Figures de l'histoire et du temps dans l'oeuvre de Rabelais*) was published in 2011. Her latest article (Rabelais's Anatomy of Emotion and the Soul) will appear shortly in the quarterly journal *Renaissance and Reformation*.

**Carlos Soler Montes**, The University of Edinburgh, United Kingdom

Carlos Soler Montes joined The University of Edinburgh as E-Learning Co-ordinator in January 2015. As a trained specialist in teaching Spanish as a Foreign Language, he has directed numerous academic projects in Spain, Canada, Mexico and the United States. Carlos is very committed to teaching. Prior to joining The University of Edinburgh, he worked for ten years at Instituto Cervantes as academic coordinator, curriculum specialist and quality evaluator. Dr Soler Montes is a Fellow of the Higher Education Academy.

## Do foreign language learners and native speakers mark coherence and cohesion in a different way?

Coherence and cohesion are often a criterion used to assess productive skills. The CEFR suggests that higher level students will use less frequent connectors and will use cohesive devices more accurately. The Interuniversity Test of Dutch as a Foreign Language (ITNA) development team is currently examining how coherence and cohesion is generally established in a C1 writing task, in this case a 300-word argumentative essay. To gain better insight into what we can reasonably expect from ITNA test takers, we will compare their use of linking words with the performance of 18-year-old native speakers who performed the same task. For this purpose, two corpora have been developed in #Lancsbox. The first one is a learner corpus comprising C1 ITNA performances of participants who have passed a B2+ level test and thus all already have a relatively high level of proficiency. The second corpus will be filled with ITNA task executed by native speaker first year Bachelor students of the Language faculty (Antwerp University). The

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



investigation focuses on a number of linking words and their occurrence and usage in both corpora. This way, we can compare the occurrence and use of both highly frequent linking words (such as 'and', 'but' and 'or') with less frequent ones (such as 'first', 'however', 'to conclude'). The results will then be used to validate the rating scale and possibly adjust the criteria to assess coherence and cohesion in the writing test.

**Sabine Steemans**, IUTC - University of Antwerp, Belgium

Sabine Steemans is co-ordinator of the Dutch as a Second/Foreign courses taught at the University of Antwerp language institute. She has been involved in language testing since 1999. After having attended seminars in language testing by Dr. Alan Davies, she developed several Dutch entrance examinations. She is part of the IUTC team responsible for the 'Interuniversity Test of Dutch for Speakers of Other Languages' (ITNA).

**Catherine Verguts**, IUTC - Ghent University, Belgium

Catherine Verguts co-ordinates the Germanic Languages department at the Ghent University Language Centre. She also has experience giving coaching sessions on test development and has been part of expert groups giving feedback on high-stakes tests. Moreover, she is part of the test development teams of Interuniversity Test of Dutch for Speakers of Other Languages (ITNA) and test of academic English for lecturers and students (ITACE).

**Ann Vlasselaers**, KU Leuven, Belgium,

Ann Vlasselaers is a member of the test development team of Interuniversity Test of Dutch for Speakers of Other Languages (ITNA). She teaches in the Bachelor's and Master's programme in Applied Linguistics at the KU Leuven (Faculty of Arts) in Brussels.

## Évaluer la compétence à communiquer en français dans l'entreprise

La Chambre de commerce et d'industrie de région Paris Ile de France (CCI Paris Ile-de-France) conçoit et diffuse depuis les années 60 les Diplômes de français professionnel. Ces diplômes s'adressent aux étudiants ou professionnels qui sont ou seront appelés à communiquer en français dans un contexte professionnel et qui souhaitent valider leurs acquis par un diplôme de langue. La CCI Paris Ile-de-France a entrepris en 2015 et 2016 une importante refonte des Diplômes de français professionnel Affaires (5 diplômes, du A1 au C1). La méthodologie mise en œuvre a renforcé l'approche actionnelle de ses examens : dans chaque activité, le candidat réalise une tâche de communication professionnelle définie par un contexte (situation professionnelle) et un ou plusieurs objectifs de communication. Pour réaliser cette tâche, il mobilise, de manière stratégique, des compétences intégrées de réception, de médiation, de production et/ou d'interaction. Il ne s'agit plus de « lire/écouter pour comprendre », mais de « comprendre (ce qu'on lit/écoute) pour agir ou interagir ». L'évaluation porte sur la maîtrise des moyens langagiers et sur le degré de réalisation de la tâche professionnelle. L'adéquation aux réalités professionnelles transparaît dans le choix des documents supports, l'authenticité de leur forme et de leur contenu, mais également dans le caractère réaliste des mises en situation et des tâches de communication. Celles-ci renvoient à des situations de communication transversales, communes aux domaines d'activité les plus courants du monde de l'entreprise et des affaires : ressources humaines, management, marketing, finances, logistique, etc. La présentation vise à montrer, par des exemples d'activités extraites des Diplômes et des retours

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



d'expérimentation, comment ces principes s'actualisent dans les modalités de conception et de mise en œuvre opérationnelle de ces nouveaux Diplômes de français professionnel Affaires.

**Dominique Casanova**, CCI Paris Ile-de-France – Centre de langue française, France

Dominique Casanova is Head of the Scientific and Digital Unit of the Centre de langue française, CCI Paris Ile-de-France. He has been working on psychometric and IT issues regarding assessments since 2005 and conducted the complete renewal of the IT system in recent years. He is an engineer in computer science and applied mathematics and he has managed a French Language Centre in Morocco for 6 years. He has published various papers on psychometric issues regarding the Test d'évaluation de français.

**François Renaud**, CCI Paris Ile-de-France – Centre de langue française, France

François Renaud is a Pedagogical Project Manager in the Centre de langue française of the CCI Paris Ile-de-France. He is in charge of the Diplôme de français professionnel design coordination and of exam centres pedagogical advice and support. He also co-ordinates the creation of various teaching resources in professional French in partnership with TV5Monde and RFI. He regularly gives seminars on French for Specific Purposes Methodology.

**Alexandre Holle**, CCI Paris Ile-de-France – Centre de langue française, France

Alexandre Holle is Head of Educational innovation and Partnerships at the Centre de langue française of the CCI Paris Ile-de-France. He taught French as a foreign language and French for Specific Purposes for 10 years in France and abroad, in French institutes, international institutions and business schools. He also served as an attaché for cooperation in French Language and as Director of a French institute. He is co-author of books and articles on high-stakes testing.

## From global student populations to localised HE settings: An example of an IMDP screening process for academic English and readiness

At the University of Jyväskylä (JYU) there are approximately 20 International Master's Degree Programmes (IMDPs), and the number of offerings continues to increase. However, the graduation rate for each two year programme has not been realised, for example, possibly due to inadequate academic English skills of the students, in spite of positive results from standardised language tests (Final Report Internal Evaluation of the International Master's Degree Programmes at the University of Jyväskylä, 2014:15-18). Furthermore, according to Murray and Nallay (2014), amongst others, academic English skills link to students' academic performance. Addressing the language hurdle in particular would require universities to set higher entry requirements for academic English (Breta, 2007:19). Therefore, JYU piloted a new screening process for student applicants. The pilot rounds of 2016 and 2017 were both conducted with multimodal technologies and included two of the IMDPs. These two programmes shortlisted candidates for screening and candidates were first assessed in academic reading and writing, followed by listening and speaking. Both parts of the test were assessed by two Language Centre English lecturers, who then communicated their feedback concerning each candidate to the programmes with a suggestion for which students might succeed, struggle or might not be able to study in the programmes. An additional aim of the test was to assess readiness (e.g. Murray and Nallaya 2014) in order to check candidates' academic abilities for study at the Master's level specifically at JYU. This presentation will cover the testing process in more detail and its results. A more tailored approach for screening academic language and readiness of

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



IMDP candidates is then presented as an example of what could be done in order to better ensure student success and graduation rates in the IMDPs.

**Miia Konttinen**, University of Jyväskylä Language Centre, Finland

Miia Konttinen is a university teacher of English at the Language Centre of the University of Jyväskylä. She teaches students and staff the use of English in field specific contexts, and she specialises in screening and teaching global student populations in the International Master's Degree Programmes (IMDP). Additionally, she is a PhD student researching academic English learning and teaching of the IMDPs in the Centre of Applied Language Studies at the University of Jyväskylä.

