European language testing
in a global context
Proceedings of the ALTE Barcelona Conference
July 2001
Also in this series:

An investigation into the comparability of two tests of English as a Foreign Language: The Cambridge-TOEFL comparability study
Lyle F. Bachman, F. Davidson, K. Ryan, I.-C. Choi

Test taker characteristics and performance: A structural modeling approach
Antony John Kunnan

Performance testing, cognition and assessment: Selected papers from the 15th Language Testing Research Colloquium, Cambridge and Arnhem
Michael Milanovic, Nick Saville

The development of IELTS: A study of the effect of background knowledge on reading comprehension
Caroline Margaret Clapham

Verbal protocol analysis in language testing research: A handbook
Alison Green

A multilingual glossary of language testing terms
Prepared by ALTE members

Dictionary of language testing
A. Brown, C. Elder, N. Iwashita, E. Grove, K. Hill, T. Lumley, T. McNamara

Learner strategy use and performance on language tests: A structural equation modelling approach
James Enos Purpura

Fairness and validation in language assessment: Selected papers from the 19th Language Testing research Colloquium, Orlando, Florida
Antony John Kunnan

Issues in computer-adaptive testing of reading proficiency
Micheline Chalhoub-Deville

Experimenting with uncertainty: Essays in honour of Alan Davies
A. Brown, C. Elder, N. Iwashita, E. Grove, K. Hill, T. Lumley, K. O’Loughlin, T. McNamara

An empirical investigation of the componentiality of L2 reading in English for academic purposes
Cyril Weir

The equivalence of direct and semi-direct speaking tests
Kieran O’Loughlin

A qualitative approach to the validation of oral language tests
Anne Lazaraton

Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913 – 2002
Edited by Cyril Weir and Michael Milanovic

Unpublished

The development of CELS: A modular approach to testing English Language Skills
Roger Hawkey

Testing the Spoken English of Young Norwegians: a study of testing validity and the role of ‘smallwords’ in contributing to pupils’ fluency
Angela Hasselgren

Changing language teaching through language testing: A washback study
Liying Cheng
European language testing
in a global context
Proceedings of the ALTE Barcelona Conference
July 2001
Contents

Series Editor’s note vii

Section One
Issues in Language Testing

1 The shape of things to come: will it be the normal distribution? 1
   Charles Alderson

2 Test fairness 27
   Antony Kunnan

Section Two
Research Studies

3 Qualitative research methods in language test development and validation 51
   Anne Lazaraton

4 European solutions to non-European problems 73
   Vivien Berry and Jo Lewkowicz

5 Validating questionnaires to examine personal factors in L2 test performance 93
   James E. Purpura

6 Legibility and the rating of second-language writing 117
   Annie Brown

7 Modelling factors affecting oral language test performance: a large-scale empirical study 129
   Barry O’Sullivan

8 Self-assessment in DIALANG. An account of test development 143
   Sari Luoma
Section Three
A European View

9 Council of Europe language policy and the promotion of plurilingualism
Joseph Sheils 157

10 Higher education and language policy in the European Union
Wolfgang Mackiewicz 173

Section Four
Work in Progress

11 TestDaF: Theoretical basis and empirical research
Rüdiger Grotjahn 189

12 A Progetto Lingue 2000 Impact Study, with special reference to language testing and certification
Roger Hawkey 205

13 Distance-learning Spanish courses: a follow-up and assessment system
Silvia María Olalde Vegas and Olga Juan Lazáro 221

14 Certification of knowledge of the Catalan language and examiner training
Mònica Pereña and Lluís Ràfols 237

15 CNaVT: A more functional approach. Principles and construction of a profile-related examination system
Piet van Avermaet, José Bakx, Frans van der Slik and Philippe Vangeneugden 249

16 Language tests in Basque
Nicholas Gardner 261

17 Measuring and evaluating competence in Italian as a foreign language
Giuliana Grego Bolli 271
Series Editor’s note

The conference papers presented in this volume represent a small subset of the many excellent presentations made at the ALTE conference – European Language Testing in a Global Context – held in July 2001 in Barcelona in celebration of the European Year of Languages – 2001. They have been selected to provide a flavour of the issues that the conference addressed. A full listing of all presentations is attached at the end of this note.

