THE UNIVERSITY OF
WARWICK

University of Warwick institutional repository

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Marilyn Hunt, Sean Neill and Ann Barnes
Article Title: The use of ICT in the assessment of modern languages: the English context and European viewpoints
Year of publication: 2007
Link to published version:
http://dx.doi.org/10.1080/00131910701255012
Publisher statement: None.

The use of ICT in the assessment of modern languages: the English context and European viewpoints

Marilyn Hunt, Sean Neill, Ann Barnes, University of Warwick

Marilyn Hunt is Lecturer in Modern Foreign Languages (Teacher Education) at the University of Warwick

Dr Sean Neill is Senior Lecturer in Research Methods

Dr Ann Barnes is Senior Lecturer in Modern Foreign Languages (Teacher Education) at the University of Warwick

Word count: 6217

Abstract: 144

Total word count: 6361

Address for correspondence:

Marilyn Hunt

Lecturer in Modern Foreign Languages (Teacher Education)

Institute of Education

University of Warwick

Coventry CV4 7AL

E-mail: m.j.hunt@warwick.ac.uk

Telephone : 02476 522120

The use of ICT in the assessment of modern languages: the English context and European viewpoints

Abstract

The ever increasing explosion of highly attractive multimedia resources on offer has boosted the use of ICT in the teaching and learning of modern languages. The use of ICT to assess languages is less frequent, however, although on-line testing is starting to develop. This paper examines the national context for the assessment of modern foreign language proficiency in England, outlines the kinds of assessment currently available and the development of electronic forms of assessment and compares the above with the survey results of an EU funded project on current good practice in on-line assessment of languages in other European countries. The findings indicate that speaking is inadequately served by on-line testing as tests currently focus primarily on receptive language skills. The implications for future successful on-line testing include the incorporation of interactive skills and effective formative feedback.

Key words: assessment, computers, languages, testing

Introduction

Assessment of course plays a vital role in successful learning and certainly remains the focus of a great deal of local and national attention, with debate around standards, external examinations, national tests and international comparisons. The apparently ever growing emphasis on evidence of learners' achievements in tests and examinations demonstrates that assessment is seen outside as well as within the educational context as crucial in judging both learners' and teachers'

performance. The aims of this paper are:

- to set out the national context for the assessment of modern foreign language (MFL) proficiency in England

- to outline the kinds of assessment currently available and the development of electronic forms of assessment

- to compare the above with the survey results of an EU funded project on current good practice in other European countries.

Languages, ICT and assessment

Although 'computers have been used for language teaching since the 1960s' (Warschauer and Healey, 1998:57), widespread use of Information and Communications Technology (ICT) in language teaching and learning has only really occurred over the past decade with the development of multimedia resources and the Internet. Warschauer and Healey (1998) maintain that computer-assisted language learning (CALL) has mirrored the broad progression of educational theory as regards MFL learning from behaviourist to communicative to the post-communicative phase. With the greater emphasis currently placed on an e-learning culture, developments are now taking place to find ways of assessing learning electronically (and effectively).

Some of the principal examples of Computer Aided Assessment (CAA) in languages, as outlined in the ICT4LT website (Atkinson and Davies 2005) include:

- interactive exercises and tests completed on a computer

- use of computers to produce coursework

- on-screen marking of students' word-processed writing

- use of revision software

- using a spreadsheet or database to keep a record of students' marks

- use of email to send coursework and to receive marks and feedback

- use of Web pages to set tasks for students and to provide tutor support

- use of plagiarism detection software.

Atkinson and Davies, the authors of this regularly up-dated site, claim that the importance of CAA is intimately bound up with raising achievement since it can be argued that the role of ICT in raising achievement cannot be fully measured unless ICT is also used in the assessment process. CAA will often be used for formative assessment because it is excellent for giving immediate feedback in terms of "right", "wrong", "try again" without learners feeling this is judgemental.   Multimedia can provide opportunities for pupil creativity, and learner confidence is increased by its non-judgemental correction of errors, allowing for 'experimentation in a supportive environment' (Dugard and Hewer, 2003: 29), and providing instant gratification. However, there are limitations to this feedback as it is not necessarily adequately discriminatory or differentiated, nor does it set future targets or provide guidance on how to improve one's learning in order to achieve these goals. Atkinson and Davies cite the most common test types as:

- matching / Pelmanism

- multiple-choice

- Cloze

- sentence reordering / jumbled words

- text reconstruction

- true or false

- crosswords

- wordsearch

- text entry quiz.