**Lisa Lahtela**, University of Jyväskylä Language Centre, Finland

Lisa Lahtela is a university lecturer of English at the Language Centre of University of Jyväskylä. She works with students and staff in the use of English for field specific purposes, and her responsibilities also include language testing for outgoing exchange students as well as incoming International Master's Degree Programme (IMDP) students. In addition she is also involved in the Finnish testing service, The National Certificates of Language Proficiency.

## Generating the assessment literacy profile of EAP teachers in the Mainland China: An evidence-based approach

We are currently working on an evidence-based approach to understanding the Language Assessment Literacy (LAL) needs of specific language assessment stakeholders in order to build group-specific profiles that generate targeted LAL development programmes. This paper reports on the conceptual-empirical approach to building a LAL profile for Chinese English for Academic Purposes (EAP) teachers and the research tools used in the process. The presentation draws on a combination of theoretical study and an 'in-practice' example. From a theoretical perspective, we consider the link between LAL and test validity. This serves to establish why LAL is necessary for all those involved in language assessment, wherever they may be situated in the test cycle, and why it is significant for a particular test and testing in general. We then describe an evidence-based approach that employs participatory, collaborative instruments to generate LAL profile of EAP teachers while raising their LAL awareness in the process. Several groups of Chinese EAP teachers (over 100 participants so far and the project is ongoing) were invited to attend face-to-face LAL workshops. Input on the concept of 'validity', with concept tailored for their specific context, was delivered. The teachers then collaboratively worked through a series of tasks designed to relate their role directly to test validity and to identify what areas of knowledge are necessary for them to safeguard the value of the test and mitigate potential threats to overall validity. Questionnaires were administered before and after the workshops. The grounded-theory analysis was conducted on data obtained during the workshops and from the questionnaires. We demonstrate the EAP teacher LAL profile generated through the process described above, review the approach and make recommendations for future iterations. The presentation concludes with a proposal for the potential use of the approach with different stakeholder groups to contribute to a cross-stakeholder LAL inventory.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



**Olena Rossi**, Lancaster University, United Kingdom

Olena Rossi is a PhD student at Lancaster University, UK. Before moving to full-time studies she worked for the British Council China as an assessment consultant and academic trainer. She is an experienced teacher trainer and has an MA in Language Testing from Lancaster University. Her areas of interest include assessment of productive skills, item writing and assessment literacy for language teachers.

## How value stream mapping at a Kaizen workshop at Cambridge English lead to a digital transformation in our test analysis

ALTE's minimum standards for test analysis demand that assessment organisations collect and analyse data on an adequate and representative sample of candidates in order to ensure the reliability of test items. Cambridge English fulfils this requirement by means of pre-testing, whereby candidates from a variety of backgrounds sit pretests, enabling the performance of test items to be analysed against a range of criteria to evaluate if they are fit for purpose. In order to increase the efficiency of test analysis through the pretesting process, Cambridge English undertook a value stream mapping exercise via a lean Kaizen workshop. The word Kaizen means 'change for the better' in Japanese, and is an approach which aims to involve all individuals in an organisation in seeking to continuously improve standardised activities and processes and eliminate waste. Whilst traditionally pretesting at Cambridge English has been paper-based, the Kaizen workshop identified opportunities for deriving significant efficiency gains by moving to an alternative delivery model, i.e. computer-based pre-testing, with the potential to bring significant benefits to candidates, test centres and also Cambridge English. During the course of the presentation, the case study will explain techniques using Kaizen methodology which can have a wide range of applications, e.g. helping achieve the necessary quality profiles in ALTE examinations and also implementation of effective digital solutions for language learning and assessment.

**Hugh Moss**, Cambridge English Language Assessment, United Kingdom

Hugh Moss has worked at Cambridge English since 2003, now in Operations, he previously spent 12 years within Assessment. He holds an MEd in English Language Teaching from Sheffield University, with his dissertation focusing on oral error correction. His previous careers include English Language Teaching and Town Planning. Hugh is committed to process improvements and firmly believes in the impact incremental changes can have on efficiency.

## Language assessment in teacher education programmes in Colombia

This paper reports on a small scale research project in English teacher education in Colombia aimed to inquire about the correlation between the rather limited effectiveness of Colombian bilingual initiative and the parameters of foreign language proficiency applied to language assessment of prospective teachers. The issue of measuring language knowledge of student teachers in reliable and valid ways becomes particularly important in view of the fact that Colombia seriously aspires to become a knowledge-based economy. Since undergraduate English education programmes are receiving students with limited English proficiency, public

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

sector higher education institutions are required to assess foreign language competency level of prospective teachers before they graduate. Within this framework, test takers are expected to demonstrate proficiency in reading comprehension, vocabulary and language use. However, no provision is made to assess their speaking skills even though consensus exists that these are crucial to effective language teaching. It was therefore intriguing for the researcher to find out what is the view of student teachers of this situation. A survey conducted among fifth year students of an undergraduate teacher training programme aimed at establishing students' views on the role of the speaking component of language testing in general, and in testing teachers' language proficiency in particular. As the data presented in this paper concern students involved in English teacher education, the results may have a broader policy implications which could be helpful in re-defining graduation requirements for English teachers in Colombia.

**Bozena Lechowska**, Universidad Industrial de Santander, Colombia

Bozena Lechowska studied English Philology at the Jagiellonian University, where she received her MA in 1986. She is an Associate Professor of English and Applied Linguistics at the Universidad Industrial de Santander, Bucaramanga, Colombia since 1988. She teaches courses in English phonetics and phonology, British cultural studies and translation. Her current research interests focus on language policy and planning and intercultural communication.

## Online language testing of immigrant languages: A nightmare or just a challenging reality check for the CEFR?

Is it possible to use an existing online placement test based on the CEFR for testing immigrant languages? What is the outcome if this test is a combination of self-assessment and self-testing? How can we use the test for testing Arabic dialects without using written language? Studieskolen has already developed an approved, reliable and user friendly online test for German, English, French, Italian, Spanish, Norwegian, Danish and Swedish. We have experienced a growing need for documentation of mother tongue language skills, especially from children and grandchildren of immigrants who need documentation of their knowledge of their parents' and grandparents' mother tongue for work and educational use. As we wanted to develop online assessment tools for immigrant languages, we set out to explore if we could further develop our existing online test-system based on self-assessment for Turkish and Arabic. We have succeeded in achieving this goal, even if it was much harder than expected. Furthermore, we set out to develop a test version for three Arabic dialects. In this workshop, we want to invite you to join us on our journey to develop a test for immigrant languages, especially Arabic dialects. It has been an exciting and challenging process as we had to find solutions and to make compromises we had never tried before. In the end we got our test. The test is brand new and does not yet fulfill all our requirements for quality assurance, but this is the next step of the process. The outcome has also resulted in a much better tool for immigrants that could use self-evaluation in their mother tongue when testing their second language assessment. The test was developed with funding from a Leonardo Transfer of Innovation project by the largest providers of second language teaching for adults in Norway, Sweden and Denmark.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



**Christoph Schepers**, Studieskolen, Denmark

Christoph Schepers is Senior Consultant and Leader of the Development Department for Foreign Languages at Studieskolen, Denmark's largest language school for adult education. He has a master's degree in Language and Literature from University in Münster, Germany and a degree in adult education from Danish School of Education. He has been working with language education, teacher training and assessment for about twenty-five years.