The volume is divided into three parts. The first, with two papers, one written by Charles Alderson and the other by Antony Kunnan, has a focus on more general issues in Language Testing.

Alderson looks at some key issues in the field; he considers “the shape of things to come” and asks if it will be the “normal distribution”. Using this pun to structure his paper, he focuses on two aspects of language testing; the first relates to the technical aspects of the subject (issues of validity, reliability, impact etc.), the second relates to ethical and political concerns.

Most of his paper chimes well with current thinking on the technical aspects and, as he admits, much of what he presents is not new and is uncontroversial. Within the European context he refers to the influential work of the Council of Europe, especially the Common European Framework and the European Language Portfolio; he describes a number of other European projects, such as DIALANG and the national examination reform project in Hungary, and he praises various aspects of the work of ALTE (e.g. for its Code of Practice, for organising useful conferences, for encouraging exchange of expertise among its members, and for raising the profile of language testing in Europe).

In focusing on the political dimension, however, he positions himself as devil’s advocate and sets out to be provocative – perhaps deliberately introducing a “negative skew” into his discussion. As always his contribution is stimulating and his conclusions are certainly controversial, particularly his criticism of ALTE and several other organisations. These conclusions would not go unchallenged by many ALTE members, not least because he misrepresents the nature of the association and how it operates.

Kunnan’s paper discusses the qualities of test fairness and reflects his longstanding concerns with the issues involved in this area. The framework he presents is of great value to the field of Language Testing and Kunnan has contributed significantly to the on-going debate on the qualities of test fairness within ALTE.

The second part of the volume presents a number of research studies. Anne
Lazaraton focuses on the use of qualitative research methods in the development and validation of language tests. Lazaraton is a pioneer of qualitative research in language testing and her involvement dates back to the late eighties and early nineties when such approaches were not yet widely used in the field. It is in part due to her efforts that researchers are now more willing to embrace approaches that can provide access to the rich and deep data of qualitative research. Readers are encouraged to look at her volume in this series (*A Qualitative Approach to the Validation of Oral Language Tests*).

Vivien Berry and Jo Lewkowicz focus on the important issue of compulsory language assessment for graduating students in Hong Kong. Their paper considers alternatives to using a language test alone for this purpose and looks at the applicability of variations on the portfolio concept. Jim Purpura’s work on the validation of questionnaires, which addresses the interaction of personal factors and second language test performance, represents an interesting and challenging dimension of validation in language testing. Readers may also wish to refer to Purpura’s volume in this series (*Learner strategy use and performance on language tests: A structural equation modelling approach*), which looks in more depth at the development of questionnaires to determine personal factors and a methodology that can be used to investigate their interactions with test performance.

Annie Brown’s paper is particularly relevant as we move towards greater use of computers in language testing. Such a move is of course fraught with issues, not least of which is the one of legibility that Brown addresses here. Her findings are interesting, giving us pause for thought and indicating, as she suggests, that more research is required. In the context of IELTS, such research is currently being conducted in Cambridge.

Barry O’Sullivan’s paper attempts to model the factors affecting oral test performance, an area of particular significance in large-scale assessment. The paper is part of on-going research commissioned by the University of Cambridge Local Examinations Syndicate and it is hoped that a collection of research studies into the dimensions of oral assessment will be published in this series in due course.

Finally, Sari Luoma’s paper looks at self-assessment in the context of DIALANG. The DIALANG project, also referred to in Alderson’s paper, has been one of the key initiatives of the European Commission in relation to language testing. As such it has benefited from significant funding and generated much research potential.

The last two parts of the volume cover aspects of work in progress. On the one hand, Joe Shiels and Wolfgang Mackiewicz summarise aspects of the on-going work of the Council of Europe and the European Union in relation to language policy. On the other, a number of researchers bring us up-to-date with test development work largely, though not exclusively, in the context of ALTE. These papers provide the reader with a reasonable overview of what is going on in a number of European countries.
In the context of the conference reflected in this volume, it is appropriate to overview how ALTE has developed over the years and what is of particular concern to the members of ALTE at the moment.