There is a range of software available on the web which helps with the reinforcement of language learning and can be used as formative assessment by testing learners on vocabulary, phrases, gap-fill and grammar. Many schools have developed their own websites (for example, http://www.rgshw.com/index2.htm or http://www.ashcombe.surrey.sch.uk/Curriculum/modlang/index.htm with exercises of this kind and there is a range of simple authoring packages with ready made exercises or the facility to create one's own, for example Hot Potatoes, http://web.uvic.ca/hrd/hotpot or Quia, http://www.quia.com.

English Assessment frameworks

The National Curriculum (NC) for MFL which was implemented in September 1992 for Year 7 pupils (age 11), demonstrated an increased recognition of foreign language needs (at least for Europe) and a belief in the necessity of languages for all. As well as laying out some indication of methodology, the National Curriculum for England (DfEE/QCA, 1999) included for the first time national criteria for assessment for pupils studying a MFL at Key Stage 3. This forced teachers to think more clearly about assessment, especially as there is a statutory obligation to award a NC level at the end of Key Stage 3 (age 14), and since there are no national tests as in English, Maths and Science, pupils' work is teacher assessed. NC levels, which consist of 8 Level Descriptions plus one of exceptional performance for each of four Attainment Targets, (Listening and responding, Speaking, Reading and responding, and Writing), are used to assess pupils' work and to plan for progression in pupils' learning. The benefits of these nationally specified criteria for different levels are clear (despite continuing debate around the accuracy of the Level Descriptions and the relative 'distance' between Levels): they provide a common understanding of standards, expectations and pupils' achievements and they are used to achieve consistency in teacher assessment within and across schools in all regions.

If pupils study a MFL to age 16, the majority take the GCSE (General Certificate in Secondary Education) public examination. Equal weighting is placed on listening, speaking, reading and writing (25% marks for each skill) and students opt for Foundation or Higher Tier examinations. The specifications in all languages are based on four themes, with some variation in the topics within the themes to take account of cultural differences. Tasks have a clear communicative purpose but there is a premium on using appropriate structures and achieving a high degree of accuracy. The cultural and societal impact of ICT is acknowledged: e-mail messages feature in writing papers and authentic texts may come from a website, although the examinations themselves remain paper-based. The Certificate of Achievement, aimed at NC Levels 2/3, was developed as an alternative accreditation to cater for pupils who are considered unlikely to be successful in achieving a GCSE grade.

A new GCSE qualification in Applied French features on-screen external assessment and places greater emphasis on productive language skills. This qualification, currently being piloted for first examination in summer 2006, enables learners to develop language skills that they can apply to a specialist context (Business, Leisure and Tourism or Media and Communications).

There are 3 'pathways' to qualifications post 16: the academic (GCSE/A level), the applied school-based vocational (GNVQ) and the workplace vocational (NVQ). The majority of students follow academic courses leading to A level where candidates are required to demonstrate knowledge of contemporary society and linguistic competence; all tests are currently paper based. GNVQ and NVQ Language Units generally form part of a wider qualification, for example, Travel and Tourism/Business Studies and cannot be studied as a qualification in their own right. Accreditation for these qualifications is available in 20 languages and is portfolio based where students collect evidence of performance.

Overall, formal examinations in the English system place considerable emphasis on interaction, including spoken interaction, practical application of language in leisure and, especially for vocational qualifications, in the workplace. While this makes them more relevant and, hopefully, attractive to learners, it does emphasise aspects which are more difficult to incorporate in e-learning systems.

The Qualifications and Curriculum Authority (QCA) E-learning strategy

Over recent years assessment has evolved from traditional methods, mainly externally awarded grades, to include teacher assessed course work, portfolios and profiles of achievement. However, community expectations, national standards and accreditation all demand ever increasing accountability. Assessment frequently defines goals for learners and teachers alike. The English government is clearly driven by goals to raise standards, improve quality, widen participation, and give people the skills and confidence to be successful in the workplace. With this in mind, their vision is to use e-learning systems to enhance the value of assessment through data analysis for the teacher and interactive feedback for the learner.

The QCA (the body charged with overseeing and monitoring all awarding bodies in England) views on-line assessment as a means to overcome barriers such as time, location and cost through 'on demand' testing. The latter would enable learners who lack confidence to progress at their own pace and demonstrate knowledge specific to their interests or to a task in the workplace and thereby become more motivated.

The e-assessment vision

Although currently national MFL examinations remain paper-based apart from a few exceptions, the strategic objective for QCA, for Awarding Bodies and for learning providers is that by 2009:

- all new qualifications must include an option for on-screen assessment;

- all awarding bodies should be set up to accept and assess e-portfolios;

- all existing GCSEs, AS and A2 examinations should be available on-screen;

- the first on-demand assessments are starting to be introduced;

- and at least ten new qualifications, specifically designed for electronic delivery and assessment, should be developed, accredited and live.