## Performance of multilingual speakers of Dutch on the ITNA admission test for higher education

The Interuniversity Test of Dutch as a Foreign Language (ITNA) is an ALTE-approved, high-stakes admission test developed by Flanders' Interuniversity Testing Consortium (IUTC). The ITNA has been administered since August 2010 and assesses whether candidates have reached the B2 or C1 level of the Common European Framework of Reference for Languages (CEFR). The test comprises three parts: a computer test, an oral examination and an additional writing part for C1-candidates. To determine the ITNA's construct validity, a pilot study was conducted in 2010 in which native speaker performance was compared with ITNA candidate performance, as in TOEFL (Chapelle et al 2007). The linguistic requirements for foreign language speakers should not be more strict than those for native speaker students. Therefore, the ITNA has been developed to assess only linguistic competences and not academic skills. In the pilot study, a group of 44 secondary school students passed the ITNA computer test, which confirms the validity of the set ITNA cut-off score. The aim of the present study is to replicate the 2010 pilot study with a specific subgroup of native speakers, namely 50 multilingual secondary school students. At home, these students speak another language than Dutch, sometimes supplemented by Dutch. However, they attend secondary school in Dutch. Although these students often face more difficulties successfully completing their first year in higher education (Nicaise I 2011), they are allowed to start higher education without passing a language test. The test performance of these multilingual secondary school students is interesting in various ways. If the majority of the test group passes, it can be assumed that the admission policy in Flanders does not impose unreasonable demands on international students. In the worst-case scenario only a minority passes. This might call for reflection about possible implications for admission policy in Flanders as well as for the secondary education system as a whole.

**Ines Blomme**, Interuniversity Testing Consortium, Belgium

Ines Blomme is a teacher of Dutch as a foreign language at the ILT (Leuven Language Institute, KU Leuven), and is specialised in pronunciation teaching. She is also part of the test development team of the Interuniversity Test of Dutch as a Foreign Language (ITNA), which is a collaborative project of four university language centres in Flanders.

**Leen Verrote**, Interuniversity Testing Consortium, Belgium

Leen Verrote teaches Dutch as a foreign language at the ILT (Leuven Language Institute, KU Leuven). She is also part of the test development team of the Interuniversity Test of Dutch as a Foreign Language (ITNA). In addition, she is involved in developing NedBox, an online platform for students who are learning Dutch as a foreign language.

## Puglia integrante – dalla vulnerabilità all'integrazione: percorsi di inclusione per rifugiati e minori sperimentando un toolkit del Consiglio d'Europa

Ci si propone di presentare i primi risultati della sperimentazione che ha realizzato percorsi integrati di alfabetizzazione, formazione linguistica ed educazione civica rivolti a cittadini di Paesi Terzi appartenenti ai due target vulnerabili dei richiedenti asilo/rifugiati e dei minori a rischio esclusione, adottando il Toolkit messo a punto dal Consiglio d'Europa (Language Policy Unit – LIAM group); tale Toolkit consiste in una serie di strumenti ideati per il supporto linguistico e culturale del profilo utenza sopra richiamato, la cui struttura mira a rispondere ai bisogni dei volontari che in particolare operano nel settore della prima accoglienza e che cercano risposte concrete in relazione a tre macro aree: 1) le informazioni da acquisire: lo stato giuridico del richiedente asilo/rifugiato con riferimento alla normativa europea, il contesto geopolitico delle aree di provenienza, le caratteristiche di base delle lingue dell'utenza vulnerabile; 2) le competenze da possedere: tecnico-procedurali nelle pratiche di accoglienza, motivazionali, relazionali, ivi compresa la gestione dei gruppi; 3) la conoscenza delle risorse del territorio: la rete locale dei soggetti che possono supportare il lavoro quotidiano e i servizi da loro erogati. Più nel dettaglio il toolkit contiene tanto linee guida generali, quanto vademecum operativo, quanto ancora risorse "pronte all'uso" quali strumenti per l'analisi dei bisogni e soprattutto attività linguistiche basate sul concetto di scenari comunicativi.

**Gianvito Ricci**, Associazione Quasar, Italy

Gianvito Ricci was born in 1969 at Gioia del Colle in Apulia. In 1993, he graduated with a degree in Philosophy from the University of Bari. In 2011, he co-ordinated Action 1 of the project "Officine Linguistiche" within the Fondo Europeo per l'Integrazione di cittadini di paesi terzi (FEI). He also co-ordinated Action 1 of the project "Le Nuove Officine Linguistiche" in 2012, Action 1 of the project "Italiano in Rete" in 2013, and Action 1 of the project "Io, Voi, Noi - Integrazione Sociale in Rete" in 2014.

## Scoring writing digitally – a tool for raters and test administrators

Digital language tests present new possibilities, as well as new challenges, for test developers aiming to ensure quality in all stages of test administration and assessment. Vox – the Norwegian Agency for Lifelong Learning – has developed and administered a digital test of Norwegian as a second language since 2014, of which the writing proficiency test is administered and marked digitally by human raters. A comprehensive digital tool has been developed in order to allow the raters and the rater-administrator to work efficiently together. Having digital solutions for a writing proficiency test and the marking process, creates new possibilities for the test organisation to efficiently and flexibly administer tests and guide the raters, as well as to maintain control over all stages of the rating process and results. However, a digital test system also presents new challenges, for instance ensuring a well laid-out design and an easy to understand user interface, a secure and reliable administration and storing of results, and tools for preventing and detecting cheating. This presentation is intended as a practical case study, with the aim of showing the digital solutions Vox has designed to make the most of the new possibilities as well

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



as facing the challenges of digital language testing. It includes a presentation of the rater website, where the raters read and mark the essays, and a discussion forum where the raters get technical and professional support from the test developers. The presentation also goes into some detail about the rater-administration website, where the test developers can allocate and administer texts to the raters, supervise the marking progress and check and ensure quality in the results. The solutions for the rater and rater-administrator website are discussed with regards to efficiency and ensuring quality in our digital writing proficiency test.

**Hanne Lauvik**, Vox, Norwegian Agency for Lifelong Learning, Norway

Hanne Lauvik has a master's degree and post-graduate teacher education from the University of Bergen. She works as a test developer for Vox - the Norwegian Agency for Lifelong Learning, focusing on the writing-proficiency tests. She is in charge of item-writing, rater-training, rater-support and the digital rater-tool for the Norwegian writing proficiency test. She has previously participated in a project specifying the Common European Framework of Reference for Languages (CEFR) for Norwegian (Carlsen 2012), and a research project on examination aids in writing proficiency tests.

**Ingvild Kleiveland Vevle**, Vox, Norwegian Agency for Lifelong Learning, Norway

Ingvild Kleiveland Vevle has a master's degree in Scandinavian Languages and Literature and post-graduate teacher education from the University of Bergen. She has been working with language tests for adult immigrants in Norway for several years, including Test of Norwegian – advanced level. Currently, she works as a test developer for Vox - Norwegian Agency for Lifelong Learning, focusing on the oral-proficiency tests. She is in charge of item-writing, rater-training, rater-support and the digital rater-tool for the Norwegian oral proficiency test.

## Tatar exams needs analysis: The case of Kazan

This paper reports on a needs analysis held to determine the demands of Tatar course students' taking the Tatar language test at Kazan Federal University, Russia. In 2013, the Ministry of Education of the Republic of Tatarstan, Russia, teamed up with Kazan Federal University and introduced free Tatar language courses with the complex aim to provide the general public (including non-Tatars) with learning materials of Tatar as a second language, disseminate Tatar traditions and improve learners' understanding of the Tatar culture. As the course gained popularity, it was decided to design, develop and conduct Common European Framework of Reference for Languages (CEFR) A1 and A2 Tatar Language Tests for Speakers of other Languages. However, piloting of the tests held in 2015 at Leo Tolstoy Institute of Philology and Intercultural Communication, Kazan, proved the tests to be too simple as more than 80 % of the test-takers received maximum scores. As part of the needs analysis, the authors of the presentation developed an open-ended Tatar Language Test Questionnaire and in the fall of 2015 and the fall of 2016 administered it to all students of Tatar language courses who took the test at the end of the course (N=457). The respondents reported on pro and cons of the test they were involved in. The content analysis of the responses showed a broad variety of test-takers'/students' needs (from reading street signs to taking a University course), although the demand mentioned in over 74% of the Questionnaires was acquisition of culture and Tatar traditions. The findings also show that Year 2015 group of Tatar course students found that the Reading and Listening texts were unchallenging and lacked a distinctive ethnic identity. Based

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



on the results of the needs analysis, the Tatar Language Test Specifications (CEFR A1, A2) that address cognitive, linguistic and affective issues were designed and recommended for practice.