ALTE has been operating for nearly a decade and a half. It was first formed when a few organisations, acknowledging the fact that there was no obvious forum for the discussion of issues in the assessment of one’s own language as a foreign language in the European context, decided to meet with this aim in mind. The question of language assessment generally is an enormous one and dealt with in different ways by national and regional authorities throughout Europe and the world. Trying to bring together such a large and diverse community would have been a very significant task and far beyond the scope of ALTE’s mission. ALTE’s direct interests and aims are on a much smaller scale and it important to underline that it seeks to bring together those interested in the assessment of their own language as a foreign language. This is often in an international context, particularly with the more widely spoken languages but also in a national context, as is the case with lesser spoken languages in particular. While some ALTE members are located within ministries or government departments, others are within universities and cultural agencies. The members of ALTE are part of the international educational context and ALTE itself, as well as the members that form it, is a not-for-profit organisation. As a group, ALTE aims to provide a benchmark of quality in the particular domain in which it operates. Should ALTE’s work be of relevance outside its own context, then so much the better, but ALTE does not set out to establish or police the standard for European language assessment in general.

The recent history of language testing in the European context is very mixed. In the case of English we are fortunate that there has been significant interest and research in this field in English speaking countries for many years. In relation to some other European languages this is not the case. ALTE recognises that the field of language testing in different languages will be at different stages of development and that developing a language testing capacity in the European context, albeit in a relatively narrow domain, is an on-going venture. Similarly, progress, in contexts where participants are free to walk away at any time, cannot be achieved through force or coercion but rather through involvement, greater understanding and personal commitment. ALTE operates as a capacity builder in the European context, albeit in a relatively narrow domain.

As with any association, ALTE has a Secretariat, based in Cambridge and elected by the membership. The Secretariat has a three-year term of office and is supported by a number of committees, made up from the membership, who oversee various aspects of ALTE’s work. The group is too large for all members to be involved in everything and there are a number of sub-groups, organised by the members and focusing on particular areas of interest. The sub-groups are formed, reformed and disbanded as circumstances and
interests dictate, and at the moment there are several active ones. We will briefly describe the work of some of these here.

The whole of ALTE has been working for some time on the ALTE Framework which seeks to place the examinations of ALTE members onto a common framework, related closely through empirical study, to the Common European Framework. The process of placing examinations on the framework is underpinned by extensive work on the content analysis of examinations, guidelines for the quality production of examinations and empirically validated performance indicators in many European languages. This work has been supported by grants from the European Commission for many years and is now being taken forward by a number of sub-groups which are considering different domains of use such as language specifically for work purposes, for young learners or for study through the medium of a language.

A group has been established to look at the extent to which teacher qualifications in different languages can be harmonised and placed on some kind of framework. The group is not looking specifically at state organised qualifications but rather those common in the private sector for example, those offered by The Alliance Francaise, the Goethe Institute, the Cervantes Institute or Cambridge amongst others. It seeks to provide greater flexibility and mobility for the ever growing body of language teachers often qualified in one language and wishing to teach another while having their existing qualifications recognised as contributing to future ones in a more systematic way than is possible at present.

The Council of Europe has made and continues to make a substantial contribution to the teaching, learning and assessment of languages in the European context and in recent years has developed the concept of the European Language Portfolio as an aid and support to the language learning and teaching community. ALTE and the European Association for Quality Language Services have collaborated on the development of a portfolio for adults, which is now in the public domain. It is hoped that this will be a valuable aid to adult learners of languages in the European context.

An ALTE sub-group has been working with the Council of Europe and John Trim in the elaboration of a Breakthrough level which would complement the Waystage, Threshold and Vantage levels already developed. ALTE’s work in this area has also been supported by the European Commission in the form of funding to a group of members from Finland, Ireland, Norway, Greece and Sweden who have a particular interest in language teaching and testing at the Breakthrough level.

Another ALTE sub-group has been working on the development of a multilingual system of computer-based assessment. The approach, which is based on the concept of computer adaptive testing, has proved highly successful and innovative, providing assessment in several European languages and recently won the European Academic Software award in 2000.
ALTE members have developed a multilingual glossary of language testing terms. Part of this work has been published in this series (A multilingual glossary of language testing terms) but is ongoing, and as new languages join ALTE, further versions of the glossary are being developed. The glossary has allowed language testers in about 20 countries to define language testing terms in their own language and thus contributes to the process of establishing language testing as a discipline in its own right. The European Commission has supported this work throughout.

