

ResearchNotes

Editorial Notes

Welcome to Issue 9 of *Research Notes*, our quarterly publication reporting on matters relating to research, test development and validation within UCLES EFL.

To coincide with the introduction of the Certificates in English Language Skills (CELS) in May 2002, this issue contains a special focus on the new suite of modular English examinations. In an introductory article Lynda Taylor discusses the developing notions of *plurilingualism* and *partial competence* within the context of language learning/teaching and she explains how the modular approach of CELS fits in well with such concepts. Roger Hawkey describes the way in which CELS can trace its ancestry back to the early communicative tradition of language teaching/testing, and discusses how CELS owes a considerable debt to precursor tests such as the CCSE and Oxford-ARELS examinations. Nick Saville summarises the process of test development which was adopted for the new CELS suite; and in articles focusing on the CELS Writing and Speaking Tests, Stuart Shaw, Sharon Jordan and Lynda Taylor describe some of the test development and validation activity undertaken, in particular work done to confirm the relationship to one another of the different proficiency levels for CELS and their positioning within the Cambridge/ALTE/CEF Framework. A full account of the CELS development project will appear early next year in an edited volume in the Studies in Language Testing series published by UCLES/CUP.

Previous issues of *Research Notes* have reported on development and validation activity relating to the revised IELTS Speaking Test, introduced in July 2001. In addition to monitoring the success of the revised Speaking Test since July 2001, the IELTS team have also been concentrating their efforts on a review and revision of the assessment criteria and rating scales for the IELTS Writing Modules. In the first of a series of articles, Stuart Shaw reports on Phase 1 of this new revision project – Consultation, Initial Planning and Design.

2002 is proving particularly busy for us in terms of introducing revised and new examinations. The previous issue of *Research Notes* focused on our Business English examinations, with the recent revisions to BEC and BULATS. In this issue – in our section on recent validation studies – we report feedback on the examiner retraining programme for revised BEC and also on development work for the CPE revision. The first half of this year has also proved busy in terms of attendance and presentation by our staff and consultants at international conferences – see page 22 for more details. Lee Knapp also reports on the CELS launch conference – ‘Passports to Participation’ – held earlier this year in Birmingham.

Finally we are pleased to announce that from October 2002 UCLES EFL will officially be known as ‘Cambridge ESOL Examinations’. The section on page 19 explains the rationale for this change as well as giving other items of news which we hope you will find interesting.

Contents

Editorial Notes	1
Plurilingualism, partial competence and the CELS suite	2
Background to CELS: the communicative construct and the precursor exams	5
The test development process for CELS	8
CELS Writing: test development and validation activity	10
CELS Speaking: test development and validation activity	13
IELTS Writing: revising assessment criteria and scales (Phase 1)	16
Review of recent validation studies	18
Other news	19
Report on the CELS Launch	20
LTRC 2002 – Reminder	21
Conference activities	22
The Cambridge ESOL staff seminar programme	23

The URL for reading/downloading issues of

Research Notes is:

http://www.cambridgeESOL.org/rs_notes

The URL for subscribing to *Research Notes* is:

http://www.cambridgeESOL.org/rs_notes/inform.cfm

Plurilingualism, partial competence and the CELS suite

LYNDA TAYLOR, RESEARCH AND VALIDATION GROUP

Language learning within the European context

In 2001 the continent of Europe celebrated the business of language learning through the European Year of Languages, sponsored jointly by the Council of Europe and the European Union. A wide range of activities took place throughout the 45 participating countries with the aims of:

- raising awareness of and promoting the linguistic heritage of Europe (in the broadest sense);
- motivating all European citizens to learn languages, including those less widely used;
- supporting the development of *plurilingualism* and *pluriculturalism*;
- supporting lifelong language learning as a way of responding to economic, social and cultural changes in Europe.

The face and work of the European Union may be relatively familiar to some but the work of the Council of Europe is perhaps less well known. Founded in 1949, the Council of Europe is an intergovernmental organisation based in Strasbourg, France; its main role is to strengthen democracy, human rights and the rule of law throughout its 40 or so member states. To this end, and especially in order to improve international understanding, co-operation and mobility, the Council of Europe has helped to promote a communicative approach to language learning and teaching (see Roger Hawkey's article on page 5) which seeks to facilitate 'the freer movement of people and ideas'. Over the past 30 years or so, its work has been widely used in curriculum and examination reform, in course and textbook design and in teacher training. Models for defining language learning objectives have been published for twenty or more of the languages widely spoken in Europe. The first and perhaps best known of these – *Threshold Level* – appeared in 1980; other levels followed, such as *Waystage* and *Vantage*, and more – such as the *Breakthrough Level* – are currently under consideration. These models are now widely used to provide the basis for graded curriculum objectives and syllabus design as well as for testing and certification systems in many national and pan-European contexts.

The notion of plurilingualism

More recently, the Council of Europe has highlighted the concept of *plurilingualism* in its approach to language learning and teaching. *Plurilingualism* differs from the more well-established concept of *multilingualism*. The term *multilingualism* generally refers to the knowledge of a number of languages, or the coexistence of different languages in a given society; and it relates directly to the implementation of language policy in educational or

social contexts. *Plurilingualism*, however, takes a rather different focus: it recognises that an individual's experience of language in its cultural contexts expands – from the language of the home, to that of the society at large, and then to the languages of other people (whether learnt at school or by direct experience). As it does so, the individual does not keep these languages and cultures in strictly separated mental compartments; instead, he or she builds up a communicative competence to which *all* knowledge and experience of language contributes and in which different languages (including regionalised dialects) interrelate and interact. This means that in different situations, an individual can call flexibly upon different parts of this competence to achieve effective communication. For example, interlocutors may switch from one language to another, exploiting the ability of each to express themselves in one language and to understand the other.

Plurilingualism has itself to be seen within the larger context of pluriculturalism. Language is not only a major aspect of culture; it is also a means of access to the cultural aspects of our daily living. Much of what can be said about language applies equally to a person's cultural competence. The various national, regional, and social cultures to which a person has gained access do not simply coexist side by side: they are compared, contrasted and actively interact to produce an enriched, integrated pluricultural competence, of which plurilingual competence is one component.

Plurilingual and pluricultural competence tend to be uneven in one or more ways:

- learners generally attain greater proficiency in one language than in the others;
- the profile of competences in one language is often different from that in others;
- the pluricultural profile differs from the plurilingual profile.

As well as being uneven, plurilingual and pluricultural competence are not fixed or stable but tend to be transitory and changing in their nature. Depending upon our career path, family history, travel experience, or reading and hobbies, significant changes can take place in our linguistic and cultural biography.

Despite its uneven and changing nature, competence in less than the full range of language skills – known as *partial competence* – means that effective communication can take place. In many situations learners can bring *all* their linguistic resources to bear (in different languages) in order to communicate effectively.

The changing aims of language education

In the light of notions such as plurilingualism and partial competence, the aim of language education may need to be

modified. The goal is no longer the 'mastery' of one or two, or even three, languages, each taken in isolation, with the ideal native speaker as the ultimate model and target for achievement; instead the aim is to develop a linguistic repertory in which all linguistic abilities are of value and have a role to play, however partial or restricted some of these may be. This implies, of course, that the languages offered in educational institutions should be diversified and students given the opportunity to develop plurilingual competence, both during their school education years and beyond. The full implications of such a paradigm shift have yet to be worked out and translated into action. But it's interesting to note that some educational authorities in Europe are already encouraging children in schools to develop partial competence in a second or third foreign language, often with the focus on oral skills. And the Council of Europe's language programme has for some time been designing and producing tools for the language teaching profession to use in the promotion of plurilingualism and partial competence. One of these is a Council of Europe initiative piloted during the 1990s and launched in 2001 – the *Common European Framework of Reference for Languages: Learning, teaching, assessment* (2001).

The Common European Framework of Reference

The Common European Framework (CEF) provides a common basis for developing language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe. Its aim is to describe in a comprehensive way what language learners have to learn to do in order to use a language for communication, and what knowledge and skills they have to develop so as to be able to communicate effectively. It also defines levels of proficiency which allow the progress of language learners to be measured at each stage of learning throughout their lives; examination providers, such as ourselves, are one of the key target groups for which the CEF was developed.

The Framework includes the description of 'partial' qualifications, which are appropriate when only a restricted knowledge of a language is required (e.g. for understanding rather than for speaking) or when a limited amount of time is available for the learning of a third or fourth language and more useful results can perhaps be attained by narrowing the focus. For example, the priority may be to acquire reading comprehension skills for work or study purposes, or perhaps oral skills for social or tourist reasons. Educators are increasingly realising that their approach to teaching, and to testing, needs to take account of this trend.

Assessing language proficiency

Testing remains an important aspect of educational provision for language learning:

- it gives feedback to those involved as administrators, teachers or learners on their progress and the effectiveness of methods;

- it gives employers and other 'users' of language skills accurate and reliable information on the level of proficiency of the skills of those they may wish to employ or to admit to educational institutions.

For some years now, Cambridge ESOL has provided a wide variety of English language assessment tools directly linked to the CEF levels. Most of our tests are based upon the traditional 'mastery' view of language proficiency in which multiple competences – reading, writing, listening, speaking, use of English, etc – function in a more or less similar way at more or less the same level. For this reason most of our examinations include four or more skill-based components pitched at a similar level in terms of their difficulty; candidate scores on these different components are combined to produce a final overall grade on the examination. However, this widely accepted approach to assessing language proficiency does not readily suit learners whose language learning biography means they have acquired a partial competence in a given language, or for whom competence across the full range of skills may be very uneven. Clearly, an alternative approach is required to meet the needs of such learners and to provide a suitable method for certificating their progress. The suite of Certificates in English Language Skills (CELS) is just such an approach.

The modular approach in CELS

The Certificates in English Language Skills (CELS) make up a modular system of skill-based examinations. They allow for competence in reading, writing, listening and speaking to be assessed separately at one of three levels linked directly to the CEF levels. So candidates have the flexibility to choose to do one skill at a particular level or to build up a profile of skills at one or more levels. CELS is therefore able to offer partial competence testing by skill and by level.

The CELS suite is linked to the ALTE/Cambridge levels for assessment and to the Council of Europe's Framework for Modern Languages; it is also aligned with the UK Qualification and Curriculum Authority's National Standards for Literacy, within the National Qualifications Framework (NQF).