**Marina Solnyshkina**, Kazan Federal University, Russia

Marina Solnyshkina is a well-known EFL innovator in Russia. She conducts research and lectures on Language Policy, Language Assessment, Academic Writing, and Professional Languages. Dr Solnyshkina has been actively involved in developing and item writing of the first Russian standardised National Exam in English for high-school leavers since 2009. She has twenty eight years' experience in pre- and in-service teacher training. She is the author of two books, a dictionary, an edited collection and a number of textbooks.

**Gulnara Sadykova**, Kazan Federal University, Russia

Gulnara Sadykova is a Deputy Director for International Affairs in Leo Tolstoy Institute of Philology and Intercultural Communication and an Associate Professor at the Department of Germanic Philology, Kazan Federal University (Russia). She holds a PhD degree from University at Albany – SUNY (USA). Her research interests include multilingual education, cross-culture distance education and technology-assisted language learning. Dr Sadykova also heads Kazan International Linguistic Centre that offers language tests including Cambridge examinations, TOEFL, and certification for Tatar language.

**Alsu Ashrapova**, Kazan Federal University, Russia

Alsu Ashrapova is an Associate Professor in Leo Tolstoy Institute of Philology and Intercultural Communication, Kazan Federal University. In 2003, she graduated from Kazan State Pedagogical University with a dual degree in teaching the Tatar language and literature and teaching English. In 2006, she received her PhD in Linguistics and accepted a position at her alma mater. Besides teaching, she leads a team of professionals who develop Tatar language tests for the Ministry of Education of Tatarstan Republic.

## The role of language exams in Colombian higher education

In Colombia, the Ministry of Education (MEN) declared that graduates from universities must certify a B2 in a foreign language, English being the preferred one by most institutions. However, a great percentage of universities require a B1 level since the majority of their students begin their degree programmes with A1 or A2 English levels, except for that 8% of graduates from private 'bilingual' and intensive English schools who reach a B2 or higher English level (MEN Programa Nacional de Inglés, 2014). Therefore, language exams in higher education are used as a graduation requirement in all programmes in most universities. Such a requirement can be met in three different ways: first, by presenting a B1 or higher level standardised language test certification; second, by taking a proficiency language exam at the institution; or third, by passing a certain number of language courses which include either an institutional or a standardised language exam. This test, in some institutions, is the final exam of the last course. For others, it is a proficiency exam taken after studying all the language courses. In some institutions, the exam only assesses reading comprehension. Besides, language exams are used in higher education institutions to classify students according to their language level so that they can either enroll on one of the language in-house courses or take the proficiency exam chosen by the institution to meet the language requirement. In English Language Teaching (ELT) education programmes specifically, language exams are used for admission and graduation requirements. Regarding the former, some universities use the English component of the national examination (Saber Pro) taken at the end of secondary education. It assesses

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

vocabulary, language use and reading comprehension. Others use an institutional exam. As for the latter, universities use either an institutional or a standardised language proficiency exam.

**Bozena Lechowska**, Universidad Industrial de Santander, Colombia

Bozena Lechowska studied English Philology at Jagiellonian University, where she received her MA in 1986. She has been an Associate Professor of English and Applied Linguistics at the Universidad Industrial de Santander, Bucaramanga, Colombia since 1988. She teaches courses in English phonetics and phonology, British cultural studies and translation. Her current research interests focus on language policy and planning and intercultural communication.

**Edga Uribe Salamanca**, Universidad Industrial de Santander, Colombia

Edga Uribe holds a PhD in Didactics of Languages and Cultures from Sorbonne-Paris 3 University. She has done research on language policies and the teaching of French and Spanish as foreign languages. She also has experience as an Academic Consultant and Material Developer. She is currently an Associate Professor at the Universidad Industrial de Santander and the director of the School of Languages offering both ELT and Spanish and Literature education programmes.

**Olga Uribe Enciso**, Universidad Industrial Santander, Colombia

Olga Uribe holds a master's degree in Teaching English as a Foreign Language (TEFL), a master's degree in Teaching Spanish and a Diploma in Teaching English to Speakers of other Languages (DELTA) certification. She has also worked as a teacher trainer, academic consultant, and curriculum designer for the British Council and the Colombian Ministry of Education. She has done research on reading strategies and the role of L1 reading and writing skills in university students. She has published papers on ELT topics. She is currently a professor on the ELT Language Education Program at the Universidad Industrial de Santander.

## The use of blogs, Skype and authentic tasks in the modern language classroom

How can teachers bring the world into their classrooms by using modern tools? Today's teachers have tasks that are wider than teaching pupils how to read and write. The wider task is to raise the awareness of the world around us. The refugee situation in Europe, for instance, is immense and we can see that nationalist parties are gaining ground. In our classrooms, we now have pupils who come from countries ravaged by war. You have to change the way you teach to handle the situation and digital tools are necessary. I try to find and create authentic situations where the students can speak, read, listen and write English to a real audience. When speaking to a guest speaker through Skype, who only speaks English and is also able to add a different perspective, the pupils are motivated and able to understand that the world is in front of them and that they need to embrace it. We also need to teach our pupils 21st century skills. Today it is also necessary to know how to handle a computer or iPad and be able to communicate through it. It is my job to prepare my pupils and make them ready to, for instance, fulfil a job interview through Skype. We combined literature about World War II with digital tools such as Skype in the classroom, guest speakers and a blog. This helped the pupils understand both the past and the present and taught them that they are the future and the only ones who can have an impact on it. By letting them speak to a guest speaker who is a Vietnam war veteran the pupils were able to hear the perspective of a soldier. A perspective that I wouldn't be able to give them without using modern tools.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



**Sara Bruun**, Ringsjöskolan, Sweden

Sara Bruun is Lead Teacher of English and ICT at Ringsjöskolan in Höör, Sweden. She is a very appreciated lecturer in Sweden. In 2014, she wrote her first book about how to use ICT in the classroom. In 2015 she was awarded the European Language Label. In 2016 she was awarded the Royal Swedish Academy of Letters, and her second book: Bring the world into the classroom, will be released.

# Poster presentations

---

## 25 Years of UNlcert® – Quality Assurance in university language teaching and learning

In 1992, German university language centres developed UNlcert® as a quality label for university language teaching & learning as well as testing & assessment, in order to certify language skills that are relevant for future academics. 25 years later, it is time to analyse whether our aims and objectives have been reached, what makes UNlcert® special and which new challenges we have to face. UNlcert® is a decentralised framework system for the certification of language teaching and testing programmes, and applies the Common European Framework of Reference for Languages (CEFR) to the university context. The member institutions have agreed to follow the UNlcert® minimum standards for teaching and assessment and adapt the given framework to their specific needs. This presentation will present the preliminary results of an internal study carried out with the following three aims: First, to identify the impact of the UNlcert® accreditation on quality assurance at the individual member institutions. Second, to assess the relevance and the added value of the accreditation for the institutions. Third, to identify possible improvements to the current system. This study focuses on different aspects of quality assurance management within UNlcert®: Developing standards in teaching and assessment of university-specific language skills across languages and levels, introducing competence-based and task-based approaches to teaching, learning and assessment, enhancing assessment literacy among language teachers, increasing the satisfaction rate of various stakeholders (students, teachers, universities, future employers etc.), analysing the impact of our quality assurance procedures and ensuring sustainability in teaching, learning and assessment. This study follows a qualitative and quantitative approach and collects data from an online study carried out during the Winter Semester of 2016/2017 at accredited institutions and aims to implement changes to the UNlcert® system in order to make it more user-friendly and cater to the needs of the various stakeholders.

**Johann Fishcer**, Zentrale Einrichtung für Sprachen und Schlüsselqualifikationen (ZESS), Georg-August-Universität Göttingen, Germany

Johann Fischer studied English, French and Italian and holds a PhD in Romance Sociolinguistics. He was previously director of the languages centres at the Universities of Hohenheim, Tübingen and Würzburg and has been Director of the Centre for Languages and Transferable Skills at Göttingen University since 2009. He has been a member of the UNlcert® Committee since 1997, and its Head since 2007. His research activities focus on language testing, task-based language teaching (TBLT) and task-based assessment (TBA); recently, he co-ordinated the projects EXPLICS (on TBLT) and GULT (on task-based university language testing).