In the early 1990s, ALTE developed a code of professional practice and work has continued to elaborate the concept of quality assurance in language testing through the development of quality assurance and quality management instruments for use initially by ALTE members. This work has been in progress for several years and is now in the hands of an ALTE sub-group. As noted above, developing the concept of quality assurance and its management has to be a collaborative venture between partners and is not prone to imposition in the ALTE context. ALTE members are aware that they carry significant responsibility and aim to continue to play a leading role in defining the dimensions of quality and how an effective approach to quality management can be implemented. This work is documented and has been elaborated in ALTE News as well as at a number of international conferences. Details are also available on the ALTE website: www.alte.org.

Members of ALTE are also concerned to measure the impact of their examinations and work has gone on in the context of ALTE to develop a range of instrumentation to measure impact on stakeholders in the test taking and using constituency. Roger Hawkey discusses the concept of impact in the context of the Lingua 2000 project in one of the papers in this volume – see contents page.

ALTE members meet twice a year and hold a language testing conference in each meeting location. This is an open event, details of which are available on the ALTE website. New ALTE members are elected by the membership as a whole. Members are either full – from countries in the European Union – or associate – from countries outside. For organisations which do not have the resources to be full or associate members or who operate in a related field, there is the option of observer status. Information on all of these categories of membership is available on the ALTE website.

Finally, following the success of the Barcelona conference ALTE has agreed to organise another international conference in 2005. Details are available on the website.

Mike Milanovic
Cyril Weir
March 03
Presentations at ALTE Conference Barcelona, 2001

Karine Akerman Sarkisian, Camilla Bengtsson and Monica Langerth-Zetterman
Uppsala University, Department of Linguistics, Sweden
Developing and evaluating web-based diagnostic testing in university language education

J. Charles Alderson
Lancaster University, Department of Linguistics and Modern English Language
The shape of things to come: Will it be the normal distribution?

José Bakx and Piet Van Avermaet
Katholieke Universiteit Nijmegen and Leuven, The Netherlands and Belgium
Certificate of Dutch as a Foreign Language: A more functional approach

David Barnwell
ITÉ
Using the Common European framework in an LSP setting: Rating the Italian, Spanish, French and German abilities of workers in the teleservices industry

Marsha Bensoussan and Bonnie Ben-Israel
University of Haifa, Department of Foreign Languages, Israel
Evaluating reading comprehension of academic texts: Guided multiple-choice summary completion

Aukje Bergsma
Citogroep, The Netherlands
NT2-CAT: Computer Adaptive Test for Dutch as a second language

Vivien Berry and Jo Lewkowicz
The University of Hong Kong, The English Centre, Hong Kong
European solutions to non-European problems

Geoff Brindley
Macquarie University, NCELTR, Australia
Investigating outcomes-based assessment

Annie Brown
The University of Melbourne, Language Testing Research Centre,
The impact of handwriting on the scoring of essays

Annie Brown (Co-authors: Noriko Iwashita and Tim McNamara)
The University of Melbourne, Language Testing Research Centre,
Investigating rater’s orientations in specific-purpose task-based oral

Peter Brown and Marianne Hirtzel
EAQUALS, ALTE
The EAQUALS/ALTE European Language Portfolio

Jill Burstein and Claudia Leacock
ETS Technologies, USA
Applications in Automated Essay Scoring and Feedback

Ross Charnock
Université Paris 9 – Dauphine
Taking tests and playing scales – remarks on integrated tests

David Coniam
The Chinese University of Hong Kong, Faculty of Education, Hong Kong
Establishing minimum-language-standard benchmark tests for English language teachers in Hong Kong
Margaretha Corell and Thomas Wrigstad
Stockholm University Department of
Scandinavian Languages and Centre for
Research on Bilingualism.
What’s the difference? Analysis of two paired
conversations in the Oral examination of the
National Tests of Swedish as a
Second/Foreign Language

Benó Csapó and Marianne Nikolov
University of Szeged and University of Pécs,
Hungary
Hungarian students’ performances on English
and German tests

John H.A.L. de Jong
Language Testing Services, The
Netherlands
Procedures for Relating Test Scores to
Council of Europe Framework