Equivalent Main Suite Exam	CELS	Council of Europe Level	UK NQF Level
CPE		C2 (ALTE 5)	
CAE	CELS Higher	C1 (ALTE 4)	Level 2*
FCE	CELS Vantage	B2 (ALTE 3)	Level 1
PET	CELS Preliminary	B1 (ALTE 2)	Entry 3
KET		A2 (ALTE 1)	
		A1	

* the level typically required for employment purposes to signify the successful completion of compulsory secondary education in the UK

The development of CELS, and the history of the four tests which have helped to shape its development, are described in more detail in articles by Roger Hawkey and Nick Saville (see pages 5 and 8). The new suite builds upon earlier teaching/testing traditions in order to provide a modern suite of modular examinations covering the four skills at three levels. It allows candidates to enter only for those components which are relevant to their needs or interests, and to enter each component at a level at which they have good chance of meeting the specifications. Different candidates can enter very different combinations of tests and levels, e.g.

	Reading	Writing	Listening	Speaking
Higher			✓	✓
Vantage	✓			
Preliminary		✓		

or they can choose to omit some tests altogether, e.g.

	Reading	Writing	Listening	Speaking
Higher				
Vantage		✓		
Preliminary	✓			

(In this regard the new suite increases the options for special needs candidates, some of whom may prefer not to take a test in a skill area relating to their disability.)

In short, CELS offers adult candidates (i.e. 16+) for whom English is not their first language the chance to gain certification for what they **can** do, without being penalised for what they **cannot** do. Can-do statements have been developed by the Association of Language Testers in Europe (ALTE) and these are directly related to the CEF levels to help describe what language users can actually do with the language.

With its modular approach and its direct link to the CEF levels, CELS fits in well with the notions of plurilingualism and partial competence discussed earlier. It also aligns well with another key initiative of the Council of Europe's languages programme – the European Language Portfolio. The European Language Portfolio is a personal document in which learners can record their qualifications and other significant linguistic and cultural experiences in an internationally transparent manner; this makes it easier for users, such as teachers and employers, to understand and interpret the language achievements of learners in an international context. And in turn, this helps to motivate learners and encourage life-long learning.

Conclusion

In a world of increasing population movement – whether this comes through free choice on the part of an individual or whether it is forced upon them as a result of adverse circumstance – there is no doubt that linguistic and cultural competence are of increasing importance, both within Europe and beyond European borders; linguistic and cultural competence are now recognised as key factors which enable individuals and their families to gain access to opportunities – whether these are social, cultural, educational, or to do with employment; they constitute a type of passport to participation.

Plurilingualism and partial competence are concepts which help us to understand better the role of languages and culture in a person's life, and especially the extent to which diversity of linguistic and cultural experience can be enriching and life-enhancing, rather than impoverishing and life-limiting. Interestingly, the Common European Framework document suggests that the promotion of respect for linguistic diversity and of learning more than one foreign language in school is not simply a linguistic policy choice at an important point in European history; nor is it even just a matter of increasing future opportunities for young people competent in more than two languages. It is also a matter of helping learners:

- to construct their own linguistic and cultural identity by integrating into it a rich and positive experience of otherness, and
- to enrich their potential for future learning as a direct result of relating to several languages and cultures

In one sense it might be argued that notions of plurilingualism and partial competence help to strengthen and preserve individual human rights; and perhaps they also contribute in some measure to social stability and economic prosperity within the wider community.

References and further reading

- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, teaching, assessment*, Cambridge University Press
- Ek, J A van and Trim, J L M (1991) *Threshold 1990*, Cambridge University Press
- Ek, J A van and Trim, J L M (1991) *Waystage 1990*, Cambridge University Press
- Ek, J A van and Trim, J L M (2000) *Vantage Level*, Cambridge University Press
- ALTE website: www.alte.org

Background to CELS: the communicative construct and the precursor exams

ROGER HAWKEY, CONSULTANT, CAMBRIDGE ESOL

The introduction in the *Handbook* for the Certificates in English Language Skills (CELS) examinations emphasises three aspects of the new exam:

- its modularity;
- its inheritance from precursor exams;
- its aim to assess English language competence “through authentic tasks based on authentic texts” (*CELS Handbook* page 7).

This article will focus on the latter two of these aspects of CELS, namely the test’s antecedents and construct.

Background to the exams from which CELS developed

The CELS exam has clearly been shaped not only by its function as a test of English language proficiency, but also by the nature of the exams from which it has been developed.

The examination in the Communicative Use of English as a Foreign Language (CUEFL), launched in 1981, was covered by a collaboration agreement in 1988 between UCLES and the Royal Society of Arts (RSA) under which the exam would be administered from Cambridge. November 1990 saw the first sitting of the extensively revised version of the CUEFL, now the Certificates in Communicative Skills in English (CCSE), which ran until December 2001. Both the CUEFL and the CCSE exams have been influential in the development of CELS.

In May 1995, the University of Oxford Delegacy of Local Examinations (UODLE) became part of the University of Cambridge Local Examinations Syndicate, then UCLES, now Cambridge ESOL. The principal reason for this merger was to rationalise the development of ‘A’ Level examinations, but it also offered the opportunity for the eventual rationalisation of overlapping UODLE and UCLES English language examinations. From 1996, UCLES administered the Oxford EFL exams which, as part of the merger agreement, were to continue in their existing form for three years. The last Oxford EFL exams were in fact administered in November 2000, but they too have been a significant influence on the new CELS exam.

The communicative approach to language teaching (CALT)

The *CELS Handbook* states that it “was decided in 1999 to merge the two sets of examinations” (Oxford and CCSE), implying a significant role for each in the development of CELS. It is interesting that an exam of the new millennium such as CELS, first administered in May 2002, should be seen to acknowledge so

significantly exams dating from as early 1978 (in the case of the Oxford EFL exam) and 1981 (in the case of CUEFL). But the dates of both these exams are significant, coinciding as they do with the beginning of the era of the communicative approach to language teaching. It was, after all, in the mid to late 1970s that the change of learning and teaching emphasis from language form to language function, advocated by psycho- and sociolinguists, discourse and needs analysts, gave impetus to the communicative approach to language teaching (CALT) and it really took off.

The message of CALT, as demonstrated in curricula and materials following its principles and approaches, can be summarised thus:

If we are to try to help our learners develop their competence to fulfil their communication needs for their academic or occupational, as well as their social, purposes, on the topics and in the settings with which they wish to operate, then we should try to ensure they experience and use the communicative notions, functions, micro-skills and activities defined as typical of and authentic to the domains concerned.

Figure 1 summarises CALT in terms of its main objectives, sample activity types and possible problems.

Figure 1: The communicative approach to language teaching

Objectives

- to develop in learners the ability to communicate in target language (TL) according to their needs
- to help learners develop relevant communicative notions, functions, and micro-skills
- to provide access to relevant authentic TL texts and tasks
- to develop communicative fluency

Sample classroom activities

- teacher and learners agree appropriate simulation project (e.g. promotion of a local product)
- teacher and learners discuss project media and modes
- learner groups work on an aspect of the promotion (e.g. leaflet, broadcast ad.)
- teacher and learners meet to discuss progress and problems, including teacher advice on language, media etc involved in the projects

Problems

- neglect of usage at the expense of use
- neglect of accuracy for fluency
- difficult activity and learning management

A key implication of the communicative approach has always been that learners should be exposed as much as possible to *authentic* discourse, media and tasks. This represented a clear departure from

the use of target language which was controlled or adapted (especially in terms of grammar and vocabulary), and which reflected the more stimulus-response patterns of earlier structuralist/audio-lingual approaches to language teaching.

All three of the CELS precursor exams claimed to adhere to the communicative approach. In the *Rationale, Regulations and Syllabuses* booklet (e.g. 1985) the 'nature' of the Oxford EFL examinations was defined with the following claim:

Wherever possible the materials used in the tests are based on authentic tasks and situations. Candidates are encouraged to see the purpose and relevance of what they are asked to do...

Under similar influences, the *CUEFL Information Booklet* (e.g. 1988) introduced the test thus:

This series of examinations is the result of a conscious effort on the part of the Royal Society of Arts Examinations Board to develop new testing procedures to match developments in the communicative teaching of foreign languages...

And the *CCSE Handbook* (e.g. 1998) claimed the following:

The aim of CCSE is to assess English language competence through a variety of authentic tasks based on authentic texts.

Since the communicative approach, with its focus on authentic language needs, is also the key construct for the CELS exam, it is no longer surprising that CELS is informed by the Oxford EFL, CUEFL and CCSE exams.

The communicative approach to language testing (CALTe)

Of course, the features of the CALT had profound implications for language testing. Milanovic and Saville (1996:3) note that language testing "followed in the wake of language teaching" and that "by the late 1970's there was considerable interest in testing communicative competence and in performance tests". Morrow (e.g. 1979:16–17), a key figure in both CALT and CALTe from the 1970s on, always considered that revolutionary change was required for language testing and evaluation "to take account of (communicative) development, in any systematic way":

Asking the question, 'What can this candidate do?' clearly implies a performance-based test. The idea that performance (rather than competence) is a legitimate area of concern for tests is actually quite a novel one and poses a number of problems, chiefly in terms of extrapolation and assessment (1979: 151).

Despite the problems inherent in testing language by means of authentic communicative tasks and texts, performance assessment tasks could and should be designed, in Morrow's view, to be:

- interaction-based;
- rooted in the context of a situation;
- linguistically contextualised;
- characterised by purposes, authenticity, and behavioural outcomes on which the performance of participants may be evaluated.

Figure 2 summarises key features of communicative performance test design, using the kinds of features proposed by Morrow and others, and influencing the development of communicative language tests such as Oxford EFL, CUEFL, CCSE and, indeed, the new CELS test.

Figure 2: Features of communicative performance test design

Global communicative tasks analysed into the underlying performance-based enabling (or micro-) skills required to be mobilised to complete them	
<ul style="list-style-type: none"> • operational scales of attainment, with levels of candidate proficiency defined by matching performance against operational specifications taking account of parameters such as the following (see Munby, 1978/B J Carroll, 1980): <ul style="list-style-type: none"> – size } of text which can be handled – complexity } of text which can be handled – range } of enabling skills, structures, functions which can be handled – speed } at which language can be processed – flexibility } in dealing with changes of, e.g. topic – accuracy } with which, e.g. enabling skills, structures, functions are handled – appropriacy } with which, e.g. enabling skills, structures, functions are handled – independence } from reference sources and interlocutor – repetition } in processing text – hesitation } 	
<ul style="list-style-type: none"> • assessment guidelines: <ul style="list-style-type: none"> – candidates to be assessed not in pass/fail terms but in terms of what they <i>can do</i> – candidate performance assessed in different communicative areas and skills, their proficiency reported as a communicative profile – types of operation, content areas and criteria adopted in assessment to be published 	

It is clear from Figure 2 that language tests with a communicative construct, and assessing life-like and usually holistic performance tasks across the skills of listening, reading, writing and speaking, may be very demanding in terms of validation, at the levels of construct, content validity, and marking and marker reliability checks. The pre- and post-test validation of the CELS exam according to Cambridge ESOL's latest validity, reliability, impact and practicality (VRIP) assurance systems are described in Nick Saville's article on page 8.