**Astrid Reich**, Ruhr-Universität Bochum, Germany

Astrid Reich holds a PhD in Psycholinguistics from Freie Universität Berlin. She has been responsible for teaching and assessment at the Language Centre of Ruhr-Universität Bochum and a member of the UNlcert® Steering Committee. She is a project partner of the Institute for Educational Quality Improvement (IQB) in the context of nationwide learning assessments of secondary school students (project VERA 8 for French). She has

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



run teacher training courses on various topics concerning testing and assessment, recently with a focus on listening comprehension, writing and task types.

## Advanced proficiency: How to get there?

In this poster we present results from a large project of proficiency assessment at a Midwestern US university. During spring 2015, we tested more than 1,100 students who were studying one of the four languages: Chinese, French, Spanish, Russian. Of those more than 700 were taking first or second year university language courses and more than 400 were taking third or fourth year classes. Students were assessed using standardised tests made available to us from the American Council on the Teaching of Foreign Languages (ACTFL). In addition to the testing, students completed an extensive questionnaire that sought to determine basic demographics, motivation for language study, use of language in study abroad settings, and use of language technology in language learning. In this poster we consider those students ( $n=174$ ) who reached Advanced level on the ACTFL scale (low, intermediate or high) on one or more of the language skills tested (listening, speaking, reading) and investigate possible correlations that may have contributed to those levels of proficiency. We also classified the students by one out of seven Advanced profiles: Advanced in all skills (one profile), two skills (three additional profiles), or one skill (three additional profiles). The primary question asked is: Are there clusters of individual-difference profiles that are common to advanced level learners? To address this question, we use cluster analysis to determine profiles of advanced learners. Preliminary results suggest that those who are able to reach Advanced level proficiency in at least one skill area have had significant experience with the second language prior to university-level study (heritage status or prior language study), they have had a boost from study abroad, and/or they take advantage of opportunities outside of the classroom. The poster will also make curricular suggestions based on these results that programmes can adopt to promote advanced-level-skill development.

**Susan Gass**, Michigan State University, United States

Susan Gass received her PhD from Indiana University. She has published widely in the field of second language acquisition (more than 30 books and 100 articles) and is the Co-Principal Investigator (with Paula Winke) on a 3-year Proficiency grant awarded by the Flagship program of the US Department of Defense. She has served as President of the American Association for Applied Linguistics and of the International Association of Applied Linguistics (AILA). She has won many local, national, and international awards.

## Developing and validating a reading strategy scale for Chinese tertiary EFL learners

The China State Council issued the document The Implementation Opinions in September 2014 which stipulated the construction of the China Standards of English (CSE). According to the theoretical foundation of CSE, strategic competence is one component of language ability. It is a consensus that strategic competence is an indispensable part of language ability. The aim of this study was to develop a theoretically-based and empirically-developed reading strategy scale and to evaluate whether such a scale functions reliably and validly. Three

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



questions were addressed: (1) What strategies do the Chinese tertiary EFL learners adopt in reading comprehension? (2) Are the reading strategies gradable? (3) How reliable and valid is the scale of EFL reading strategies? The study was undertaken in three phases. During Phase 1, the strategy items were elicited from EFL learners through verbal reports and written diaries, generated from EFL instructors through exemplar generation and collected from literature surveys as well, then all the items were transformed into descriptors and these descriptors were compiled into questionnaires and administered to a sample of 1,200 EFL tertiary learners to address the first question; Nvivo 7 is used for the qualitative analysis. In Phase 2, the results of Phase 1 were used as the basis for the development of a reading strategy scale through 6 experts' judgment and an analysis of the judgments of 2,000 EFL learners using the many-facet rasch model. Phase 3 involved the validation of this empirically-developed scale. The Rasch-measurement software package Facets (Linacre 2010) was used for the analysis of 600 EFL learners' judgment. Interviews were also administered to EFL learners and instructors to elicit the raters' perceptions of the efficacy of the scale. This study is supposed to answer the three research questions; one overall reading strategy scale and one self-assessment reading strategy scale are expected.

**Zhou Yanqiong**, Guangdong University of Foreign Studies, China

Zhou Yanqiong is a PhD candidate in Guangdong University of Foreign Studies in China, specialising in language testing under the supervision of Professor Liu Jianda. She has participated in the huge project of the construction of the China Standards of English (CSE) for more than two years and has published several papers pertaining to this project domestically.

## El español y la certificación lingüística en la Universidad de Salamanca

Cursos Internacionales de la Universidad de Salamanca, socio cofundador de ALTE, está vinculado a la evaluación y certificación del grado de competencia y dominio del idioma español desde finales de los años 80. Desde entonces, ha desarrollado y participado en la evaluación del español con diferentes pruebas de certificación. Su trayectoria incluye la colaboración en proyectos como Surveylang o Study on comparability of language testing in Europe. En la actualidad colabora con el Instituto Cervantes en la elaboración de los modelos de examen y en la evaluación de las pruebas para la obtención de los Diplomas de Español como Lengua Extranjera (DELE) que otorga el Instituto Cervantes en nombre del Ministerio de Educación, Cultura y Deporte del Gobierno de España. También es centro examinador de este sistema de certificación. Del mismo modo, promueve, junto con el Instituto Cervantes, la Universidad Nacional Autónoma de México y la Universidad de Buenos Aires, el Servicio Internacional de Evaluación de la Lengua Española (SIELE), examen progresivo que certifica el grado de dominio del español a través de medios electrónicos. Para la certificación de los fines específicos, Cursos internacionales de la Universidad de Salamanca cuenta con el examen BULATS español, un examen que permite evaluar las aptitudes lingüísticas en el mundo profesional que crea y distribuye a través de su red de centros autorizados por todo el mundo.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



**Juan Miguel Prieto**, Cursos Internacionales, Universidad de Salamanca, Spain

Juan Miguel Prieto holds a PhD in Spanish Philology from the University of Salamanca and is currently in charge of the Assessment Department of Cursos Internacionales de la Universidad de Salamanca. He is highly experienced in running teacher training courses and seminars at Spanish and international universities and institutions. He is responsible for the teams that work on the development and marking of the Diplomas de Español como Lengua Extranjera (DELE) and the Servicio Internacional de Evaluación de la Lengua Española (SIELE).

**Marta García**, Cursos Internacionales, Universidad de Salamanca, Spain

Marta García holds a PhD in Assessment and a master's degree in teaching Spanish as a Foreign Language, both from the University of Salamanca. She is an experienced teacher of Spanish as a foreign language. Aside from teaching, Mrs. García has also contributed her expertise to the production, development, revision and assessment of the DELE tests and is currently the Academic Coordinator of the BULATS Spanish tests. She has participated in ALTE projects and she is currently working in the ALTE LSP group.

## English teachers' perceptions of China's Standards of English for speaking

China is developing its national framework for English language proficiency scales. The framework, called China's Standards of English, describes in a comprehensive way what Chinese English language learners can do. The construction of the Standards follows three steps: (a) establish descriptor pools from existing language scales; (b) classify and assign levels to descriptors through expert judgments; and (c) validate the scales through an analysis of the judgments of a large sample of teachers and students using Many-facet Rasch model. While developing the speaking part of China's Standards of English, we classified the descriptors we collected into three broad categories: performance features, language activities that learners can accomplish, and cognitive strategies. And we initially divided the descriptors into nine levels based on the educational system in China. The classification and division into nine levels were all based on expert judgment. It was vital to hear from English teachers' voices. This study examined how high school English teachers perceive the descriptors of the speaking part of the Standards. Altogether eight teachers attended an interview after finishing a questionnaire concerning the judgements of their students' speaking abilities. The analysis of the interview data revealed that English teachers were confident in judging their students' performance features while less confident in evaluating whether their students can accomplish certain language tasks, even less confident in assessing their students' cognitive strategies when completing an oral task. The study also found that the English teachers thought the descriptors regarding to language activities were very helpful in teaching speaking. The findings suggest that elaborating the descriptors concerning language activities learners can do and reducing descriptors of cognitive strategies may benefit English teaching in China.