Clara Maria de Vega Santos
Universidad de Salamanca, Spain
Otra manera de aproximarse a la Evaluación:
La Asociación Europea de Examinadores de
Lenguas (presentation in Spanish)

Veerle Depauw and Sara Gysen
Catholic University of Leuven, Centre for
Language and Migration, Belgium
Measuring Dutch language proficiency. A
computer-based test for low-skilled adults and
an evaluation system for primary school pupils

Urszula Dobesz
Uniwersytet Wrocławski, Poland
To know, to understand, to love (Knowledge
about Poland in the teaching of Polish as a
second language)

Ana Maria Ducasse
LaTrobe University, Spanish Department,
Australia
Assessing paired oral interaction in an oral
proficiency test

Ina Ferbčezar and Marko Stabej
University of Ljubljana, Centre for Slovene
as a Second/Foreign Language and
Department of Slavic Languages and
Literature, Slovenia
Developing and Implementing Language Tests
in Slovenia

Jésus Fernández and Clara Maria de Vega Santos
Universidad de Salamanca, Spain
Advantages and disadvantages of the Vantage
Level: the Spanish version (presentation in
Spanish)

Neus Figueras
Generalitat de Catalunya, Department
d’Ensenyament, Spain
Bringing together teaching and testing for
certification. The experience at the Escoles
Oficials d’Idiomes

Nicholas Gardner
Basque Government, Basque country
Test of Basque (EGA)

Małgorzata Gązynska
Jagiellonian University, Poland
Testing Comprehension in Polish as a Foreign
Language (paper in Spanish)

April Ginther (Co-author: Krishna Prasad)
Purdue University, English Department, Biloa
University, and Purdue University, USA
Characteristics of European TOEFL
Examinees

Giuliana Grego Bolli
Università per Stranieri di Perugia, Italy
Measuring and evaluating the competence of
the Italian as a second language

Rüdiger Grotjahn
Ruhr-Universitaet Bochum, Germany
TestDAF: Theoretical basis and empirical
research
Anne Gutch
UCLES, UK
A major international exam: The revised CPE

H.I. Hacquebord and S.J. Andringa
University of Groningen, Applied Linguistics, The Netherlands
Testing text comprehension electronically

Roger Hawkey
c/o UCLES, UK
Progetto Lingue 2000: Impact for Language Friendly Schools

Nathalie Hirschprung
Alliance Française, France
Teacher certifications produced by the Alliance Française in the ALTE context

Maria Iakovou
University of Athens, School of Philosophy, Greece
The teaching of Greek as a foreign language: Reality and perspectives

Miroslaw Jelonkiewicz
Warsaw University and University of Wroclaw, Poland
Describing and Testing Competence in Polish Culture

Miroslaw Jelonkiewicz
Warsaw University, Poland
Describing and gauging competence in Polish culture

Neil Jones
UCLES, UK
Using ALTE Can-Do statements to equate computer-based tests across languages

Neil Jones, Henk Kuijper and Angela Verschoor
UCLES, Citogroep, UK, The Netherlands
Relationships between paper and pencil tests and computer based testing

Sue Kanburogh Hackett and Jim Ferguson
The Advisory Council for English Language Schools Ltd, Ireland
Interaction in context: a framework for assessing learner competence in action

Lucy Katona
Idegennyelvi Továbbképző Központ (ITK), Hungary
The development of a communicative oral rating scale in Hungary

Antony John Kunnan
Assoc. Prof., TESOL Program, USA
Articulating a fairness model

Rita Kursite
Jaunjelgava Secondary School, Latvia
Analyses of Listening Tasks from Different Points of View

Michel Laurier and Denise Lussier
University of Montreal – Faculty of Education and McGill University
The development of French language tests based on national benchmarks

Anne Lazaraton
University of Minnesota, USA
Setting standards for qualitative research in language testing

Jo Lewkowicz
The University of Hong Kong, The English Centre, Hong Kong
Stakeholder perceptions of the text in reading comprehension tests

Sari Luoma
University of Jyväskylä
Self-assessment in DIALANG

Denise Lussier
McGill University, Canada
Conceptual Framework in Teaching and Assessing Cultural Competence

© Cambridge University Press www.cambridge.org
Elana Shohamy  
**Tel Aviv University, School of Education, Israel**  
The role of language testing policies in promoting or rejecting diversity in multilingual/multicultural societies

Kari Smith  
**Oranim Academic College of Education, Israel**  
Quality assessment of Quality Learning: The digital portfolio in elementary school

M. Dolors Solé Vilanova  
**Generalitat de Catalunya, Centre de Recursos de Llengües Estrangeres of the Department of Education**  
The effectiveness of the teaching of English in the Baccalaureate school population in Catalonia. Where do we stand? Where do we want to be?