CELS and its relationship with the Oxford EFL, CUEFL and CCSE exams

The CELS exam certainly emerges with a communicative construct (assessing "English language competence through a variety of authentic tasks based on authentic texts") and has the following main features:

- **free-standing skills tests:** meaning that candidates can enter for one, two, three or four skills if they wish;
- **three levels:**
 - Preliminary* = Council of Europe level B1, ALTE Level 2 (e.g. PET)
 - Vantage* = Council of Europe level B2, ALTE Level 3 (e.g. FCE)
 - Higher* = Council of Europe level C1, ALTE Level 4 (e.g. CAE)

- **task-based:** e.g. leave a note for a friend who has agreed to help you set up a disco; listen to an information line message and note key facts; read an article to make notes for a project;
- **authentic texts:** reading texts unsimplified, taken directly from newspapers, novels, advertisements etc; listening extracts from radio, announcements, answerphone messages etc., not slowed down or edited; candidates dealing with 'real' English, completing sorts of tasks that native speakers do.

Figure 3 summarises some of the key relationships between the CELS exams and its Oxford EFL and CCSE (CUEFL) predecessors. It should become clear from the summary that the new CELS exams benefit considerably from the constructs, content and approaches of the older tests, and from improved test development, revision and management systems.

References and further reading

- Cambridge ESOL (2002) *Certificates in English Language Skills Handbook*. Cambridge: Cambridge ESOL
- Milanovic, M, and Saville, N (1996) *Considering the Impact of Cambridge EFL Examinations*. Draft
- Morrow, K (1979) Communicative language testing: revolution or evolution? In C. Brumfit and K. Johnson (eds.) *The communicative approach to language teaching*, Oxford: Oxford University Press
- Royal Society of Arts Examinations Board. (1985) *The Communicative Use of English as a Foreign Language*, London: RSA
- University of Oxford Delegacy of Local Examinations and the British Association of Recognised English Language Schools Examinations Trust (1985) *The Oxford-ARELS Examinations in English as a Foreign Language: rationale, regulations and syllabuses*

Figure 3: Continuity and changes from CCSE and Oxford to CELS

Oxford EFL	CCSE (CUEFL)	CELS
<i>Features of potential influence on CELS</i>		<i>Changes for CELS</i>
Two tests, Reading, and Reading and Writing, at each of two levels	Four tests, Listening, Reading, Writing, Oral. Interaction at each of four levels	Four tests, Listening, Reading, Writing, Speaking, at each of three levels
Only reading and writing skills covered; candidates must take both papers and at the same level Oxford EFL exam often taken along with the ARELS Oral test to cover all four skills	Four free-standing skills exams, with options of selecting different levels across reading, writing, listening, oral interaction modules	Four skills tests modularity retained; "Oral Interaction" test renamed "Speaking"
Two levels, Preliminary and Higher; Oxford Preliminary probably between CELS Preliminary and Vantage, Oxford Higher approximately equivalent to CELS Higher	Four levels: Threshold, Vantage, Effective proficiency, Mastery, changed from CUEFL's three levels (basic, intermediate and advanced), not pegged to external levels and "Higher" retained	Three levels (like CUEFL), not four like CCSE; CCSE top level dropped: CELS has Preliminary, Vantage, Higher benchmarked to Council of Europe levels; Oxford EFL names, "Preliminary"
Three pass grades (pass, credit, distinction) and two fail grades (narrow fail, fail)	Pass/ fail only, in terms of communicative effectiveness	Two pass and two fail grades: pass and pass with merit, fail and narrow fail
Papers contain realistic tasks, e.g. writing articles, letters, short notes, messages etc	Texts all authentic at all levels and tasks adapted to suit each level	Authenticity of tasks and texts retained
Reading texts presented with their original colour and illustrations	No colour	CELS Reading Test materials in colour
Text processing skills tested, including dictionary skills; use of English-English dictionary allowed	Range of reading enabling skills tested; use of English-English or bilingual dictionary in the Tests of Reading and Writing	Range of reading enabling skills tested; only English-English dictionary allowed
Candidates receive booklet with all texts, and are directed to the particular texts they need to read. Paper 2 had loose texts as well as interleaved ones; two/three tasks in Higher linked by a common situation. Candidates write answers to the Reading questions in the test booklet	Reading test candidates received a separate booklet with all texts in it. Reading and Writing test materials often based on a theme. Answers written on question papers	Texts interleaved with tasks in reading test instead of separate materials booklet no themed papers, but Writing Part 1 has related tasks based on some common input. Reading test answers directly on to answer sheet; Listening, answers on question paper, then transfer to the answer sheet; Optical Mark Reader (OMR) answer sheets
No listening / speaking tests (see above)	CCSE Oral Interaction test uses paired interviews, with Interlocutor, Assessor and an Usher, three rooms, candidates moving from room to room	CELS Speaking Certificate requires only one exam room and a waiting area; UCLES trained examiners fulfil functions of Usher, Interlocutor and Assessor
Papers follow similar but not identical patterns	Papers follow similar but not identical patterns	Set pattern for Reading, Listening and Writing papers in order to test different skills in each part of the test
Full or part tests drafted by teacher / examiners	Item writers invited to submit full draft tests	CELS items written by teams of writers, each task to be pre-tested and item banked, for greater consistency across papers

The test development process for CELS

NICK SAVILLE, RESEARCH AND VALIDATION GROUP

The late 1990s was a time of major systems development for UCLES EFL exams, with a particular emphasis on good practice, accountability, and research and development; during this period there were also four main influences on the revision and test development projects which were conducted at that time:

- the work on a Framework of levels undertaken by the Association of Language Testers in Europe (ALTE) and the related projects carried out in this context, in particular the Can Do Project;
- the work of the Council of Europe (including the revision of Waystage and Threshold Levels for which UCLES was the major sponsor), and more recently the development of the Common European Framework of Reference and related documents such as the Users Guide for Examiners;
- developments in theoretical models of Communicative Language Ability (CLA) and how to assess that ability;
- advances in technology.

It is not surprising that the development of a new suite of examinations to replace the Oxford-ARELS and the CCSE followed the approach, and built on the experience and expertise, which had previously been employed in the main suite revision projects for FCE and CPE spanning the 1990s.

This summary article is broadly based on Chapter 4 of the SILT Volume 17 which has been written to provide a full history of the CELS project (Hawkey, 2002, draft). This volume, in its turn, continues the theme of *continuity and innovation* in SILT Volume 15 which records the history and latest revision of CPE – 1913 to 2002 (Weir, ed., 2002 in press).

Previous issues of *Research Notes* have also contained articles which have dealt with this topic (for example *Research Notes 4* and *Research Notes 8*). In these articles the importance of consultation with stakeholder groups and the management of the “process of change” have been discussed.

Figure 1 presents an outline of the now well-established, UCLES model for test development or revision which was used for the CELS project.

This approach is both cyclical and iterative, and as part of the development/revision process, the essential “qualities” of the test or examination have to be balanced in order to ensure that it fulfils its intended purpose in a useful way.

A significant feature is the importance of both *quantitative* and *qualitative* methods for the establishment of validity, reliability, impact and practicality (VRIP), and the need to develop an explicit validation plan to ensure that all aspects are adequately accounted for.

Figure 1: Model for test development or revision

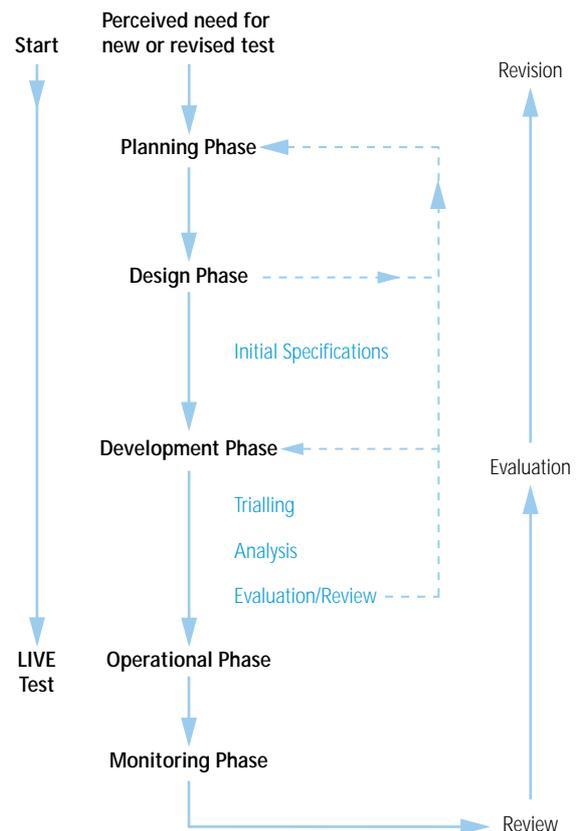


Figure 2 gives an outline of the validation plan for CELS.

In the “design and development phases” VRIP-based checklists were completed for each skill at each level of the proposed CELS system. The necessary information was collected through iterative cycles of consultation and trialling. Before the specifications and the revised procedures were finalised, the checklists were scrutinised by the Project Steering Group to ensure that all of the VRIP features were adequately dealt with and were in line with “best practice” employed for other Cambridge ESOL exams.

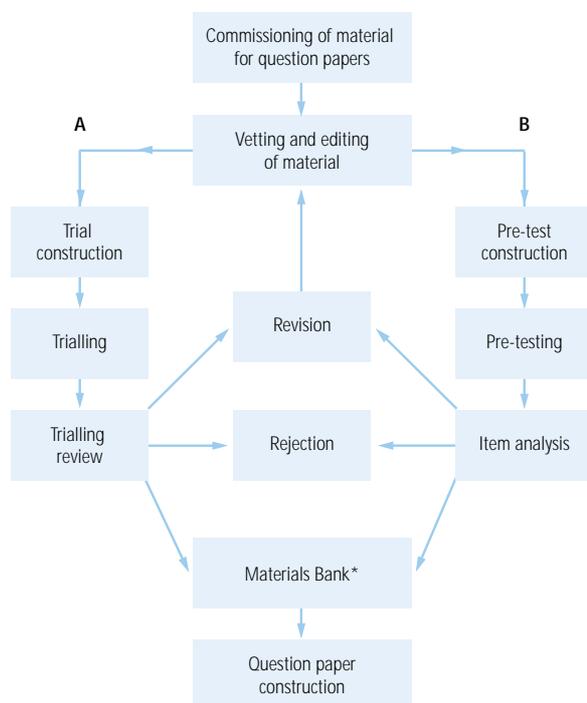
In terms of technical aspects of the system (such as mark-scheme development and rating procedures for Writing) a prioritised list of validation projects was agreed and carried out between August 2001 and April 2002, when the new exam system entered its “operational phase” (with the first of the regular administrations beginning in Summer 2002). See the articles on Writing and Speaking on pages 10 and 13.