**Wang Hua**, Shanxi University, China

Wang Hua obtained his PhD from Guangdong University of Foreign Studies and he is now a lecturer in the School of Foreign Languages in Shanxi University.

**Jie Wei**, Shanghai Jiao Tong University, China

Jie Wei is now a PhD candidate majoring in language testing in Shanghai Jiao Tong University.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



## From language assessment literacy to better teaching and learning

English language testing is a challenge for Russian universities, since the educational standards do not have any references to international frameworks (e.g. the CEFR). For this reason the universities have to act at their discretion when setting requirements for progress, achievement and final tests. Even within one university different language departments can make their own decisions on what and how to test (there are 5 such departments in Ural Federal University (UrFU)). Three different computer-based tests were used for UrFU first year students in summer 2016: level-based tests (mock Cambridge English Language Assessment exams, including only receptive skills and use of English); Federal Internet Exam in Professional Education (FEPE), designed by a testing centre in Russia for students of different majors; System of Monitoring Students' Educational Achievements (SMSEA) designed by a group of UrFU teachers. The results were measured for 3274 students in level-based exams, 5319 in FEPE, 1048 in SMSEA. 65-85 % of students taking level-based tests scored within the range of 60-85% thus confirming their level, as according to Cambridge English Scale the passing grade starts with 60%. 50-65% of students taking FEPE scored high and only some of the students scored high in SMSEA. The non-level-based FEPE demonstrates a normal distribution of students between the levels giving us a bell curve. However, given its considerable drawbacks and lack of consideration of even the basic test design principles, the test cannot be regarded as a valid and reliable assessment tool. SMSEA results give us a fluctuating line with no distinctive pattern. As the decision on the kind of test was left to the language departments, our interpretation of the results suggests it is the knowledge of the basics of test-design that could help teachers, first, with choosing the right test for their students and second, in their teaching practice.

### **Tatiana Rasskazova**, Ural Federal University, Russia

Tatiana Rasskazova is an Assistant Professor in the Department of Linguistics of the Institute of Social and Political Sciences, Ural Federal University, and has a specialist degree in Philology from Ural State Pedagogical University. She is a qualified teacher of English and a translator, and has taken the TKT, CELTA, DELTA (modules 1&2), IDLTM. In 2015 she also participated in a course on Testing, Evaluation and Assessment from NILE. She has been teaching English to university students for over twenty years; has been involved in developing tests for students for over ten years, and has published several articles (both in English and in Russian) on English language teaching and learning.

### **Maria Guzikova**, Ural Federal University, Russia

Maria Guzikova, holds an MA in «Applied Educational Leadership and Management» (UCL Institute of Education, UK), and a PhD in History (Ural Federal University, Russia). She has been working since 2004 as the Head of the Department for Linguistics and Professional Communication in Foreign Languages at the Ural Federal University. From 2012 until present, she is leading the project on foreign language excellence at the same university. This project aims at introducing learning outcomes in connection with CEFR, and assessment of the outcomes - in connection with international exams. Her research interests are multilingualism, management of foreign language education, multiliteracies pedagogy.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



## Intersubjectivity, foreign language proficiency and the development of pre-service teachers' linguistic-communicative competences in teletandem interactions

This paper is about an investigation into language description and implications for language assessment that can contribute to teacher education and professional development. The investigation focuses on the role of the establishment of intersubjectivity (McCafferty 2002, Mori and Hayashi 2006, Vygotsky 1930/1971, 1978, Duranti 2010, Sambre 2012) in the development of teaching competences. Intersubjectivity is understood as a relationship of otherness and established between individuals as they recognise each other as individual subjects in social contexts. We report on analyses of teletandem data from online interactions within the project Teletandem: transculturality in online interactions via webcam (Telles 2011). The data were generated in a context of institutional and integrated teletandem interactions (Aranha and Cavalari 2014), involving interactants from a Letters course in Brazil and Portuguese as a foreign language students in a university in the USA. The study is based on the assumption that qualitative analyses of the teletandem interactions, from a hermeneutic-methodological perspective, reveal information about the linguistic domain that is relevant for the development of teachers-to-be's language proficiency for teaching purposes. The study can bring important insights for the Examination of Foreign Language Proficiency for Teachers (EPPLE) project, which aims at the development of the EPPLE examination (Consolo and Teixeira da Silva 2014) for foreign language teachers. The teletandem interactions were audio and video-recorded. The recordings were then transcribed and analysed by means of qualitative and quantitative procedures. The categories focused in the study contemplate verbal and non-verbal communication, the fact that the interactants are establishing the grounds for mutual knowledge, and variations in discourse genre, in the linguistic codes and sociolinguistic variants. The analyses indicate that intersubjectivity is related to dimensions of sociolinguistic and strategic competences, and that the establishment of intersubjectivity is strictly connected with the mutual knowledge between the interactants in teletandem communication, with traces of an affective, cognitive and cultural nature.

**Douglas Altamiro Consolo**, State University of Sao Paulo (UNESP), Brazil

Douglas Altamiro Consolo holds a PhD in Applied Linguistics from the University of Reading, UK, and is an Associate Professor at State University of Sao Paulo, Brazil (UNESP), where he works as a lecturer and researcher. He has published articles in specialized journals, full texts and abstracts in conference proceedings, book chapters and organized books. He supervises for MA dissertations, PhD theses and scientific initiation projects in language teacher education, assessment and testing, classroom interaction and technology in language education.

**Gerson Rossi dos Santos**, Federal Institute of Education, Science and Technology of Sao Paulo (IFSP), Brazil

Gerson Rossi dos Santos is a PhD student in Applied Linguistics at State University of Sao Paulo, Brazil (UNESP), holds an MA degree from the same university and teaches both Portuguese as a first language and English as a foreign language at the Federal Institute of Education, Science and Technology of Sao Paulo (IFSP). His language teaching experience ranges from basic education to post-graduate courses. His main academic concerns are language teaching and learning, teacher training, interaction and intersubjectivity.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

## Investigating the TestDaF benchmarking process

The Test for German as a Foreign Language (TestDaF) is a language test for foreign learners of German who plan to study in Germany or who require a recognised certification of their language skills. Because it is a high-stakes test, it is particularly important to keep all procedures comparable and objective. Yet, the rating of learners' productive skills (writing and speaking) relies on human, thus subjective, decisions. To reduce this subjectivity we provide our raters with regular trainings, rating grids and calibrated materials, that is, benchmarks. The benchmarking process, however, also relies on human judgement and might put the validity of the TestDaF scores at risk. The purpose of this study is to examine the following two questions focusing on the speaking section of the TestDaF: firstly, to what extent do the judgements of two different expert groups agree with each other when rating the same performances of test takers? And secondly, which aspects influence the final decisions in a group discussion? The first question will be addressed by a correlational study of two expert groups' final ratings. The second question will be investigated by a qualitative analysis of the recorded group discussions using the software NVivo.

**Claudia Pop**, TestDaF-Institut Bochum, Germany,

Claudia Pop has worked as a Test Development Officer at the TestDaF Institute since 2012. She is responsible for the speaking section of the TestDaF exam, organises rater trainings, and coordinates the inspections of test centres. Before joining the TestDaF Institute, she has worked as a language coach at Ruhr-University Bochum. She holds a BA in Philosophy and Educational Science and an MA in Second Language Acquisition.

## Language assessment and effective teaching and learning for English language learners in Florida

There are almost 5 million English language learners (ELLs) in the public schools in the United States, and many of them have no access to the optimal educational programs (Anyon 2005, Orfield and Lee 2004). For example, between the scores of ELLs and native English speakers there was a 36-point gap at the 4th-grade level, and a 44-point gap at the 8th-grade level, and this achievement gap has remained constant each year between 2002 and 2011 (National Center for Education Statistics 2013). According to the US Census Bureau, in 2011, 28% of school-aged students in Florida spoke a language other than English at home (2013). In Florida, the META consent decree, which is the framework for compliance with the federal and state laws, regulates the identification, eligibility and programmatic assessment for ELLs. Considering the stagnating achievement gap, questions related to the role of assessment and language instruction across the curriculum in a multilingual classroom are essential to ask to explore the current status, effective practice, and areas for improvement. This presentation reports on the findings of an empirical study that examined elementary school teachers' perceptions about the ELLs' program placement assessment, and progress assessment. Specifically, interviews were conducted to explore teachers' views on the effectiveness of these assessments, and the connection between language assessment and curricular decisions. In

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



addition, classroom observations were done to gain insights into the pedagogical approaches and language strategies that promote better teaching and learning for ELLs. In this presentation, first we will provide an overview of the current assessment system for ELLs in Florida, then we will share findings regarding the value of assessment as a drive for better teaching and learning, and conclude with recommendations for effective ways of integrating ELLs into the learning environment.