(source: http://www.qca.org.uk/6877.html)

Indeed, the ICT test (http://www.qca.org.uk/15350.html) to assess pupils' ICT capability at the end of KS3 (age 14) in England is currently being piloted as an on-screen test to be taken within a four week window of dates. This test is to become statutory from 2008 with results reported in achievement and attainment tables. This on-screen test redesigns the way testing happens in schools and represents a major cultural change for most schools, being a new experience for both pupils and teachers and is likely to have an impact on the use of ICT in the assessment of other subjects in the UK curriculum.

To achieve their objective, QCA intend to ensure that e-assessments are fair, reliable and valid, to quality assure e-assessment, and to develop technical and assessment standards. Within this e-assessment vision QCA aims to promote the use of ICT in formative assessment for all sectors and in all publicly-funded materials and to support action research pilots to test appropriate use of formative assessment, and improve assessment techniques. QCA also recognises the need to align e-

assessment methods to the needs of pedagogy and each specific subject, its content and how it is taught. It is clearly important to explore subject-area standards and how online assessment could transform the way in which learners' attainment is examined.

The use of ICT in assessment offers the potential to increase efficiency, decrease the labour intensive aspect of assessment and administration, and streamline and safeguard data transfer processes.

The Languages Ladder

A new voluntary recognition system in England to complement existing national qualifications frameworks and the Common European Framework (CEF) was proposed in the Green Paper, *Languages for All: Languages for Life, A Strategy for England* (DfES, 2003). The rationale for this system is to give people credit for their language skills in listening, speaking, reading and writing at all levels of competence for all ages in a wide range of languages. A ladder of recognition with 6 stages from Beginners' level to degree level is currently being developed to endorse foreign language learning competence and will allow learners to progress and be assessed in one or more of the four skills in one or more languages. Each stage is assessed with a series of "can do" statements for each skill which can be used by learners to assess their own levels of language competence, to be then endorsed by the teacher/tutor or by external assessment at the interface between stages. Some of the assessments are being developed for future use as on-line and just-in-time testing. The first on-line materials will be computer-based versions of the comparable 'paper' test and the first round of specifically computer-based assessment will be available in 2006. The scheme will be available in 8 languages at the first three stages in 2005/6 and a suite of 15 additional languages is under development for the following year. Whilst it is easy to imagine that listening and reading could be tested on-line, the

testing of speaking and writing is considerably more problematic and would still require some form of human intervention.

The Project

The ON-LANG project is designed to produce an e-learning resource which can be used by teachers to formulate tests which are validly calibrated against the Common European Framework (CEF); at the moment many tests are matched to national or proprietary frameworks and are not easily comparable across countries. This drastically reduces their value in permitting mobility, for example to take up courses or jobs in other countries, because potential course providers or employers cannot easily assess the language ability of applicants. For example, in England, although national criteria have been equated to levels within the CEF, this framework is not commonly used as a means of describing pupils' achievement in schools in England. Members of the project (language teachers from secondary and higher education and teacher trainers from Denmark, England, Finland, Italy, Norway, Poland, Romania) agreed to report on e-learning resources for teaching and testing Modern Foreign Languages (MFL - i.e. teaching native speakers the language of another country) or at Teaching [a specified] Language as a Foreign Language (TLFL i.e. teaching the native language for educational or business purposes to speakers of another language) which were examples of good practice. Responses indicate what members of the project, as experts in language teaching, thought was educationally valuable, or, conversely, problematic. The aim of the survey was to identify what aspects of current practice were valuable and should be included in the ON-LANG resource, and what features were currently not available but should be included in the ON-LANG resource. When completed, the ON-LANG resource is being piloted with students and teachers, with feedback questionnaires to assess its success in meeting their needs. A future report will assess its success.

Language teaching raises particular problems for the application of ICT, as many aspects of language rely on social understanding (Levinson 1995) and the last fifty years have shown that there are great problems in trying to program computers, which continue to operate in the mechanistic way of the original Turing machines (Copeland 2004), to take account of the subtleties of language (e.g. Aitchison 1989, Pinker 1994). Clearly it is important to analyse the limitations of such a new assessment method. Modes of testing supported best by computers tend to depend on the very convergent (only one right answer) and discrete-item styles of gap fill and multiple choice, whereas the weight of professional opinion after three decades of communicative language teaching would prefer assessment to be more concerned with judging the capacity to engage in interaction, to handle discourse (texts and contexts beyond sentence level) and to produce speech and writing in a variety of contexts.