As an operational examination system, CELS now benefits fully from the standardised procedures which are used in the production of other Cambridge ESOL exams – e.g. for main suite exams such

Figure 2: CELS Validation Plan

1 INTRODUCTION	
2 SALIENT ISSUES FOR CELS	
3 VALIDITY	
3.1 Content-related Validity	
	3.1.1 Format
	3.1.2 Situational authenticity
	3.1.3 Topics and themes appropriate to intended target-use contexts
	3.1.4 Makes use of appropriate, genuine sources
3.2 Construct-related Validity	
	3.2.1 Covers appropriate ability range
	3.2.2 Reflects underlying model of language ability and linguistic target-use contexts (interactional authenticity)
	3.2.3 Test specifications reflect underlying construct(s)
	3.2.4 Constructs can be operationalised using the formats
3.3 Concurrent validity	
	3.3.1 Compares well with existing paper and retains distinctive features
4 RELIABILITY	
4.1 Reliability estimates are adequate for the type of test	
4.2 Discrimination is adequate for the level and purpose	
4.3 Procedures in place to apply standard to marking and scaling of paper	
	4.3.1 Pre-testing and test construction procedures
	4.3.2 Maintenance of historical grading information
	4.3.3 Provision of information on paper difficulty and population ability from anchoring, pretesting, standards fixing
	4.3.4 Information enabling comparison of performance across related syllabuses and components
	4.3.5 Candidate information data
	4.3.6 Establish reliability of marking
	4.3.7 Establish procedures for examiner scaling
	4.3.8 Establish procedures for difficulty scaling of versions of papers
5 IMPACT	
5.1 Candidate feedback is positive from trials	
5.2 Teacher feedback is positive	
5.3 Chair and Principal Examiner feedback is positive	
5.4 Invitational meeting feedback is positive	
5.5 Key Item Writer feedback is positive	
5.6 Senior Team Leader feedback is positive	
6 PRACTICALITY (resource implications: human, material, time)	
6.1 Sustainability of format and items	
6.2 Layout	
6.3 Question paper production	
6.4 Packaging and despatch	
6.5 Cost effectiveness of marking	
6.6 Administration of CELS exam	
7 SCHEDULE	

Figure 3: The question paper production process (QPP)



* Electronic bank for pre-tested materials, storage of unpre-tested materials

as FCE. These standardised procedures for question paper production (QPP) were put in place for CELS as part of the development project. The operational QPP process, as illustrated in Figure 3, goes through a number of pre-determined stages, starting with the commissioning of material and ending up with the production of the live papers which meet very high standards of quality.

The first live examinations were conducted in May 2002 and, as for other EFL exams, the papers for CELS were returned to Cambridge for marking, analysis, grading and issue of results. The post-examination evaluation completed in June/July provided an early opportunity to carry out “monitoring” of the new exams under live conditions. This, in turn, will now feed into the on-going process of validation. In future sessions during 2002 and 2003, feedback on the *impact* of CELS on key stakeholder groups will be collected and evaluated (future articles in *Research Notes* will follow up on this).

And so the cycle begins again!

CELS Writing: test development and validation activity

STUART SHAW, RESEARCH AND VALIDATION GROUP
SHARON JORDAN, SPECIALISED EXAMINATIONS GROUP

Background to the CELS Writing test

CELS Writing is a test of writing in the context of general English proficiency in various employment, educational and training contexts. The aim of the writing test is to assess written language competence through a variety of tasks which as far as possible reflect real-life situations and test a wide range of writing skills and styles. The range of task types is specified for each level and increases in difficulty across the three levels.

In October 1999, a working group comprised of experienced CCSE and Oxford examiners, item writers and Cambridge ESOL Subject Officers, was set up to review CCSE and Oxford Writing papers, their task types, format and difficulty, marking criteria and levels, and to make recommendations for the CELS Test of Writing. From the research and discussions of this group, decisions were taken regarding CELS Writing task types, marking and assessment.

It was agreed that aspects common to both examinations would be carried forward to the new exam:

Task types

Task type and format would reflect the ethos of both CCSE and Oxford examinations and would not change dramatically. The function of very short tasks in the Oxford exam (fewer than ten words) would be covered through bullet points in longer tasks.

The authenticity of each task would be paramount. Tasks would always be contextualised – for each task, there would be a real-life purpose of some kind. Candidates would be asked to produce pieces of writing which would be authentic responses to communicative ‘real-world’ situations.

Number of tasks

There would be a fixed number of tasks at each level: four at Preliminary, five at Vantage and six at Higher. For a stand-alone communicative test of writing there would be a need for a variety of authentic communicative tasks in order to test a wide range of writing skills and styles and to demonstrate overall competence in writing.

Functions

Tasks would attempt to cover a wide variety of functions at each level.

Timing

The timing would reflect the number of tasks the candidates would have to complete. There would also be a ten-minute reading time allocation, though this would be included in the overall time allocation for the test.

Marking

It was decided that the CELS Test of Writing would be criterion-marked according to the system used for CCSE. The CCSE criteria and marking system were reviewed and compared with marking approaches used for the Main Suite examinations. It was then proposed that the Main Suite model would be used for the marking of CELS Writing (with modifications where necessary). Reasons for this included the facts that the Main Suite model is already successfully in use, allows for two levels of pass and fail, and that the use of Main Suite marking criteria and system of grading would give increased uniformity across the Cambridge exams.

A marking project was therefore set up to see if Main Suite descriptors could be used to mark CCSE and Oxford task types. From this project, FCE descriptors were found to be the most suitable, with modifications to wording and further research necessary in order to develop mark schemes for each of the three levels of the CELS Test of Writing.

Validating the CELS Writing test

The Cambridge ESOL examinations are designed around four essential qualities, collectively known by the acronym VRIP : Validity, Reliability, Impact and Practicality.

Successful examinations cannot be developed without due consideration being given to all qualities. The process of establishing the validity of a test is one of the basic concerns of language testing. Validity is generally considered to be the most important examination quality and it is desirable to establish validity in as many ways as possible as it concerns the appropriateness and meaningfulness of an examination in a specific educational context and the specific inferences made from examination results. Confidence in a test is directly proportional to the amount of evidence that is gathered in support of its validity. The validation process then is the process of accumulating evidence to support such inferences.

According to Weir (2002, forthcoming), approach to test development is based on a cyclical model :

“In this approach it is important for the test developer to consider, as part of the process, how the VRIP qualities of the test or examination can be balanced in order to achieve the overall validity or usefulness of the examination and to ensure that it fulfils a useful purpose in the contexts it is to be used”

With the increasing emphasis on the validation process reflected by this model, it is unsurprising that the proposed plan for the development of a new examination, CELS, follows a similar model (see Nick Saville’s article on page 8).

Examinations which test at a level, with passing and failing grades, are attractive particularly in the context of language study, within a school or some other setting. The examination can select material and elicit performance which is appropriate to the level, and which can impact positively on the learning process. But this can also raise issues for the vertical comparison of lower with higher levels. In validating CELS it is important to establish that horizontally each skill is tested at a similar level, and that vertically the three levels are related systematically. In the case of Writing it has been recognised that the setting of different kinds of task, with different degrees of challenge, complicates the comparison of performance across levels. As part of the validation of CELS Writing, two studies were undertaken:

1 The CELS Writing mark scheme reliability study

The comparative markscheme study was intended to equate the CELS global and task-specific markschemes to their Main Suite counterparts. The purpose of the study was, therefore, four-fold:

- To verify suitability of the CELS mark scheme;
- To ascertain how usable and transparent the CELS mark scheme was;
- To ascertain the range of score divergence across CELS raters; and
- To equate CELS and MS mark schemes.

The study focused on the standardisation process as the variable most critical to improving the assessment of writing and emphasises examiner collaboration in the refining of existing CELS mark schemes. In particular, it was intended to address a number of validation questions :

- Is the Main Suite markscheme operable on CELS scripts produced in response to pretesting?
- How do the Main Suite and CELS markschemes relate to each other?
- Can human raters use the CELS markscheme?
- Is the CELS mark scheme discriminating/functioning well?
- Should changes be made to existing CELS markschemes on the basis of their use?

An initial trial, Phase 1, took place using Preliminary Writing CELS scripts from Pretesting Round 1(March/April 2001). Phase 2 of the trial investigated a number of Vantage and Higher sample scripts representing a wide variety of tasks, range of scores and functionalities. The scripts were multiply re-marked by two teams of experienced, independent raters. One team used existing CELS global and task-specific mark schemes whilst the second team applied Main Suite mark schemes to the same scripts.

In Phase 1 a quantitative functional comparison of CELS and

Main Suite mark schemes was not suitable at Preliminary Level/PET (ALTE Level 2) as respective task formats for each examination are widely different and employ task-specific mark schemes which address entirely different language functions. Comparability of tasks at Vantage/FCE (ALTE Level 3) and Higher/CAE (ALTE Level 4) mean that Vantage and Higher scripts, on the other hand, can be compared with their Main Suite counterparts. Appropriate scripts from Pretesting Rounds 2 (July/August 2001) and 3 (October/November 2001) were used to form the basis of the Phase 2 comparative study.

In general, both global and task-specific mark schemes seemed to generate the same band scores. Where there were discrepancies, these were discussed between examiners and agreement quickly reached.

2 The CELS vertical writing re-marking study

The vertical writing re-marking study sought to establish, empirically, the relation of the three separate CELS rating scales to each other.

A sample of scripts collected during trialling of CELS at all three levels was re-marked by independent raters using one of the three markschemes. Level 2 scripts were re-marked using Level 2 and 3 rating scales, Level 3 scripts using all three rating scales, and Level 4 scripts using Level 3 and 4 rating scales (see Table 1).

Table 1: Mark scheme application design

Levels	Mark scheme		
	Preliminary	Vantage	Higher
Preliminary	✓	✓	✗
Vantage	✓	✓	✓
Higher	✗	✓	✓

Each script was thus rated using two or three rating scales; each rater used all the rating scales, and each rater rated every script in the sample. Each candidate, however, was rated on the basis of performance elicited by tasks at a single level, which corresponds to operational live test situation.