**Tunde Szecsi**, Florida Gulf Coast University, United States

Tunde Szecsi, PhD, is a professor of Early Education and English for Speakers of Other Language (ESOL) at Florida Gulf Coast University. For 13 years she has taught undergraduate and graduate courses in culture, language learning and child development. She earned her doctorate at the University of Buffalo, NY. Her research interests and scholarly publications focus on multicultural teacher education, heritage language maintenance, culturally responsive education, and effective teaching English as second/foreign language.

**Janka Szilagyi**, The College at Brockport, State University of New York, United States

Janka Szilagyi, PhD, is an associate professor and chair of the Department of Education and Human Development at the College at Brockport, State University of New York. She earned her doctoral degree at the University at Buffalo, NY, USA. Her recent research has focused on heritage language maintenance, language learning and teaching, developmental progressions in the learning of mathematical concepts, and job-embedded professional development for in-service mathematics teachers.

**Melissa Meehan**, Florida Gulf Coast University, USA

Melissa Meehan is a doctoral student at Florida Gulf Coast University focusing on Educational Leadership. She is a kindergarten teacher, teaching linguistically diverse children, and children with special needs. Her research interest is play and alternative seating in classrooms, and cultural and language diversity. She is a co-author of the research study titled Teacher candidates' views about play and its role in kindergarten classrooms. She presented at the Association for Childhood Education International's Global Summit in Costa Rica in 2016.

## Linking performance assessment to language scales: Challenges of the rating method effect

In test alignment with language scales, standard setting plays a central role in setting and validating cutscores which make test-based decisions on placing test-takers at appropriate levels in relation to specified performance standards (Kane 1994). In performance assessment, particularly, variances of cutscores not only arise from the standard-setting methods being used, but also rating-related issues which predominantly contribute to score variation and different score interpretations. However, no trace of such discussion has been found in the handful studies on standard setting (e.g., Papageorgiou 2010, Shin and Lidster 2016), whose primary focus is given to processes and methods of standard setting. The present study, therefore, examined effects of rating method on standard setting regarding 1) resulting cutscores, 2) accuracy of alignment, and 3) judgmental behaviours of standard-setting participants. The research context was the linking of a national oral English proficiency test administered to tertiary-level learners of English in China to China's Standards of English (CSE) (National Education Examinations Authority 2014). Thirteen language testing experts participated in two sessions of standard setting using the basket method (Council of Europe 2009). Each session focused on a representative sample of test-takers who were scored for their performances on the entire test (criterion-based scoring) or on individual tasks (task-based scoring) with the same set of rating

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

criteria and scales. Results showed criterion-based scoring led to significantly higher cutscores than task-based scoring. Moreover, criterion-based scoring outperformed task-based scoring in discriminating test-takers in line with the nine band levels set by CSE. Different points of focus and strategies for decision-making were reported by experts in the two standard-setting sessions. Results of the study indicated discrepancies in score meaning and misplacements of test-takers by different rating methods, which could affect experts' judgmental behaviors in standard setting, hence instability and inaccuracy in the result of test alignment.

**Xiaoyi Zhang**, Shanghai Jiao Tong University, China

Xiaoyi Zhang is a PhD candidate in language testing and assessment from Shanghai Jiao Tong University (China).

**Shaoyan Zou**, Shanghai Jiao Tong University, China

Shaoyan Zou is a PhD candidate in language testing and assessment from Shanghai Jiao Tong University (China).

## Looking into listening: The influence of context videos in computer-based assessment of listening comprehension

Authenticity is often considered to be one aspect of language assessment that may increase the validity of a test. Following this assumption, test tasks should show relevant characteristics of the target language use domain. In the case of academic contexts, learners are, in most situations, not only able to listen to but also see the speakers, their body language, lip movements, etc. Including these visual cues using input videos may therefore be considered construct relevant for academic listening and increase the generalisability of test results to real-life comprehension in a typical classroom environment. Most existing research investigating the influence of different input types (e.g. audio-only vs context or content video) on test takers' performance in listening assessment has been based on test scores and self-reported verbal data. Only a few studies have taken into account test takers' viewing patterns (Ockey 2007, Wagner 2010) and even fewer used eye-tracking as a research method (Suvorov 2015). This paper presents the results of an explorative qualitative study with 5-10 participants examining the use of context videos in two computer-based listening tasks for German as a foreign language, one monologue and one dialogue. The research design envisages the following methods: explorative study using eye tracking data for analysing test takers' viewing patterns while working on the two tasks followed by stimulated recall interviews with the participants. Perspectives for future research in this field will be offered for discussion.

**Leska Schwarz**, TestDaF-Institut, Germany

Leska Schwarz obtained her master's degree in Instructional Design and German Linguistics from the University of Freiburg im Breisgau, Germany. She's currently a Test Validation Officer at the TestDaF Institute. Her research interests include the assessment of the receptive skills with specific reference to the testing of listening comprehension, and the impact and possibilities of technology on testing.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



## Meeting student needs through informal assessment OR do I know what I need to know?

Commonly used external, standardised assessments that assess student performance in relation to national standards may not necessarily provide teachers with specific enough information to respond to learner weaknesses with immediate, appropriate instructional practices (Bolon, 2000; Martinek, 2011). One reason relates to unwanted backwash effects of standardized testing (Hudson, 2013; Jang 2013). Therefore, informal assessment tools that are specifically related to curriculum content allow teachers to identifying student needs directly and design appropriate learning or intervention support accordingly. This poster presentation introduces the audience to an informal, online curriculum-related assessment tool that an individual teacher can use in the English as a Foreign Language (EFL) classroom. While originally designed for the EFL context in Austria, the audience can gain insights into how to adapt this format for EFL teaching contexts in other countries. Additionally, ways to use the gained informal assessment information for follow-up interventions or remedial practices are discussed. This includes differentiations of a topic/content for different learner needs in one and the same classroom and guidelines to foster student self-monitoring skills (Friedl-Lucyshyn, 2011).

**Andrea Kulmhofer**, Federal Institute BIFIE, Austria

Andrea Kulmhofer is a trained EFL teacher currently working at the federal institute BIFIE in Austria and involved in the development of the national assessment of English at grade 8, levels A2 –B1 in the CEFR.

**Christina Schimböck**, Federal Institute BIFIE, Austria,

Christina Schimböck is a trained EFL teacher currently working at the federal institute BIFIE in Austria and involved in the development of the national assessment of English at grade 8, levels A2 –B1 respectively in the CEFR. She has experience in teaching middle and high school students as well as adults in various contexts. She is also trained to teach German as a second or foreign language as well as Spanish as a foreign language.

## The Language Centre at Pisa University faces the challenges of digital assessment in an evolving community

Over the last 25 years, language testing at the Centre has evolved from paper-based tests to computer testing and, more recently, on to digital assessment of students with specific needs, i.e. dyslexia. The Language Centre at Pisa University offers language courses and language testing within the university. During the 1990s, the Centre provided paper-based language testing for the Faculty of Engineering. The first challenge facing the Centre came about in 2003 when the Centre was asked to offer language testing for all the students at Pisa University who did not already hold an internationally recognised certificate providing evidence of their competence in a foreign language at B1 or B2 level. The high numbers of students who availed themselves of the services of the Centre for their language certificate obliged the Centre to move towards digital assessment with monthly sessions of around 600 students. Texts were selected and adapted for B1, B2 and C1 levels and examined reading skills. These were corrected automatically by the computer. Subsequently, with an ever-increasing number of

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

students participating in Erasmus programmes, the Centre assumed the role of language tester for outgoing students. In this case, testing was extended to all four abilities. The biggest challenge to date faced by the language testers is the increase in awareness of problems faced by students with learning disabilities such as dyslexia. In collaboration with the “Sportello Dislessia e DSA”, the Centre is now working on a project with researchers from the Department of Linguistics to create an English language course and test at B1 level tailored to their special needs. The first testing session for these students took place in June 2016 and provided essential feedback for further work in this area.