Methodology: ON-LANG survey of European resources

The ON-LANG survey data were collected by a web-based questionnaire, hosted on the ON-LANG website by Hochfeiler (the partner responsible for IT support to the project). As mentioned above, the respondents were the project members, who were experts in language teaching in their own countries. Question topics included:

- a description and overall rating of each resource

- resource coverage

- whether the resource provided formative and/or summative assessment

- the question format for assessment

- whether the resource was primarily theory-based or balanced with practice

- whether students could use the resource on their own, and if not, what level of support was needed

- whether the resource had been designed for class work, group work or individual work

11

- whether the resource allowed direct interaction between teacher and student, or interaction with a forum of users

- what receptive or productive facilities were needed: video, audio, photos, text or others

- whether the resource provided interactivity and/or room for student creativity

- whether additional / support material was available on the net.

Most of the questions were closed, but each section included an opportunity for open-ended comment. The responses represented the considered view of project members on what constitutes a valuable resource: the survey did *not* aim to provide a representative sample of ICT-based resources, good, bad and indifferent, across Europe. Its primary purpose was to identify existing good practice which could be incorporated into the ON-LANG.NET resource, or which should not be duplicated by ON-LANG.NET or to identify gaps in existing practice which ON-LANG.NET should fill.

Questionnaire outputs were collected as an Excel file, which was analysed by researchers at the University of Warwick. Quantitative aspects of the responses were analysed using SPSS version 12.01, with qualitative commentaries being transferred verbatim from the original records. Each open-ended question was accompanied by a closed question which asked how important the open-ended answer was, from 'very important' to 'not important at all'. Respondents were told in the guidance notes for completing the questionnaire that only qualitative comments marked 'very important' would be used in the analysis; the survey aimed to identify aspects which the project members considered critical to good practice in language teaching, and the qualitative comments were used to interpret the quantitative data. Respondents were free to make comments marked 'not important' but these were not used to

guide the analysis. Only a small number of questions were made compulsory, because some questions were likely to be irrelevant to particular resources, either because they lacked a specific feature or were not intended for a specific purpose. In general the text refers to the percentages of valid responses; for most questions these do not differ greatly from the percentages for the whole sample, but where a high proportion of respondents did not answer a question, this is highlighted.

E-learning resources covered by the ON-LANG survey

On-line language learning resources (not necessarily designed for assessment) were selected by professionals in the language field: teacher trainers and language teachers in secondary/vocational/language schools and higher education. Slightly over half the resources (59%) were aimed at MFL (Modern Foreign Languages - i.e. teaching native speakers the language of another country) usage – a slightly smaller proportion (50%) were aimed at TLFL (Teaching Language as a Foreign Language i.e. teaching the native language for educational or business purposes to speakers of another language) – of course it is possible for a resource to serve both purposes. Most countries reported on more MFL than TLFL resources – Italy and Romania (but not the UK), where the national languages were to be included in the project outputs, were exceptions and reported on more TLFL resources. Ratings for MFL and TLFL resources were compared statistically; there were few statistically significant differences. Those which did exist are discussed below, but where no difference is described, responses were generally closely similar for MFL and TLFL resources, reflecting the fact that the problems of language learning are not greatly affected by the context involved. However, MFL resources are more likely to be used in a class situation in a standard state school situation, whereas TLFL resources are more likely to be used with smaller groups – immigrant children in schools, or older learners who are voluntarily taking courses in a country to learn its language. This is reflected in the fact that statistical differences between MFL and TLFL courses are

concentrated in the area of teaching and learning usage; levels covered and resources used, for example, differed little between the two types.

The credibility of the survey obviously depends critically on how much experience respondents have of the resource on which they are reporting. For most resources, respondents had first-hand practical experience of using them (Table I); the remaining responses were fairly evenly divided between second-hand experience (i.e. respondents had discussed the value of the resource with colleagues who had used it in practice) and first-hand review (i.e. respondents had examined the resource, but had not had the opportunity to corroborate their professional judgment by using it in practice) rather than second-hand review. It should be pointed out that respondents are less likely to have had the opportunity for practical experience with the newest (and, hopefully, most educationally valuable) resources. Resources were most likely to be rated good (43%), with satisfactory (23%) the next most frequent rating; fewer were rated excellent (15%) with less than a fifth being rated fair (11%) or poor (8%) (Table II). One respondent noted that the technical quality of the sites investigated varied from the

> 'primitive to the highly professional, with an equally diverse and usually corresponding variability in user-friendliness; generally speaking, the html driven pages were less satisfactory than those using third-party test generating software, while those using Macromedia or similar had the best presentation and greatest user-friendliness.'
> Insert Tables I and II about here

Two-thirds of resources (69%) in the ON-LANG survey originated from commercial enterprises, with one-third originating from university language centres. Where commercial enterprises are concerned, the aim is to promote and sell the courses, and for language schools and university institutions, to persuade students to

14

subscribe to courses on-line. In these cases rather than testing to measure knowledge acquired, many different types of entry test evaluate linguistic ability by measuring grammatical and lexical skills and then suggest courses appropriate to different levels. The most common form of test consists of straightforward single or multiple choice questions. In slightly less than half the cases (44%) there was no indication on the web-pages of when they had been compiled, but slightly more than half (53%) showed who was responsible for updating the material. However only a quarter (27%) showed how users could send commentary or feedback. Thus the majority of resources were presented on a 'take it or leave it' basis, with little attempt to engage with their users.