Three raters applying three mark schemes rated nine batches of scripts, A to I (each batch comprising 35 scripts) covering all three CELS levels as shown by the marking matrix (see Table 2).

The resulting dataset was analysed using FACETS (a multi-faceted Rasch analysis program). The findings from the analyses informed the design and interpretation of the rating scales and thus strengthened the vertical definition of the overall CELS Writing scale.

Although the coverage of issues within the validation plan is comprehensive not all of the questions can be answered by actual research and some questions are discussed and decided by working groups and the CELS Steering Group. The following are various considerations broached and, often, decided by such groups :

Table 2: Rater/mark scheme/level design

Raters	Levels			Mark scheme
	Preliminary	Vantage	Higher	
Rater 1	A	B		Preliminary
	G	E	C	Vantage
		H	I	Higher
Rater 2	D	H		Preliminary
	A	B	F	Vantage
		E	C	Higher
Rater 3	G	E		Preliminary
	D	H	I	Vantage
		B	F	Higher

- **VRIP** (an investigation of reliability levels – previous research on targets for internal reliability coefficients and factors affecting them; the possible use of CAN DO statements for CELS);
- **Levels** (linking CELS, ALTE levels and the Common European Framework; the relationships between CELS boundaries and equivalent Main Suite papers given that CELS is testing individual skills rather than the overall language ability; the use of a common anchor for as many levels as possible);
- **Skills Areas** (the relevance of CELS to partial competence testing; the format of all CELS modules – specific and specified text types, a specific length, authentic/realistic appearance, amenable to item banking and pre-testing);
- **Grading and scoring** (the number and name of CELS grade boundaries; an investigation of the CCSE ‘table-top’ model versus current trends in marking; how CELS Writing papers should be marked within the current IT systems).

Summary

A checklist has been completed, based upon the validation work undertaken on CELS Writing, in order to ensure that certain satisfactory standards are met in line with the established principles of good practice, covering validity, reliability, impact and practicality. The qualities of validity and reliability are widely understood in the language testing world. Whilst impact and practicality have been important considerations for Cambridge examinations, they have only recently emerged in the language testing literature (Bachman and Palmer, 1996). It is now generally recognised that each of the four individual examination qualities cannot be evaluated independently and that the relative importance of each needs to be determined, as has been the case with CELS, in order to maximise the overall usefulness of the examination.

References

- Alderson, J C and Wall, D (1993) Does Washback Exist? *Applied Linguistics* 14, 2: 115-129
- Alderson, J C and Buck, G (1993) Standards in Testing: a Study of the Practice of UK Examination Boards in EFL/ESL Testing. *Language Testing* 10, 1: 1-26
- Bachman, L F (1990) *Fundamental Considerations in Language Testing*. Oxford University Press: Oxford
- Bachman, L F and Palmer, A S (1996) *Language Testing in Practice*, Oxford University Press: Oxford
- Davies, A, Brown, A, Elder, C, Hill, K, Lumley, T and McNamara, T (1999) Dictionary of Language Testing. *Studies in Language Testing* 7, University of Cambridge Local Examinations Syndicate / Cambridge University Press: Cambridge
- Messick, S (1980) Test validity and the ethics of assessment, *American Psychologist* 30, 955-966
- Weir, C J (2002) Continuity and innovation: the history of an EFL examination 1913-2002, *Studies in Language Testing*, University of Cambridge Local Examinations Syndicate/Cambridge University Press: Cambridge

CELS Speaking: test development and validation activity

LYNDA TAYLOR AND STUART SHAW, RESEARCH AND VALIDATION GROUP

Introduction

The CELS Test of Speaking is a standalone test of speaking in the context of general English proficiency. The test lasts a total of 20 minutes and is divided into two parts each lasting 10 minutes (see Table 1).

Table 1 : Test format and timing

Part	Time	Interaction pattern	Input
<i>Preparation time – 1 min 30</i>			
1	10'	Candidates talk individually with Interlocutor on prompts they have chosen.	<ul style="list-style-type: none"> Written stimulus from task Oral stimulus from Interlocutor
<i>Preparation time – 1 min 30</i>			
2a	10'	Interlocutor sets up task; candidates talk together.	<ul style="list-style-type: none"> Written stimulus from task
2b		Three-way discussion between Interlocutor and candidates.	<ul style="list-style-type: none"> Written prompt on task sheet Oral prompts from interlocutor Two-way discussion from 2a

In designing the CELS Test of Speaking, the test development team took account of a range of factors. The test retains various features which were characteristic of its predecessor – the CCSE Test of Oral Interaction, e.g. paired format, task-based approach. (The Oxford examinations had no speaking component). The CCSE speaking test had its roots in the communicative language teaching

and testing tradition of the 1970s and the CELS Test of Speaking continues to reflect aspects of this 'heritage' (see article on page 5). However, some changes were made to the shape and content of the test in order to streamline procedures and to make the examination more straightforward for the centres to administer. Experience gained over recent years of designing and administering the Cambridge Main Suite Speaking tests at comparable proficiency levels (i.e. PET, FCE and CAE) also informed the development process; for example, the language of the Interlocutor has been more heavily scripted for standardisation purposes, and the assessment criteria and scales have been redeveloped in the light of Main Suite experience and work on the UCLES Common Scale for Speaking¹.

The CELS Test of Speaking is offered at three levels – Preliminary, Vantage and Higher; these three levels are designed to equate to the existing Cambridge/ALTE levels 2, 3 and 4, as well as to the Common European Framework levels B1, B2 and C1 (see Table 2).

Table 2: CELS and the framework of levels

CELS	Cambridge Main Suite	ALTE	Common European Framework of Levels
	CPE	Level 5	C2
Higher	CAE	Level 4	C1
Vantage	FCE	Level 3	B2
Preliminary	PET	Level 2	B1
	KET	Level 1	A2
			A1

¹ The UCLES Common Scale for Speaking was developed during the 1990s in order to provide a set of user-oriented descriptions of L2 spoken language proficiency at the 5 levels relating to the Cambridge Main Suite tests.

The CELS rerating study

A key goal in developing the new suite of CELS Speaking Tests was to confirm their horizontal link to equivalent levels or tests within a wider framework as well as their vertical relationship within the suite. A rerating study was therefore carried out in order to :

- confirm the relationship of CELS speaking proficiency levels to the UCLES Common Scale for Speaking;
- provide an empirical link between CELS speaking proficiency levels and performance levels as described by the ALTE CAN DO statements²;
- verify that CELS rating scales provide a sound vertical equating.

Method

Nine samples of CELS speaking test performance on video were selected for the study. The sample tests were taken from the CELS standardisation videos and covered the three CELS levels: Preliminary (two tests), Vantage (three tests) and Higher (three tests). Candidate performances in each sample test had previously been rated using the Main Suite Global and Analytic scales to place them at the same proficiency level as performances in PET, FCE or CAE.

Two experienced raters were asked to rate candidate performances in the nine sample tests using:

- (a) the UCLES Common Scale rating descriptors;
- (b) the ALTE CAN DO statements.

The two raters, in consultation, observed and 'blind' rated each test twice. The test performances were presented to them in randomised sequence to avoid the levels being self-evident. Both examiners were very familiar with the Main Suite levels and assessment approaches for PET, FCE and CAE, but were unfamiliar with the test format for CELS.

During the first viewing raters were asked to match the observed performances in each sample test to the Common Scale for Speaking levels 2, 3 and 4 using Common Scale Descriptors (see Table 3). The raters were required to make notes as they assessed each performance; following each test, the raters discussed their results and their discussions were audio recorded.

During the second viewing the raters were asked to link the observed performances to the CAN DO statements for level and skill. Speaking CAN DO statements corresponding to Preliminary, Vantage and Higher CELS levels had been selected and randomised to form one, new group of statements (Table 4). Again, the raters were required to make notes as they assessed each test and their follow-up discussions were audio-recorded.

² The ALTE Can Do statements provide a comprehensive description of what language users can typically do with the language at each level, in the various language skills and in a range of contexts. See previous articles in Issues 2 and 5 of *Research Notes* for more information on their development.

Table 3: Common Scale for Speaking

CPE – CAMBRIDGE LEVEL 5

Fully operational command of the spoken language

- able to handle communication in most situations, including unfamiliar or unexpected ones
- able to use accurate and appropriate linguistic resources to express complex ideas and concepts, and produce extended discourse that is coherent and always easy to follow
- rarely produces inaccuracies and inappropriacies
- pronunciation is easily understood and prosodic features are used effectively; many features, including pausing and hesitation, are 'native-like'

CAE – CAMBRIDGE LEVEL 4

Good operational command of the spoken language

- able to handle communication in most situations
- able to use accurate and appropriate linguistic resources to express ideas and produce discourse that is generally coherent
- occasionally produces inaccuracies and inappropriacies
- maintains a flow of language with only natural hesitation resulting from considerations of appropriacy or expression
- L1 accent may be evident but does not affect the clarity of the message

FCE – CAMBRIDGE LEVEL 3

Generally effective command of the spoken language

- able to handle communication in familiar situations
- able to organise extended discourse but occasionally produces utterances that lack coherence, and some inaccuracies and inappropriate usage occur
- maintains the flow of language, although hesitation may occur while searching for language resources
- although pronunciation is easily understood, L1 features may be intrusive
- does not require major assistance or prompting by an Interlocutor

PET – CAMBRIDGE LEVEL 2 (Threshold)

Limited but effective command of the spoken language

- able to handle communication in most familiar situations
- able to construct longer utterances but is not able to use complex language except in well-rehearsed utterances
- has problems searching for language resources to express ideas and concepts resulting in pauses and hesitation
- pronunciation is generally intelligible, but L1 features may put a strain on the listener
- has some ability to compensate for communication difficulties using repair strategies but may require prompting and assistance by an Interlocutor

KET – CAMBRIDGE LEVEL 1 (Waystage)

Basic command of the spoken language

- able to convey basic meaning in very familiar or highly predictable situations
- produces utterances which tend to be very short – words or phrases – with frequent hesitations and pauses
- dependent on rehearsed or formulaic phrases with limited generative capacity
- able to produce only limited extended discourse
- pronunciation is heavily influenced by L1 features and may at times be difficult to understand
- requires prompting and assistance by an Interlocutor to prevent communication from breaking down

Pre-Waystage Level

Zero

Table 4: Speaking CAN DO Statements for CELS Validation Study

1	CAN give a clear presentation on a familiar topic, and CAN answer predictable or factual questions.
2	CAN follow up questions by probing for more detail. CAN reformulate questions if misunderstood.
3	CAN ask for clarification, but this needs to be given sympathetically in order for it to be understood.
4	CAN keep up a casual conversation for a reasonable period of time, provided that this is of a mainly familiar, predictable nature.
5	When making requests, CAN deal with unpredictable replies and difficulties.
6	CAN ask questions, for example for reasons, clarification etc.
7	CAN keep up conversations of a casual nature for an extended period of time and discuss abstract/ cultural topics with a good degree of fluency and range of expression.
8	CAN express opinions on abstract/cultural matters in a limited way.
9	CAN present her/his own opinion, and justify opinions.
10	CAN keep up a conversation on a fairly wide range of topics, e.g. personal and professional experiences, events currently in the news.
11	CAN make critical remarks/express disagreement without causing offence.
12	CAN give simple information to a visitor about familiar places, for example her/his own school, city etc. CAN answer simple, predictable questions.