## **Susan Corrieri, Pisa University, Italy**

Susan Corrieri graduated with an MA from Edinburgh University in 1982. She has postgraduate certificates in English Language Teaching and has taught English at all levels since then. She joined Pisa University as an English language lecturer in 1998 and is a founding member of the university's English language certification team. She has participated in different projects within the field of language testing including the Consortium for Inter-University Integrated Access Systems (CISIA) for Italian university entry tests and B1 certification for university students with dyslexia.

## **Ida Brucciani, Pisa University, Italy**

Ida Brucciani holds a first-class honours degree in Foreign Languages from Pisa University and post-graduate qualifications in Linguistics. She is an English Language Lecturer at Pisa University where she has taught for 30 years, and has worked at the Language Centre since 1994. She has participated in different projects within the field of language testing including B1-C1 English Language Certification for Pisa University, the Consortium for Inter-University Integrated Access Systems (CISIA) for Italian university entry tests, and B1 certification for university students with dyslexia.

## **Verifica delle possibilità per una certificazione della lingua araba**

La ricerca mira alla verifica delle possibilità per una certificazione della lingua araba, come lingua straniera, e allo sviluppo di un proto-test basato su uno specifico profilo di apprendente in un preciso contesto di acquisizione, strutturato sugli attuali test delle lingue europee, in base agli standard e i principi del QCER, a partire da un'analisi storico-linguistica e sociolinguistica. La valutazione della lingua araba rappresenta un complesso oggetto di ricerca: il fenomeno della “diglossia” (Ferguson 1959) e la dimensione sacrale e di prestigio della lingua giocano un ruolo primario nell'analisi sociolinguistica di questo idioma, allo stesso tempo, oggi l'arabo è considerato una delle ‘lingue emergenti’ all'interno del cosiddetto “mercato globale delle lingue” (Calvet 2002, De Mauro, Vedovelli, Barni, Miraglia 2002). Tuttavia, la marcata variazione linguistica fra ‘norma’ e ‘uso’ rappresenta una notevole sfida per stabilire uno standard linguistico uniforme e condiviso, utilizzabile in un test linguistico, che possa assicurare il massimo grado di spendibilità sociale, obiettivo principale di una certificazione, in un mondo dove la crescente mobilità delle persone e il conseguente rilevante contatto linguistico incoraggiano la formazione di nuove identità. La prima fase della ricerca è incentrata sulla denotazione del modello teorico di lingua su cui basare il test, individuato nell'Arabo Standard Contemporaneo che meglio riflette i principi del QCER, definito a partire dalla descrizione degli aspetti sociolinguistici di questa lingua: la dimensione di prestigio, la variazione linguistica e la diglossia. Con la seconda fase, vengono delineati i principali profili di apprendimento tramite indagini costruite ad hoc, per poi

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era

definire i bisogni comunicativi, i contesti di acquisizione e d'uso e i parametri di valutazione della competenza. L'ultima fase comprende la messa a punto di un test pilota da somministrare ad un campione di studenti di lingua araba.

**Aisha Nasimi**, Università per Stranieri di Siena, Italy

Aisha Nasimi is a PhD student at the University for Foreigners of Siena. She graduated from the University for Foreigners of Siena in Linguistics and Intercultural Communication; her specialisation was Arabic language and culture. Her research interests include Arabic/Islamic studies and Arabic language assessing, learning and teaching. She has worked as translator, linguistic-cultural mediator and taught Italian as a foreign language. She has worked on the following publications: *Italianismi nel mondo arabo: l'italiano fra le altre lingue nei panorami linguistici urbani*, Cosenza L., Nasimi A. - *Studi Italiani di Linguistica Teorica e Applicata (SILTA)*.

## What can we learn from language assessment results with a statistical perspectives?

The Turkce Yeterlik Sinavi (TYS; Turkish Proficiency Exam) is a standardised exam administered to individuals from more than 38 countries by Yunus Emre Institute. YYS has been developed considering the proficiency levels under the CEFR-Turkish, and administered to give certificates for education and hiring applications. The first exam was administered in 2013, and since then the number of individuals who took the exam in different countries has been increased dramatically. This research aims to show the transition procedures from classical test theory to item response theory applications on the analysis of the YYS items. First, items in the Reading and Listening sections of YYS were examined using both classical test theory (CTT) and item response theory (IRT) techniques. In CTT, item difficulty and discrimination levels were examined, and the descriptive statistics of the selected distractors were reported to be used in the next step of the research. In IRT modeling, several models were tested based on the model fit indices, and items were again analysed with respect to the item difficulty, discrimination, and guessing levels. The origins of examinees based on their countries and native languages were then considered in order to perform item bias analysis towards the different groups of examinees using differential item functioning in IRT. All these findings were used to review items with the item writers and research group of YYS in order to analyse YYS's comparability across different backgrounds in native languages. In addition, the findings of this research will be beneficial for finding out how to teach the Turkish language to mixed linguistic backgrounds in the language education centres of Yunus Emre Institute. Finally, the research also led to the development of the computer-based version of the YYS to be administered next year.

**Mehmet Kaplan**, Artvin Coruh University, Turkey

Mehmet Kaplan has a MS degree in Statistics and PhD in Educational Statistics and Measurement from Rutgers University. His main research interests include adaptive testing, statistical modeling, and formative assessment procedures. Currently, he works as an assistant professor at Artvin Coruh University, and also as a researcher at the Measurement, Selection, and Placement Center in Ankara, Turkey.

**Nursel Tan Elmas**, Yunus Emre Institute, Turkey

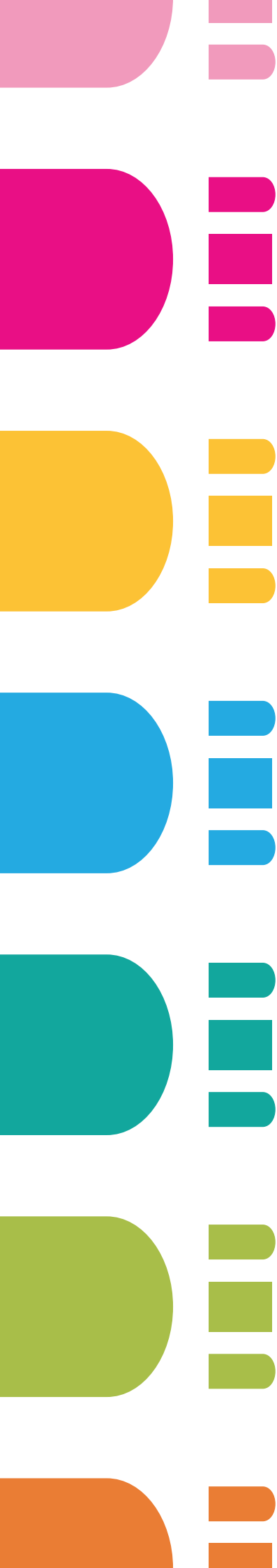
Nursel Tan Elmas has worked for Yunus Emre Institute for six years. Currently, she works as a Director of Test Center in Yunus Emre Institute in Ankara, Turkey. She has been working in her current position for around a year. She has a B.Soc.Sci. degree in Linguistics, and an M.Soc.Sci Degree in Teaching Turkish to Foreigner Students.

# Learning and Assessment: Making the Connections

... in a globalised economy ... in a multilingual world ... in the digital era



Her professional interests include effective teaching of Turkish as a foreign language, linguistics, semantic, collocations and web/internet-based assessment and evaluation.



Under the auspices of the Secretary General  
of the Council of Europe, Mr Thorbjørn Jagland

COUNCIL OF EUROPE