Almost three-quarters of resources commented on were free, with the remainder fairly evenly spread over cost bands up to 500€ (Table III); over half the respondents (57%) who commented felt cost was a very important factor affecting whether schools would adopt a resource. In some cases, as will be apparent from the comments below, the freely available resource was 'a sprat to catch a mackerel' with the intent of inveigling users into the main, heavily charged resource.

Insert Table III about here

Most respondents who gave an opinion on whether the resource they had reported on was self-explanatory felt that resources were fairly clear in their explanations, with only a few resources being not at all or not very clear (Table IV). Almost all respondents thought that it was important that a resource was self-explanatory (Table V). For over a third of resources, additional or support material was fairly readily available on the Net; for a fifth of resources it was very available, but none at all was available for a quarter of resources (Table VI).

Insert Tables IV, V and VI about here

Most respondents considered that the availability of support material was very or fairly important (Table VII). In some on-line courses there are very rich resources for students' help-pages, and special interest pages for a deeper understanding and knowledge of the other country's culture and civilization.

<div align="center">Insert Table VII about here</div>

The resources covered by the ON-LANG survey covered resources for adult learners, including those who needed language skills for university or work purposes, as well as those for school use. Reflecting this, many were intended largely for individual autonomous learning: half the resources were considered very well suited to individual work, and most of the remainder fairly well suited (Table VIII). Four-fifths of respondents considered the design of the resource for individual work was very (44%) or fairly (38%) important. Over half the resources were not considered to be not at all (38%) or not very much (17%) designed for class work, but over half the respondents thought that design for class work was fairly (37%) or very (12%) important. Similar figures applied for group work. Respondents thought group work was significantly more important for MFL than for TLFL work. Since language is essentially social, this reflects the belief of language educators that it should be learnt in a social context. Nearly half the resources did not require the presence of a tutor at all, and only for slightly over a fifth was the presence of a tutor considered essential or necessary (Table IX). However two-thirds of the respondents thought the need for a tutor was very or fairly important, presumably because they thought students would not gain full benefit from the resource on their own (Table X). A tutor was thought significantly more necessary for MFL than TLFL work.

<div align="center">Insert Tables VIII, IX and X about here</div>

The four language skills

Reading

A majority of resources reported on by the ON-LANG survey provided a course on reading at a good (40%) or satisfactory (33%) level, which respondents considered fairly (47%) or very (39%) important. This reflects the large number of packages available, some with on-line support in terms of dictionaries, some which are specifically designed to help learners develop their reading skills or which are particularly suitable for reluctant readers (Slater and Varney-Burch, 2001). However, research indicates that on-screen reading is difficult with longer texts which require scrolling down and that 'reading from computer screens is about 25% slower than reading from paper' (Nielson, 1997) TLFL courses were more likely to be rated 'good' for reading and writing than MFL courses, which attracted a wider range of scores.

Writing

Respondents were less likely to consider a course on writing very (30%) or fairly (41%) important, and resources were less likely to be good (27%) or satisfactory (28%) in providing this. Given the technical difficulties involved in creating interactive testing of more extended writing, teacher assessment is more probable, although students might be asked to hand in their writing on-line.

Listening

Though most respondents thought listening very (44%) or fairly (45%) important, a quarter of courses (24%) did not provide this at all, though a third (34%) were good. Again this reflects the development of packages containing sound icons which when clicked enable the learner to hear the vocabulary, phrase or longer text; pause and replay are possible for listening activities.