Following the two rating activities, both raters were asked to give general feedback regarding the study, CELS test format, findings and any possible recommendations. Their views were recorded.

Data analysis

Rater judgements together with any accompanying written reports were collected and analysed, along with the recorded consultations. The rater assessments derived from the Common Scale Band Descriptors were compared with the original assessments based on Main Suite Band Descriptors for the CELS standardisation process. Attempts were made to account for any discrepancies between the two and a list of issues raised by examiners was documented.

Results and discussion

Inter-rater agreement between the two examiners (calculated using Pearson product-moment correlation coefficients) was good (0.748). A good level of agreement was also achieved between each rater and the 'standardised' assessment; in nearly 95% of

observations, the two raters agreed with the original assessments awarded as part of the standardisation process.

With regard to the CAN DO statements, the raters were in general agreement about what candidates were able to do in terms of speaking proficiency at the Preliminary Level. At Vantage and Higher Levels, however, there was more variation between raters about which CAN DO performance descriptors should be ascribed to which candidates.

The raters commented on a number of other features of interest, including:

- the potential for distinguishing sub-levels within the CELS suite;
- the strengths and limitations of individual tasks;
- the role of the interlocutor in relation to support/intervention;
- implications for examiner training and standardisation;
- the role of accuracy as a performance criterion;
- the timing of the speaking test;
- the consistency of rater judgements;
- the limitations of the CAN DO statements in relation to assessment.

Their observations and recommendations fed directly back into the test development and validation process for the CELS speaking test.

Conclusion

The CELS rerating study revealed good agreement between examiner assessments of CELS candidates (based on Common Scale Band Descriptors) across the three levels of proficiency: Preliminary, Vantage and Higher. Moreover, these assessments compared very favourably with 'standardised' ratings based on Main Suite Band Descriptors. In this sense, both CELS and Main Suite performance thresholds have now been specified in Common Scale terms. Performance on CELS has also been empirically linked to CAN DO statements. Ongoing monitoring of the operational test following its introduction in May/June 2002 will enable further confirmatory work to be done on the performance level thresholds for CELS speaking and the position of CELS within the wider framework.

IELTS Writing: revising assessment criteria and scales (Phase 1)

STUART SHAW, RESEARCH AND VALIDATION

The choice of appropriate rating criteria and the consistent application of rating scales by trained raters are regarded as key factors in the valid and reliable assessment of second language performance (Alderson, Clapham and Wall 1995; MacNamara 1996; Bachman and Palmer 1996). Criteria and rating scales, as well as test content and format, need to be kept under review during the operational life of any test and are likely to require revision from time to time. Approaches to rating scale revision and some of the methods Cambridge ESOL adopts when revising assessment criteria and rating scales were described in Issue 3 of *Research Notes* (November 2000).

This article is the first in a series which will report on a project to revise the assessment of the IELTS Writing Test. The revision project aims to enhance the validity and reliability of the writing assessment process for IELTS by redeveloping the criteria and scales, and by developing a more comprehensive set of materials and procedures for (re)training IELTS raters.

Assessing writing in IELTS

The International English Language Testing System (IELTS) is designed to assess the language ability of candidates who need to study or work where English is used as the language of communication. IELTS tests ability in all four skills – Reading, Writing, Listening and Speaking. (Full details of the test can be found in the IELTS Handbook or via the IELTS website at www.ielts.org).

The current Writing Test lasts one hour and consists of two tasks. In Task 1 candidates are asked to either look at a diagram or table, and to present the information in their own words (Academic Module), or to respond to a given problem with a letter requesting information or explaining a situation (General Training Module). In Task 2, candidates are presented with a point of view or argument or problem (both Academic and General Training Modules). It is suggested that about 20 minutes is spent on Task 1 which requires candidates to write at least 150 words. Task 2 requires at least 250 words and should take about 40 minutes.

Each of the two tasks is assessed independently and the assessment of Task 2 carries more weight in marking than Task 1. Task 1 scripts are assessed on the following criteria: Task Fulfilment (TF), Coherence and Cohesion (CC), and Vocabulary and Sentence Structure (VSS). Task 2 scripts are assessed on performance in the following areas: Arguments, Ideas and Evidence (AIE), Communicative Quality (CQ) and Vocabulary and Sentence Structure (VSS). Detailed band descriptors have been developed to describe written performance at each of the 9 IELTS bands. These exist in two formats: as three 'profile' or analytical scales for each task (e.g. TF, CC and VSS for Task 1), and also as a global or holistic scale (i.e. the descriptors for TF, CC and VSS are conflated

into a single set of band descriptors). Examiners are able to select the global or profile approach according to whether a script has a 'flat' or 'uneven' profile.

The Revision Project

The revision project began in June 2001 with three main objectives:

- (1) the development of revised rating scales, including definition of assessment criteria and revised band descriptors (Task 1 and Task 2 for the General Training Module and the Academic Module);
- (2) the development of new materials for training trainers and examiners;
- (3) the development of new certification/re-certification sets for examiners.

It was agreed that the IELTS Writing revision project should closely model the approach successfully used for the earlier IELTS Speaking Test Revision Project, and would be divided into the following five phases:

Phase 1	Consultation, Initial Planning and Design	June – December 2001
Phase 2	Development	January – June 2002
Phase 3	Validation	July 2002 –
Phase 4	Implementation (incl. examiner retraining)	To be decided
Phase 5	Operation	To be decided

This article reports on Phase 1 – the process of consultation, initial planning and design.

Consultation, Initial Planning and Design

Initial discussion by the Revision Working Group was informed by a review of recent commissioned and non-commissioned studies relating to IELTS Writing, and also by a comprehensive survey of the literature on holistic and analytic approaches to writing assessment. This led to a decision to explore current practice among IELTS Writing assessors, in order to gauge their attitudes towards their respective assessment practice and to highlight theoretical and practical factors which might help shape the redevelopment of the writing assessment criteria and scales.

Current writing assessment practice

The consultation phase began with a series of semi-structured interviews with groups of IELTS Academic and General Training Writing assessors in the UK and Australia. These interactions led to the construction of a survey questionnaire which was sent out to a sample of several hundred IELTS assessors based at a range of test

centres worldwide. The function of the interviews and questionnaires was to elicit from assessors individual approaches and attitudes to the assessment of IELTS writing tests, especially in relation to differing domains (Academic and General Training) and differing task genres (Task 1 and Task 2). Protocol analyses are capable of revealing rich insights on the part of assessors which can be instrumental in helping to develop assessment criteria and scales that are valid, reliable and practical.

The questionnaire, which was designed to be concise and able to be completed in a short time frame, consisted of four sections the findings for which are reported below:

Task 1 and Task 2

More than three-quarters of Academic and General Training examiners rate Task 1 responses before they rate Task 2 and many examiners prefer to assess an entire candidate – Task 1 followed by Task 2 – before assessing the next candidate. The next favoured strategy is to rate tasks in batches choosing to rate all Task 1s first. More than four-fifths of all examiners give serious consideration to response relevance and response appropriacy when assessing responses.

Approach to rating

The *Writing Assessment Guidelines* (WAG) are not always made available to practising examiners and not all examiners are fully conversant with the WAG. However, Academic examiners do tend to consult the WAG as often as their General Training counterparts.

Approaches to rating vary and do not necessarily conform to procedures in the WAG. Although there is a general tendency to Profile rate, many examiners initially Global rate in order to gain an overall impression. Profile rating is often used to confirm initial impressions. Interestingly, only half of examiners read responses at least twice and whilst reading a response most examiners tend not to concentrate on one criterion at a time. More than half of all assessors retain information related to all three profile band descriptors whilst rating.

Assessment criteria and band descriptors

Nearly all examiners believe that the band descriptors are clearly worded and over three-quarters of them believe the descriptors to be easily interpretable. However, a third of examiners do encounter difficulties applying band descriptors to scripts.

Whilst nearly three-quarters of Academic examiners claim to readily understand the assessment criteria for Task 1 and Task 2, only slightly more than half of their General Training counterparts admit to fully understanding the criteria. Virtually all examiners consult the Assessment Criteria and Band Descriptors *during* assessment of a particular script (as opposed to either before or after assessing a particular script) with more than three-quarters paying attention to *Task Fulfilment* before any other criteria.

With regard to Task 1, *Coherence and Cohesion* and *Vocabulary*

and *Sentence Structure* seem to be given equal weighting in terms of their contribution to the rating process. Academic and General Training examiners appear to adopt similar strategies when assessing *Task Fulfilment* – almost three-quarters of Academic raters compare actual and expected responses whilst nearly two-thirds of General Training examiners make the same comparison. Nearly all the examiners surveyed agree that *Coherence and Cohesion* is the most difficult criterion to assess, followed by *Task Fulfilment* and *Vocabulary and Sentence Structure*.

In relation to Task 2, *Communicative Competence* and *Vocabulary and Sentence Structure* seem to assume equal weighting in terms of their contribution to the rating process. Both Academic and General Training examiners appear to adopt similar strategies when assessing *Arguments, Ideas and Evidence* – almost two-thirds of Academic raters conduct a detailed study of the features of the response whereas slightly more than half of General Training examiners perform the same analysis. Only a third of all examiners are concerned with overall, general impression. Most examiners agree that *Communicative Quality* is the most difficult criterion to assess, followed by *Arguments, Ideas and Evidence* and *Vocabulary and Sentence Structure*.

In general, examiners are satisfied with the accuracy of their final award.

Examiner background

Number of years experience as an examiner

Academic writing examiners			General training examiners		
Mean	Max	Min	Mean	Max	Min
4.3	12.0	0.25	4.8	15.0	0.25

Number of years experience as an EFL/EAP teacher

Academic writing examiners			General training examiners		
Mean	Max	Min	Mean	Max	Min
13.1	33.0	1.0	13.5	33.0	3.0

From the information presented, it would appear that many examiners have a wide experience of teaching and examining although a number of relatively inexperienced EFL/EAP teachers have limited experience of IELTS writing assessment.