Bernard Spolsky  
**Bar Ilan University, Israel**  
Developing cross-culturally appropriate language proficiency tests for schools

Claude Springer  
**Université Marc Bloch, Strasbourg**  
A pragmatic approach to evaluating language competence: Two new certifications in France

Marko Stabej and Nataša Pirih Svetina  
**University of Ljubljana, Department of Slavic Languages and Literatures, Slovenia**  
The development of communicative language ability in Slovene as a second language

Sauli Takala, Felianka Kaftandjieva and Hanna Immonen  
**University of Jyväskylä, Centre for Applied Language Studies, Finland**  
Development and Validation of Finnish Scales of Language Proficiency

Lynda Taylor  
**UCLES, UK**  
Revising instruments for rating speaking: combining qualitative and quantitative insights

John Trim  
**Project Director for Modern Languages, Council of Europe**  
The Common European Framework of Reference for Languages and its implications for language testing

Philippe Vangeneugden and Frans van der Slik  
**Katholieke Universiteit Nijmegen, The Netherlands and Katholieke Universiteit Leuven, Belgium**  
Towards a profile related certification structure for Dutch as a foreign language. Implications of a needs analysis for profile selection and description

Juliet Wilson  
**UCLES, UK**  
Assessing Young Learners – What makes a good test?

Minjie Xing and Jing Hui Wang  
**Salford University, School of Languages, UK and Harbin Institute of Technology, Foreign Language Department, China**  
A New Method of Assessing Students’ Language and Culture Learning Attitudes
Poster Presentations

Guy Bentner & Ines Quaring
Centre de Langues de Luxembourg
Tests of Luxembourgish as a foreign language

Else Lindberg & Peter Vilads Vedel
Undervisningsministeriet, Copenhagen, Denmark
The new test of Danish as a foreign language

Reidun Oanaes Andersen
Norsk Språktest, Universitetet i Bergen, Norway
Tests of Norwegian as a foreign language

José Pascoal
Universidade de Lisboa, Portugal
Tests of Portuguese as a foreign language

Heinrich Rübeling
WBT, Germany
Test Arbeitsplatz Deutsch – a workplace related language test in German as a foreign language

László Szabo
Eotvos Lorand University, Budapest, Centre for Foreign languages, Hungary
Hungarian as a foreign language examination
List of contributors

J. Charles Alderson
Department of Linguistics and Modern English Language, Lancaster University

Antony Kunnan
California State University

Anne Lazaraton
University of Minnesota, Minneapolis, USA

Vivien Berry and Jo Lewkowicz
University of Hong Kong

James E. Purpura
Teachers College, Columbia University

Annie Brown
University of Melbourne

Barry O’Sullivan
University of Reading

Sari Luoma
Centre for Applied Language Studies, University of Jyväskylä, Finland

Joseph Sheils
Modern Languages Division/Division des langues vivantes, DGIV Council of Europe/Conseil de l’Europe, Strasbourg

Wolfgang Mackiewicz
Conseil Européen pour les Langues/European Language Council

Rüdiger Grotjahn
Ruhr-Universität Bochum, Germany

Roger Hawkey
Educational Consultant, UK

Silvia María Olalde Vegas and Olga Juan Lázaro
Instituto Cervantes

Mònica Pereña and Lluís Ràfols
Direcció General de Política Lingüística, Departament de Cultura, Generalitat de Catalunya

Piet van Avermaet (KU Leuven), José Bakx (KUN), Frans van der Slik (KUN) and Philippe Vangeneugden (KU Leuven)

Nicholas Gardner
Department of Culture/Kultura Saila, Government of the Basque Country/Eusko Jaurlaritza, Spain

Giuliana Grego Bolli
Università per Stranieri di Perugia, Italy
Section 1

Issues in Language Testing