Speaking

Over half the resources (57%) did not provide for oral production, but nearly two-thirds of respondents thought this was very (38%) or fairly (34%) important. Similar figures applied to oral interaction, where approaching two-thirds of courses (64%) did not provide this facility, despite it being thought fairly (38%) or very (34%) important by approaching two-thirds of respondents. The oral production aspect is often weak in the structure of on-line courses; the interactivity with tutor or teacher is often insufficient, there is a general lack of production activities and very frequently creativity is not stimulated. The technical difficulties of providing for oral interaction are of course largely responsible for this deficit. Although software has been developed which compares the learner's pronunciation with a pre-recorded model version - 'Performance can be compared with the model simply by listening or by looking at a graphical representation of the two recordings and comparing them' (Slater and Varney-Burch, 2001:29) - this software is not always entirely successful as experience has shown that even native speakers cannot always match the desired model! Furthermore it lacks any kind of interaction. 'Although some software boasts speech recognition it simply cannot meet the demands which communicative approaches make of it' (Bax, 2000: 215). As far as assessment of oral output is concerned, 'basically a human teacher is needed to evaluate free responses where they occur' (Warschauer and Healey, 1998:59), although digital video, e-mail with audio/video attachments or video-conferencing can be useful in recording this. Some developmental work has been done in creating software tools for effective delayed feedback and research indicates the potential of such multimedia technology in enabling the creation of tools which can significantly enhance delayed feedback (Tsutsui, 2004). However, as Atkinson and Davies (2005:1) state: 'To assess speaking is a very complex task and, as yet, computers are not fully adequate in either automatic speech recognition or natural language processing, but research in this area is developing rapidly'. It would appear that the problem lies in the nature of language itself and therefore in the complexity of the programming required.

Provision and quality of assessment

Important work has been published recently on the value and huge potential of assessment **for** learning (see Brooks 2002 and particularly Black and Wiliam 1998), as compared to assessment **of** learning, or summative assessment as demonstrated through specified national criteria for judging success. Following Wiliam and Black's research (2002) which demonstrated the importance of formative feedback, there has been an increased emphasis on formative assessment or 'assessment for learning' in order to give specific feedback to learners on their performance and targets for improvement. Barnes and Hunt (2003: 2) explain:

> 'In MFL, this includes using specific linguistic feedback and targets when marking pupils' work across all four Attainment Targets as well as self-assessment checklists, where pupils know what they are supposed to be learning and why, and then what they need to do to get there. Such assessment can achieve excellent results in improving pupils' learning, which then feeds through to the summative assessment of which a wider public is aware and by which pupils and teachers tend to be judged.'

 In the ON-LANG survey, although two-thirds of respondents thought formative assessment very (39%) or fairly (37%) important, resources provide only good or fair formative assessment, with a fifth providing none at all (Table XI). Summative assessment was considered more important by respondents (very - 43%; fairly - 34%) but resources were less likely to provide it (Table XII). Over half the resources did not provide placement tests (tests intended to provide information to place learners in a teaching programme at a level most appropriate to their abilities) for reading (57%), writing (66%) or listening (68%). Again, most respondents thought these tests important. Very few resources provided tests for oral production (not at all - 85%) or oral interaction (not at all - 86%) and views on the tests which were provided were

fairly evenly distributed between 'not at all satisfactory' and 'excellent': though a proportion of respondents thought such tests were not at all important (26% for production, 27% for interaction), more thought they were important (45% for production, 46% for interaction).

Insert Tables XI and XII about here

Three-quarters of the resources which were reported on in the ON-LANG survey provided multiple-choice questions (73%) and a similar proportion (75%) had a fill-in format; half provided closed questions (51%) and a third open questions (34%), with a similar proportion (33%) using a drag-and-drop format. Overall, current resources provide better in the areas which match better with the mechanistic approach of computers (Copeland 2004), but linguists see the more creative and oral aspects of language performance as equally important, and these are poorly provided for at present.

The largest proportion (approaching two-thirds) of resources in the ON-LANG survey covered levels B1 (65%) and B2 (62%) of the Common European Framework (CEF), with somewhat fewer covering levels A1 (55%) and A2 (58%). In other words the resources correspond to Intermediate and Advanced levels on the Languages Ladder, or the English GCSE and A-Level examinations. This may reflect a belief that learners at Breakthrough or Preliminary levels might find it difficult to deal with e-tests (despite the evidence that ICT can support young learners or those with special needs, see for example Wilson, 1999) but a more probable explanation is that commercial suppliers find it more profitable to focus their effort on the examination cohorts: Intermediate and Advanced levels correspond to the level of skill required, for instance, by those learning English in order to take up training or work opportunities. Fewer resources covered levels C1 (43%) and especially C2 (28%), corresponding to Proficiency or Mastery – levels of performance which would

normally be achieved in higher education. Again this can be interpreted as lack of commercial opportunity in competition with higher education institutions.