Conclusion

Phase 1 of the project, which involved consultation with a range of stakeholders, was completed to schedule. It highlighted some key issues from the perspective of the assessor which have provided a valuable focus for the subsequent development phase: e.g.

- variation in sequencing of rating;
- variation in reference to WAG;
- variation in use of global/profile approaches;
- interpretability of particular criteria.

The Revision Working Group in conjunction with external participants (academic consultants and senior examiners with a strong interest in Academic Writing as well as experience with IELTS and international students in the university context) are currently completing Phase 2 of the project. Phase 2 – which entails the design and development of revised draft criteria and descriptors in preparation for trialling and validation – will be reported in *Research Notes 10*.

Review of recent validation studies

Investigating the CPE word formation cloze task

One of the consequences of the CPE revision project, and in particular the revision of Paper 3 – Use of English, has been the emergence of several key issues related to the inclusion of a word formation cloze task. One such issue relates to word formation task difficulty. A research question of profound interest and one which has wider implications for corpus data is the extent to which certain identifiable item characteristics contribute towards item difficulty.

A preliminary study into aspects of difficulty in the word formation cloze investigated certain elements of the word formation item and its associated difficulty for a data set comprising FCE, CAE and CPE items.

Multiple linear regression, a statistical technique that aims to provide valid models to describe the variation of a response variable (item difficulty) in terms of a small number of explanatory variables (key word frequency, compounding, multiple affixation, lexical transformation, etc.) in such a way that their inter-dependent effects are accounted for in each model, was applied to the data.

The results suggested that multiple regression is able to explain a proportion of the item difficulty but on this showing probably not enough to enable item writer specifications to be predictive of task difficulty. The analysis revealed that the predictor variable EXAM contributed most to the prediction of item difficulty, i.e. item writers know how to target a level in ways not wholly captured by the variables explained in the study. In the absence of the predictor variable EXAM, the frequency of key words (KeyLogFreq.), the number of transformations (Ntransforms) and the nature of those transformations (changeCode) contributed most to the prediction of item difficulty.

Reviewing the retraining of BEC Oral Examiners

The introduction of the revised Business English Certificates in March 2002 necessitated the worldwide retraining of Oral

References and further reading

- Alderson, J C, Clapham, C and Wall, D (1995) *Language test construction and evaluation*, Cambridge: CUP
- Bachman, L F and Palmer, A S (1996) *Language Testing in Practice*, Oxford: OUP
- McNamara, T (1996) *Measuring Second Language Performance*, Harlow: Longman

Examiners for BEC, to familiarise them with the new Speaking Test formats and to standardise their approach to assessment. During the retraining exercise trainees were asked to complete feedback forms which would allow us to:

- assess the impact and effectiveness of BEC examiners;
- collate examiner questions and comments arising from the training exercise for inclusion in an FAQ document for circulation to BEC trainers.

(A similar exercise was undertaken for the revision of the IELTS Speaking Test in 2001 and was reported in *Research Notes 6* – November 2001.)

The retraining materials pack included a feedback form to be completed by examiners. The form contained eight statements which respondents were asked to rate as follows: 1 – Strongly disagree, 2 – Disagree, 3 – Agree, 4 - Strongly agree. The eight statements were designed to elicit reactions to the main changes to the BEC Speaking Test format and materials, and the perceived effectiveness of the training session. A total of 166 examiner feedback forms were received and analysed. Results for the statement ratings are shown below in Table 1.

Table 1: Examiner feedback form statements - average ratings and rankings

Statement	Average	Rank
1. The revised Speaking Tests are an improvement on the current tests.	3.33	3
2. The topics and tasks are appropriate for the candidates.	3.20	7
3. The revised criteria are easy to use.	3.23	6
4. The revised frames are easy to use.	3.17	8
5. The revised Part 2 task will generate sufficient language to be assessed.	3.28	4=
6. There is sufficient support from the back-up questions.	3.28	4=
7. I am clear about the changes that have been made to the test.	3.55	2
8. I am confident that I will be able to examine in the new format.	3.61	1

All statements attracted an average rating of above 3 (i.e. agreement with the statement), suggesting that examiners are satisfied with changes to the BEC Speaking Test and feel confident about their ability to examine using the new format and assessment scales.

Qualitative feedback from the 166 examiners was analysed separately, together with comments from a further 42 examiners received on alternative forms. A total of 186 comments were recorded and coded into two main categories: *Test content and procedure*, and *Retraining session – content, experience and materials*. This type of feedback is used to inform production of

ongoing Oral Examiner training programmes, as well as potential changes to the test in future revision projects.

The most common questions/comments raised by examiners were adapted for inclusion in a document entitled *Revised BEC Speaking Test – Frequently Asked Questions II – April 2002*, circulated to all BEC trainers.

Feedback forms continue to be received and will be analysed in an ongoing manner. Plans are in hand to adopt a similar approach in order to gather data on the implementation of examiner retraining for the revised CPE Speaking Test later this year.

Other news

UCLES EFL adopts 'ESOL' identity

From October 2002 UCLES EFL will officially become 'Cambridge ESOL Examinations'. We have chosen this title to reflect the increasing diversity of our candidates, ranging from those in the 'traditional' language school sector to the growing number of people living, studying and working in English-speaking countries for whom English is not the first language. ESOL stands for 'English for Speakers of Other Languages'. It is already widely used in the UK, USA and elsewhere, as a more inclusive alternative to EFL (English as a Foreign Language).

UCLES EFL – the division of the University of Cambridge Local Examinations Syndicate that provides qualifications for learners of English – is adopting the title Cambridge ESOL for all of its activities. The full title of the EFL division of the UCLES group will change to 'University of Cambridge ESOL Examinations'. The change of identity is a recognition of the broader needs and backgrounds of learners and teachers throughout the world, and the many different reasons for which they take the examinations. It is also anticipated that the change to Cambridge ESOL will help to extend the worldwide recognition of the examinations by governments, educational institutions and employers.

IDLTM

The International Diploma in Language Teaching Management is a high level qualification which enables Language Teaching professionals to apply insights and skills derived from management theory and practice to their work as Language Teaching managers. The Diploma is run jointly by UCLES EFL, the School for International Training in Vermont and the Institute of Continuing and TESOL Education at The University of Queensland. More information is available from www.idlrm.org

Examination handbooks

Cambridge ESOL provides handbooks for each of the examinations we offer. Where an examination is new or has been updated, the handbook also includes an audio CD containing sample listening material. Handbooks can be ordered by sending an email request to eflinfo@ucles.org.uk; alternatively, they can be downloaded from our public website at www.cambridgeESOL.org/support/downloads/

First examinations for CELS

The first examinations for the Certificates in English Language Skills took place in May this year with over 10,000 entries for the first session. Full details on CELS are available at www.cambridgeESOL.org/exam/cels. Summertown Publishing are working on a series of textbooks for CELS, the first of which is scheduled for publication in August of this year; further information is available from www.summertown.co.uk/cels.html

New in-country websites

Three new Cambridge ESOL websites have recently been launched for users in:

- France (www.cambridge-efl.org/france)
- Italy (www.cambridge-efl.org/italia)
- Germany (www.cambridge-exams.de)

Similar sites are currently being prepared for Spain and Brazil.

Report on the CELS launch

LEE KNAPP, DEVELOPMENT MANAGER, UK

To mark the launch of CELS, an audience of well over 100 delegates attended "Passports to Participation", a conference which was held at the Britannia Hotel, Birmingham on 22 February 2002; the conference was designed to put a spotlight on how the new suite of examinations might contribute to the assessment of ESOL in the UK and beyond.

The audience represented a wide spectrum of interests within the ESOL community, drawn from schools, colleges and universities in both the public and private sectors, as well as representatives amongst others from the British Council, the Universities and Colleges Admissions Service (UCAS), the Prison Service and the Refugee Council. Course book authors and publishers were also present to hear four leading speakers each give a very thought-provoking but also very different perspective on meeting the needs of ESOL learners.

1. "Joint successes, shared challenges"

Dr Tony Archdeacon, Qualifications and Curriculum Authority (QCA)

The conference opened with a presentation by Tony Archdeacon, who as Principal Subject Officer, English, at the QCA has been deeply involved with work to establish national standards for ESOL in the UK, to accredit qualifications such as CELS, and to integrate ESOL awards into the National Qualifications Framework. "A Fresh Start", the report produced by Lord Moser in 1999, exercised a major influence in realising developments such as these, placing its absolute emphasis on improving national standards in education through a national curriculum, with national assessment and nationally accredited qualifications. The QCA's own three-fold role - regulating assessment and qualifications, establishing criteria for accreditation, and developing curriculum content - is specifically targeted at ensuring that the improvements in educational standards can take place.

When the QCA commenced its review of qualifications it quickly became apparent that, to make provision consistent, transparent and accessible to learners, rationalisation was essential. To consolidate provision, accreditation criteria and codes of practice were provided as guidelines, qualifications were grouped together and categorised, and titles were standardised. Tony Archdeacon was pleased to announce to the conference that over 100 awarding bodies have so far (February 2002) been accredited across a range of qualifications in different subjects.

Adult literacy qualifications were the first to be accredited, and with the arrival of many new learners in the UK the QCA was keen to make similar progress with English language qualifications for non-native speakers. For all UK residents, whatever their mother

tongue or home language, it was felt that English language qualifications should be commonly valued, and it was in this spirit that the division between ESOL (English for Speakers of Other Languages) and EFL (English as a Foreign Language) was considered no longer valid, and that indeed ESOL qualifications should be linked to mainstream English qualifications. Tony Archdeacon was pleased to acknowledge that CELS was one of the first qualifications to match up to the accreditation criteria and to be brought into the National Qualifications Framework.

Looking to the future, the impact of the recently published (February 2002) "Adult ESOL Core Curriculum" will be significant and extensive, for learners, teachers and testers amongst many others. Other issues on the QCA agenda for English included pre-entry and higher levels, qualifications for young learners, the assessment of spoken English, and the international recognition of UK accredited qualifications. In the meantime, to conclude his presentation Tony Archdeacon expressed the view that the successes of the first round of accreditation achieved by the QCA and the awarding bodies together boded well for the future.

2. "Partial competence and the contribution of the Certificates in English Language Skills to assessing ESOL"

Dr Lynda Taylor, Cambridge ESOL

The second presentation was given by Lynda Taylor, whose article elsewhere in this issue expands upon her comments at the Birmingham conference on how the key concepts of plurilingualism and partial competence are realised in CELS.

3. "ESOL - basic skill or basic right?"

