

# Conclusion? The Way Forward

Charles Alderson, Nagy Edit and Öveges Enikő

### A Summary

We hope that it is apparent to the reader how much has been achieved in this Project to date, and how useful the results of the first round of piloting have proved. We have not only been able to draw conclusions about the value of the items and tasks themselves: their empirical difficulty, problems in design, in formulation, and in administration. That is, after all, the normal purpose of piloting.

Because of the design of the piloting, we were able to develop and apply a methodology for calibrating the items onto a common scale, which enabled us to compare the difficulties of items and tasks across different papers taken by different groups of students. We were also able thereby to compare the difficulty of our tasks with those of earlier projects. As a result, the Project team is now able routinely to calibrate future pilot tasks with anchor items that have already been tested and calibrated. This will lead to the creation of a bank of calibrated items: items whose difficulty is known in advance and which can therefore be used for future administrations of the live examination, as well as for the calibration of further items.

In addition, we have developed and applied a methodology for interpreting the empirical difficulty of items, and for deciding what is meant by the terms Basic, Intermediate and Advanced, or indeed for any other scale of 'standards' that might be developed or applied. It will be important in the future that Hungary can describe the levels of its examinations, and thus the achievements of its school-leaving population, in terms of standards that other countries can understand. That in part is the function of the Common European Framework of reference of the Council of Europe, and this is why it is crucial that the Hungarian educational authorities reach decisions about the meaning of its supposed levels of achievements. We believe that the work of this Project will contribute to a debate on this issue.

The work of the Project as reported in this volume has also resulted in a significant volume of information about the achievements of Hungarian learners of English, and thus greatly augments the information available in the Baseline Study. In particular, there is now available a wealth of detail on how Hungarian learners perform on particular tests – be they of listening or reading, or whatever – and how particular tasks and items worked with this population. This is not only useful to item writers, who of course need empirical feedback on how well their items worked. It is also of immense importance for teachers themselves, for their everyday classroom work, for textbook writers, and for all those who wish to improve the level of English of the Hungarian population.

The information thus gained on the achievement of this population is also extremely valuable for the light it can throw on debates in language education that have to date been devoid of appeal to systematically collected evidence. The little evidence that is used in debates is frequently anecdotal and cannot be verified. Now we have the beginnings of a bank of data on performance of particular groups of the school population that can begin to inform not only the debates, but hopefully the decisions that are made by the educational authorities who are, after all, elected to power in part in order to improve Hungarian education. The issue of how many hours a week students

need to study a foreign language is one such issue, but there are others, like the number of years needed to achieve a certain standard, the standards achieved by different types of school, by different regions, and the quality of language teaching more generally. All these issues can be explored as part of the piloting of future valid and reliable tests of English.

One thing is already clear from the first large-scale piloting of new examinations, as documented in these pages, and that is the enormous heterogeneity of achievement in the school population. This is of course normal in any country, but Hungarian decision-makers need to consider this fact very carefully when deciding what 'levels' the school-leaving examinations should have. The Baseline Study showed very clearly that the current English examination is worthless, whatever its ceremonial value. For examination reform in English to have any impact, the new examinations must have value: they must not only be valid and reliable, they must be believed to certify valuable levels of achievement. At the same time, we firmly believe from the results of this piloting that school-leaving examinations must be capable of acknowledging the achievement of all school-leavers, however minimal that might be. It is absolutely clear that any examination that is 'failed' by a large proportion of the school-leaving population will have catastrophic impact on students, teachers and schools, and that must not be allowed to happen. If, however, the authorities do not recognise the need to address the issue of levels, then by default the notion of 'Intermediate' as relating to Rigó utca exams will prevail for better or worse – and we believe that this will cause considerable damage. Not only is there little believable evidence as to what Rigó utca levels really mean at present, the failure rate of these exams is high (even for those relatively few students who take the exams) and cannot and must not be paralleled in school-leaving exams.

We thus strongly endorse the recommendations of the Baseline Study that school-leaving examinations must be multi-level, that they cannot be confined to a simplistic two-level system, and we suspect that this is true for other foreign languages as well. School examinations must be capable of covering the range of achievement in the school population, and of certifying that achievement. We believe that this necessarily involves abandoning the traditional 5-point scale, at least for English, as has been done successfully in other Central European countries, and replacing it with a wider scale, probably 1 to 100, on which scores will be reported, and which users – employers, universities and others – can decide for themselves what result they consider adequate for their own purposes. 'Fitness for purpose' should replace the notion of 'pass' or 'fail' in the interpretation of exam results.