As mentioned in the previous sections, current European electronic language resources very seldom provide for skills other than reading and writing. The ON-LANG survey included an assessment of the receptive and productive media used by each resource. By 'receptive' materials are meant those included in the resource materials for users to learn from; 'productive' materials are those users can produce in the course of using the resource. The receptive materials in the resources investigated were conventional: almost all resources used text (91%); audio (67%) was more commonly used than graphics (55%) with a minority of resources using video (37%) and few (15%) using multimedia. As might be expected, smaller proportions of resources permitted productive materials, with text (56%) being by far the most common, followed by audio (19%). TLFL resources were significantly more likely to use productive text, than MFL resources, where there was an even split between resources which did and did not use productive text. Two-fifths of the resources did not provide for student creativity at all, with most of the remainder providing only limited or intermediate creativity; very few allowed maximum creativity or allowed students to create a new production (Table XIII). Again, a sizeable proportion of respondents did not comment on the importance of creativity; those who did had more mixed views than for previous questions (Table XIV). The significance of these results in indicating the purpose of the resources reported on in the ON-LANG survey is explored below.

<div align="center">Insert Tables XIII and XIV about here</div>

Discussion

Two major differences between the English scene and the European scene are the range of languages used and the relative importance of formative versus summative assessment.

In other European countries, English tends to be the most commonly taught MFL, because of its utility due to its status as a world language. Since no other single European language has a similar widespread utility in Europe, and due to shifting patterns especially in leisure use, the traditional dominance of French in England has been challenged, and there is also demand for teaching in the home languages of English as an Additional Language (EAL) learners. Commercial providers, especially, are likely to find the demand for these languages less rewarding than providing resources for TEFL, where adult learners require fluency for work or training purposes.

The nature of this demand for TEFL also affects the type of assessment demanded. It was notable that respondents to the ON-LANG survey placed greater emphasis on summative, as compared to formative, assessment, and less emphasis on interaction and creativity, than might be expected in English practice. The resources were also most commonly suited to individual learning rather than group or class work. All these points are consistent with resources being oriented to assessing language skills to meet the requirements of training or work, where clearly-defined skills, especially in writing, are likely to be required, and assessment resources need to validate these skills.

On-line assessment in languages is developing rapidly and needs to be evaluated carefully, just as it is important to evaluate the use of computers in teaching and learning: 'the image of learners simply switched in to computers and learning in isolation is not part of a positive vision' (Hood, 2000:134). There are a number of challenges in on-line testing. One concern is that new technology will drive

assessment back to Neanderthal methods, testing knowledge of a language instead of language skills. According to Warschauer (1996) despite its potential, multimedia software has not really made the expected impact, mainly because of the often questionable quality of the programs which have often been left to commercial developers who do not always base their programs on sound pedagogical principles. 'Technology should not merely replace current practice for the sake of novelty, but must contribute to it and improve it' (Bax, 2000: 209). In the evaluation of the KS3 pilot ICT on-screen test reporting on the reliability and validity of the test Boyle (2006) highlights the issue of potentially novel sources of difficulty. However, he draws attention to the lack of research into such sources of difficulty in such novel assessment methods and outlines future research to hypothesise, categorise and investigate sources of difficulty. Teachers are vital for planning, guidance and extra support, accordingly it is crucial to evaluate the tests available in language programs to ensure that they are able to test or assess what is the important part of learning and therefore make the assessment effective, not merely the type of tests computers are able to perform and therefore focusing on what it is possible to test. If the tests are not pedagogically sound, there is also the risk of reduced effectiveness in teaching methodology if teachers start teaching to the test.

It is also vital to consider how all the different language skills can be tested in an appropriate way. It is critical that the tests focus very much on strategic skills and productive skills, both oral and written, yet this is not what has been encountered in much of the material that has been evaluated. On the contrary, what is currently available is much more "old tests in new technology". The main task in developing on-line testing will be to secure one more step forward in the process of evaluating what is important when it comes to language learning and making intelligent and ever-changing compromise between effective pedagogy and the technical limitations of the medium.

Furthermore, there are issues in ensuring that the students taking part in an on-line test really have the IT skills needed as well as functioning hardware and software. Moreover, it will be essential that on-line testing takes into consideration learners with learning difficulties or students with disabilities.

With the current emphasis on assessment for learning it will be crucial to develop CAA so that it provides genuine formative feedback to individuals and offers practice so that learners can progress, not just a judgement on attainment. However, formative assessment using ICT, with the aim to provide the learner with immediate and relevant feedback at the point of error, is still at an innovative stage. Winkley (2006) discusses the complexity of producing on-screen feedback frames as these are an additional set of conditional screens based on the marked outcomes of the questions. Nonetheless, the DfES are clear in their vision: 'we want to see pilots of a wide range of applications of e-assessment: immediate feedback to learners and teachers, online tests, personalised diagnostics, online assessment and accreditation of e-skills, and inclusion of e-skills in the assessment of all curriculum topics.' (DfES 2005:31)

References

AITCHISON, J. (1989) **The Articulate Mammal**. (London, Unwin Hyman).