Chris Taylor, National Institute of Adult Continuing Education

Opening the afternoon session Chris Taylor, Basic Skills Development Officer at NIACE, challenged her audience to consider whether ESOL, whilst categorised as a basic skill for non-native speakers of English living and learning in the UK, should more crucially be made available as a basic right.

The plight of many ethnic minority groups in England and Wales was spotlighted by "Lost Opportunities", a report produced by the Basic Skills Agency in 1996, in which one quarter of the 1,170 respondents could not even attempt such simple tasks in English as reading a calendar or filling in a library card. Three quarters of those surveyed were described as having a level of English "below survival", and the report concluded that the findings indicated a substantial unmet need for ESOL classes. In a further report "Divided by language", produced by NIACE itself in 2001, it was revealed that while just 11% of UK white

adults speak two or more languages additional to their mother tongue, this figure rises to 45% amongst black and ethnic minority groups. From this Chris Taylor concluded that language learning is a distinct feature of migrant groups – provided that the opportunity is made available to them. However, it is broadly accepted that the current ESOL provision in the UK is of mixed quality and is not easily accessible, and research on the baseline needs of ESOL learners remains scant.

Referring to extracts from “Secure Borders, Safe Haven”, the UK Home Office White Paper on Immigration and Asylum, Chris Taylor brought her presentation to a conclusion by echoing the views that “...economic migration and the seeking of asylum are as prevalent today as they have been at times of historic trauma ...” and that “... our society is based on cultural difference, rather than assimilation to a prevailing monoculture.” Whilst recognising the need for English language tests for citizenship and acknowledging the benefits of the Highly Skilled Migrant Programme, however, these initiatives would serve little purpose if ESOL provision remains insufficient, inappropriate, and neither intensive nor tailored to need. Teaching the English to get the job must not be confused with the English to do the job, and asylum seekers and refugees must not be forced to view English as yet another barrier to cross or hoop to jump through.

4. “English as a Barrier to Employment, Education and Training”

Dr Philida Schellekens, The Schellekens Consultancy

The final conference presentation was delivered by Philida Schellekens, whose focus was a report which she completed for the then Department for Education and Employment in 2001. In it she detailed the findings of her research on why non-native speakers of English in the UK encounter so many difficulties in, specifically, finding employment and, more generally, participating in English-speaking community life.

The desk research and subsequent field research related to five areas in England and Wales, and one of the first findings was that there is an alarming lack of firm data regarding the size of the ESOL community, not only in these areas but also throughout the UK. Nationally, it is estimated that at least 3 million residents were

born in countries where English is not the national language, and that up to 1.5 million people cannot function in work or social contexts because of a lack of English language skills. However, Philida Schellekens underlined that judging the success of action taken to address problems of this nature would always be difficult whilst the precise extent of the problems is allowed to remain uncertain.

In the sample areas unemployment rates were found to be as high as 90% for some asylum seekers and refugees, but even those in work often had to face under-employment (engaged in jobs for which they were over-qualified and/or over-experienced). Not surprisingly, levels of poverty were up to four times worse amongst non-native speakers than amongst the native English population.

Given these conditions, Philida Schellekens reported that the main motivation for her respondents to learn English was to secure employment, and that inadequate competence in English was their major barrier to the labour market. Interviews with employers revealed the generally held view that all but the most menial jobs require fluent communication skills, and that all employees need to be able to express themselves well in English when dealing with the public or with colleagues. It therefore makes good economic sense to develop ESOL provision in the UK, especially if the learners have work skills which are transferable to the UK environment.

Acquiring English language skills for the workplace, however, demands more intensive and systematic teaching, higher level courses and a more challenging pace than is sometimes found in ESOL classrooms. It is a matter of critical importance to devise a curriculum that engages the learner, meeting work and study requirements as well as personal and social needs, and it was especially for this reason that Philida Schellekens was pleased to welcome the launch of the Certificates in English Language Skills. Designed specifically for non-native speakers of English, offering a flexible form of assessment on a skill-by-skill basis and across different levels, the CELS, especially at the Higher level, is precisely the type of qualification which will help people make progress in work or study and capitalise on the potential which they possess.

LTTRC 2002 – Reminder

The 24th International Language Testing Research Colloquium (LTTRC 2002) will be held at the Hong Kong Polytechnic University December 12–15 2002, immediately before the 13th World Congress of Applied Linguistics (AILA) in Singapore

(16–21 December 2002). Further information is available from the following website: <http://www.engl.polyu.edu.hk/ACLAR/ltrc.htm>

Conference activities

The past few months have been especially busy ones for our staff and consultants in terms of attending and giving academic presentations at national and international conferences. The table

below gives an indication of the range of recent events where we have been able to profile our research and validation work, as well as the breadth of topics covered.

Event	Date	Title of presentation	Presenter
British Council Florence, Italy	February	Balancing continuity and innovation: the development of an EFL examination 1913-2002	Professor Cyril Weir
British Council Florence, Italy	February	Diplomato/laureato: certificato?	Liam Vint
British Council Florence, Italy	February	Teachers as researchers in the Cambridge PL2000 impact study	Dr Roger Hawkey and Grazia Maria Bertini
British Council Florence, Italy	February	Communicative approaches to language testing: past, present and future	Nick Saville and Dr Roger Hawkey
IATEFL, UK	March	From certificate to diploma: making the journey easier	Monica Poulter
IATEFL, UK	March	The Common European Framework and its implications for language assessment	Dr Peter Hargreaves
IATEFL, UK	March	Current perspectives on corpus-informed language testing from Cambridge	Fiona Ball
IATEFL, UK	March	Getting the most out of validation information	Dr Nic Underhill and Chris Banks
IATEFL, UK	March	International training needs in language teaching management	Elaine Brown, Ron White and Andrew Hockley
IATEFL, UK	March	Some assessment-related ideas for developing young learners' language skills	Melanie Williams and Rosalie Kerr
IATEFL, UK	March	Continuity and innovation: CPE 1913-2002	Professor Cyril Weir
IATEFL, UK	March	The revision of the Cambridge Business English Certificates	Dr Barry O'Sullivan and Hugh Bateman
IATEFL, UK	March	Impact studies as action research: the role of teachers	Dr Roger Hawkey
AAAL, US	April	Assessing learners' English for international education: but whose/which English(es)	Dr Lynda Taylor
AAAL, US	April	Plurilingualism and acquisition of partial competence: implications for assessment	Nick Saville
TESOL, US	April	Evaluating the academic writing ability of L2 learners	Dr Lynda Taylor
ALTE, St Petersburg, Russia	April	Balancing continuity and innovation: the development of an EFL examination 1913-2002	Professor Cyril Weir
ALTE, St Petersburg, Russia	April	Defining the test-taker: theoretical and practical perspectives	Dr Barry O'Sullivan
CPE Launch	May	Continuity and innovation: CPE 1913-2002	Professor Cyril Weir
CPE Launch	May	CPE Reading: constructing the right challenge	Hilary Maxwell-Hyslop
CPE Launch	May	Research studies for the revised CPE Speaking	Angela French
Ethics Conference, Pasadena, US	May	A code of practice and quality management system for international language examinations	Nick Saville, Henk Kuijper and Piet van Avermaet
British Council Berlin, Germany	May	The IELTS Question Paper Production Process	Juliet Wilson
METU, Turkey	May	What makes a good test?	Dr Nic Underhill
METU, Turkey	May	Recent developments in IELTS	Dr Nic Underhill
NAFSA, San Antonio, USA	May	IELTS – English for international opportunity	Nick Charge, Andy Williams, Anne-Marie Cooper, Beryl Meiron

The Cambridge ESOL staff seminar programme

Our EFL staff seminar programme has continued in recent months with contributions from both internal staff (e.g. Subject Officers, Development Managers and Research/Validation staff) and external specialists in various ELT fields.

In April Dr Jennifer Jenkins from the Department of Education and Professional Studies at King's College, London, was able to visit us and outline the proposals contained in her recently published book – *The Phonology of English as an International Language* (published by Oxford University Press in 2000). Dr Jenkins observed that mainstream ELT is steadily being transformed from the teaching of the foreign language (EFL) to the teaching of an international language (EIL) and she highlighted the implications for pronunciation teaching which arise as a result, suggesting that this development calls into question the type(s) of English which should serve as pedagogic models. She went on to describe some of the empirical evidence on which her proposal of a 'lingua franca (pronunciation) core' is based, outlined the areas which she regards as non-core, and briefly extended the discussion to parallel issues in the teaching of lexico-grammar.

Since the Cambridge EFL speaking tests routinely include pronunciation among the criteria for evaluation, this topic was especially relevant to our work; in a follow-up discussion with Dr Jenkins we were able to consider the implications that such

proposals might have for the way in which we assess L2 spoken language ability.

In May we welcomed Professor Michael McCarthy from the School of English Studies at the University of Nottingham, UK. In his presentation Professor McCarthy looked at the problems and prospects of investigating spoken language corpora, based upon the CUP/Nottingham University CANCODE conversational corpus. He explained how spoken corpora enable us to obtain unique insights into the special grammar and lexis of everyday conversation but also enable analysis of discourse features such as turn-boundary phenomena, topic management, use of pragmatically motivated items (e.g. discourse markers) and a range of other speech-specific phenomena. He also highlighted problems with spoken corpora, not least the fragmented appearance of speech, the difficulty of recontextualising highly contextualised talk, cultural problems and purely mechanical problems of transcription and interpretation.

As we move ahead this year with our own project to build a corpus of spoken learner language, based upon recordings of our face-to-face speaking tests, we look forward to developing our relationship with Professor McCarthy and to establishing collaborative links with him and his colleagues at Nottingham University.

Further Information

UCLES provides extensive information on the examinations and assessment services referred to in this newsletter. For further information, visit the website

www.cambridgeESOL.org

or contact

ESOL Information
University of Cambridge
ESOL Examinations
1 Hills Road
Cambridge, CB1 2EU
United Kingdom

Tel. +44 1223 553355

Fax. +44 1223 460278

e-mail esol@ucles.org.uk

For information on the ALTE five-level scale and the examinations which it covers, visit the ALTE website **www.alte.org**

or contact

The ALTE Secretariat
1 Hills Road
Cambridge CB1 2EU
United Kingdom

Tel: +44 1223 553925

Fax: +44 1223 553036

e-mail: alte@ucles.org.uk

If you would like further copies of this issue of Research Notes, or if you would like to be added to our mailing list (all registered centres receive copies automatically), please complete this form using block capitals and return it.

Please photocopy the form if you wish.

Please send me extra copies of this issue of Research Notes.

Please add me to the Research Notes mailing list.

Name

Job Title

Institution

Address

Country

Please let us have any comments about Research Notes, and what you would like to see in further issues.