In order to reinforce the point, we repeat here the Recommendations made at the end of the Baseline Study, which this volume clearly endorses:

### **Recommendations for the Examination Reform**

- A central body should be responsible for designing and piloting tasks, administering and marking exams at all three levels, feeding results back to the developmental process, training of examiners and publicity
- A publicity campaign is needed to raise students', teachers' and other stakeholders' awareness of the advantages of the new exams
- The media should be exploited to keep students, teachers and parents updated on progress on the development of new examinations to prevent and reduce resistance to change
- A coordinated, standardised three-level exam should be established
- Levels of the new exams should be in harmony with those of the Council of Europe

- The new examination should record the level of achievement of all students in a valid and reliable way
- Careful level setting is necessary to pitch the three levels realistically, so that results motivate and do not frustrate students and other stakeholders
- We recommend that students' achievements should be indicated along a 1 to 100 point scale to record the wide range of performances at three levels, thus providing detailed information in addition to traditional marks
- Students' achievements should be separately indicated in the four skills and in the fifth component at the Advanced level
- The four skills should be equally weighted
- Students' performances should be carefully and systematically documented, and results sent back to schools and stored for analysis and research purposes
- Criterion-based exams should replace norm-referenced examinations
- External marking should replace internal marking as much as possible
- Examiner training should form an integral part of the Examination Reform
- Students not leaving state education after Year 10 should be encouraged to take the Basic Exam in modern languages so that they get feedback on their performance and develop test-taking skills
- The Basic Exam should be available as an option in a first and second foreign language for Year 11 and 12 students
- We suggest that better students should be encouraged to attempt the exams at two levels at the same time (Basic and Intermediate, or Intermediate and Advanced) and they should get feedback on their achievements on the 1 to 100 scale, but only their achieved level should be documented in their school reports
- The traditional timescale of school-leaving exams has to be rescheduled as externally marked exams take longer to process, similarly to current entrance exams
- Statistics on students' performances on exams should be made public
- Guidelines for task writers, interlocutors and markers should become public
- Exam preparation materials and practice tests should be published soon.

## **Professionalisation**

At a more mundane level, we hope that this volume has documented the range of different procedures that have been established as part of the process of making examination development in Hungary truly professional. We are not claiming that our procedures are perfect – far from it. But they now exist and are publicly documented, so that they can be followed by those who come after us. They can also be examined, criticised, reviewed and improved. That is part of the purpose of this document. We believe that examination reform should not proceed in secret, that documents should be openly available to the public, for the sake of informing and educating them as to what is done and how it is done. Testing all too often suffers from a reputation of being either too technical for the average teacher to understand, or something that is done to teachers and students, something to suffer, and perhaps to resist. In this Project we have followed a different philosophy: we have sought to involve school teachers – those responsible for the education of children and adolescents in real schools. The majority of Project members are practising secondary school teachers, and we have greatly benefitted from their experience and expertise, as well as hopefully having contributed to it. A considerable number of people – teachers – have now been trained in testing through this Project, and they understand fully the need for the professionalisation of assessment

as well as of teaching. Appendix II records the opinions of Team members, including practising secondary teachers, about the benefits of their involvement in this Project. It is impressive testimony.

In order to emphasise yet again the importance of involving and convincing teachers, we repeat here the Recommendations made in the Baseline Study with respect to teachers:

### **Recommendations for teacher education**

- Teachers have to be involved in the Examination Reform and their current favourable attitudes need to be strengthened and maintained
- Teachers' awareness must be raised towards the faults of the current school-leaving exam
- Courses on testing and exam preparation should be integrated into pre-service curricula
- Training of examiners should take place as part of mandatory in-service teacher development
- Methodology courses both in pre- and in-service teacher education should emphasise good practice, skills development, and efficient use of teaching materials and authentic resources
- All teachers must become familiar with new exam task types
- Techniques of developing the listening skills need special attention, as testing listening comprehension will be a new component in the school-leaving exam
- Teachers must be made aware of how classroom management in the target language develops students' listening skills
- Teachers should be encouraged to exploit input available outside the classroom
- All teachers should be aware how to motivate students and maintain their motivation in the long run
- Techniques of dealing with mixed-ability and mixed-level groups should be promoted
- Teachers must become more careful in choosing materials for teaching and exam preparation
- Classroom research techniques should be introduced to trigger reflection on teachers' own practice and to enhance cooperation among teachers.

### **And next?**

So what of the future? Where will the Project and examination reform more generally, go next?