ATKINSON, T AND DAVIES, G (2005) ICT4LT Module – Computer Aided Assessment (CAA) and language learning, http://www.ict4lt.org/en/index.htm (accessed 16.11.05)

BARNES, A. AND HUNT, M. (2003) **Effective Assessment in MFL** (London, CILT).

BAX, S. (2000) Putting technology in its place: ICT in MFL teaching, in: FIELD, K. (Ed.) **Issues in Modern Foreign Languages Teaching** (London, Routledge/Falmer, pp208-219).

BLACK, P. AND WILIAM, D. (1998) Inside the black box: raising standards through

classroom assessment. **Phi Delta Kappan**, Volume 10 (see

http://www.pdkintl.org/kappan/kbla9810.htm, accessed 20.6.02, and

www.kcl.ac.uk/depsta/education/publications/blackbox.html)

BOYLE, A. (2006) 'The validity of an innovative on-screen assessment' presented at

IAEA conference, Singapore, available at http://www.iaea2006.seab.gov.sg/,

accessed 19.9.06

BROOKS, V. (2002) **Assessment in Secondary Schools: the new teacher's**

guide to monitoring, assessment, recording, reporting and accountability.

(Milton Keynes, Open University Press).

A Common European Framework of Reference for Languages Learning, Teaching,

Assessment available at

http://www.culture2.coe.int/portfolio/documents_intro/common_framework.html

(accessed 16.12.05)

COPELAND, J. (2004) **The Essential Turing**. (Oxford, Clarendon Press).

DfES (2003) **Languages for All: Languages for Life. A strategy for England.**

(Nottingham, DfES publications).

DfES (2005) **Harnessing technology; Transforming learning and children's**

services. available at http://www.dfes.gov.uk/publications/e-strategy/ (accessed

30.11.05)

DUGARD, C. AND HEWER, S. (2003) **Impact on learning: What ICT can bring to**

MFL in KS3. (London, CILT).

HOOD, P. (2000) Logging on to learning: ICT, modern languages and real

communicative classrooms? in: GREEN, S. (Ed.) **New perspectives on teaching**

and learning modern languages (Clevedon, Multilingual matters, pp.124-137).

Languages Ladder (DfES, 2005) available at

http://www.dfes.gov.uk/languages/DSP_languagesladder.cfm (accessed 14.11.05)

LEVINSON, S. C. (1995) Interactional biases in human thinking in: E.N.GOODY (ed.)

Social Intelligence and Interaction (Cambridge, Cambridge University Press).

National Curriculum for England, Modern Foreign Languages (1999) (London,

DfEE, QCA)

NIELSON, J. *Be Succinct! Writing for the Web*, Alertbox for March 15, 1997:

http://www.useit.com/alertbox/9703b.html (accessed 15.5.06)

ON-LANG available at http://www.on-lang.net/ (accessed 12.12.05)

PINKER, S. (1994) **The Language Instinct** (London, Penguin).

QCA e-assessment vision available at http://www.qca.org.uk/6877.html (accessed

14.11.05)

SLATER, P. AND VARNEY-BURCH, S. (2001) **Info Tech-Multimedia in language**

learning (London, CILT).

TSUTSUI, M. (2004) Multimedia as a Means to Enhance Feedback**, Computer**

Assisted Language Learning, Volume 17, Numbers 3-4 / July, 2004, pp. 377-402

WARSCHAUER, M. (1996) **Computer assisted language learning: an**

introduction available at http://www.gse.uci.edu/markw/call.html (accessed

12.12.05)

WARSCHAUER, M. AND HEALEY, D. (1998) Computers and language learning: an

overview, **Language Teaching***,* Volume 31, Number 2, pp.57-71.

WILIAM, D. AND BLACK, P. (2002) **Working inside the black box** (London,

Department of Education and Professional Studies King's College).

WILSON, D. (1999) Applying New Technologies Appropriately to Modern Foreign

Languages available at http://www.tomwilson.com/david/case/Present1/Index.htm

accessed 12.12.05

WINKLEY, J. (2006) 'Developments in On-Screen Assessment Design for

Examinations' presented at IAEA conference, Singapore, available at

http://www.iaea2006.seab.gov.sg/, accessed 19.9.06

Websites:

http://www.rgshw.com/index2.htm accessed 14.12.05

http://www.ashcombe.surrey.sch.uk/Curriculum/modlang/index.htm accessed

14.12.05

http://web.uvic.ca/hrd/hotpot accessed 14.12.05

http://www.quia.com accessed 14.12.05

http://www.qca.org.uk/15350.html accessed 19.9.06