As this conclusion is being written, at the end of April 2000, further pilot tests are underway, of Writing and Speaking tasks. New tasks have been written and reviewed, formative feedback has been given to item writers, who have revised their tasks and resubmitted them. The Editing Committee has scrutinised these tasks and selected a number of them for piloting. The procedures we have developed for the production of test tasks have been tested and appear to work smoothly. The tasks have also been submitted to the scrutiny of an outside body, the University of Cambridge Local Examinations Syndicate, for their professional opinion and the valuable feedback received has been taken into account wherever possible.

Performances of students on oral tasks will be recorded and analysed in the summer of 2000. The responses of students to the Writing tasks will be marked in the summer, after

professionally respectable standardisation procedures have been presented and implemented. The tasks, the students' performances, the marking scales presented in this volume, will all be subject to detailed scrutiny, they will likely be revised and lessons will hopefully be learned, much as has been reported in this volume.

New tasks for Listening, Reading and Use of English tests have been commissioned and will be reviewed and revised in the now normal way, with the intention of piloting them in a new round in the autumn of 2000. And so the work goes on, in a regular cycle of development, administration, analysis, interpretation, and learning. As a result, banks of calibrated items will be accumulated, experience and expertise will be developed, and Hungary will be able to count on an ever-growing body of professional language testers following professional procedures, developed according to international standards and practice.

And in parallel to this development, in-service courses for teachers will continue to be developed, training will be designed for those who will deliver these courses in the future, and the various needs of school teachers of English will continue to be analysed, and provision will be made for meeting these needs, as examination reform hopefully gathers momentum. Indeed, the third volume in this series of books on *English Language Education in Hungary*, this time documenting in detail the in-service course briefly described in Chapter 18, has already begun, and publication is planned for late 2000/early 2001.

But of course, the success, and the impact, of all these developments, depends upon the resolve of the educational authorities in Hungary to implement fully and professionally the planned examination reform. Traditionally, teachers here are cynical about the success and pace of reform and in the course of this Project we have often encountered teachers who are convinced that nothing will change, and that the problems of unreliable, invalid school leaving examinations will not be adequately addressed by policy makers or those charged with implementing decisions. And indeed the signs are that reform may be faltering. Already the planned Basic Examination for students at the end of Year 10 appears to have been scrapped with the raising of the school-leaving age to 18. For English all is not lost, since from the beginning we incorporated the development of Basic items and tasks into our work, despite the opposition and indifference of some. New Basic items are currently being reviewed and hopefully will eventually be piloted. Certainly they will exist for use in a future examination that incorporates multiple levels and does not restrict itself to the unrealistic and undefined 'Intermediate' and 'Advanced' levels.

Hungarian university authorities have to date taken little notice of the reform of school-leaving examinations, having been obsessed with coping with the bizarre new Accreditation Decree that seeks to open the 'market' for foreign language examinations. Although universities are autonomous, the time will soon come when they will need to consider what they require of school-leaving examinations in terms of levels and the reporting of results, if their own needs are to be taken into account by the examination reform, and if the negative impact of Rigó utca examinations on schools is to be removed, as is currently foreseen by legislation.

Above all, the Ministry of Education and the institutions charged with implementing examination reform will need to continue to take seriously their responsibility for ensuring that valid, reliable and useful examinations can be introduced. This was originally planned for the year 2004. Four years ago, that goal appeared highly realistic and it appeared eminently sensible to plan so far ahead. Time has passed, and a sense of urgency begins to rise. If schools are to be adequately prepared for new examinations in four years, decisions will need to be taken very soon about levels, formats, procedures, infrastructure, resources, training and quality control and assurance. Decisions, and resources, not just words. For students to have been adequately prepared in 2004, they would need maximum information by 2002, and teachers would need to be trained in

advance of that. Things are indeed becoming urgent. But the one decision that has just been taken is that the new examinations will now be introduced in the year 2005!

What is happening? Right now, it appears that it is still unclear which institutions will be responsible for the development and delivery of the new exams, and rumours abound. Organisational uncertainties seriously hinder work on the Project and we must emphasise the importance of clearly defining the authority of the institutions involved, and of clearly establishing which institution is responsible for what.

All, in short, is in flux. But 'twas ever thus, and those engaged in examination reform cannot expect anything less. Examinations lie at the heart of educational policy, for better or worse, simply because they drive the curriculum, whether you like it or not. Many, of course, do not like it, and ignore the role of examinations, preferring to concentrate on teacher education, school self-evaluation, or on academic research. But examinations will not go away, and they will continue to have impact. They should play an increasingly important role in quality assurance. It is crucial, to the health and development of language education in Hungary, that all professionals continue to strive to ensure that the importance of valid, reliable and meaningful school-leaving examinations is made clear to those with responsibility for educational decisions, and that the development and implementation of respectable, professional and externally monitored school-leaving examinations be assured.