Assessing English as a Second Language: From Classroom Data to a Competence-Based Open Learner Model

Susan BULL[a*], Barbara WASSON[bd], Michael KICKMEIER-RUST[c], Matthew D. JOHNSON[a], Eli MOE[d], Cecilie HANSEN[d], Gerhilde MEISSL-EGGHART[e], Klaus HAMMERMUELLER[e]

[a] *Electronic, Electrical and Computer Engineering, University of Birmingham, UK*

[b]*Department of Information Science & Media Studies, University of Bergen, Norway*

[c]*Knowledge Management Institute, Technical University of Graz, Austria*

[d]*InterMedia, Uni Health, Uni Research AS, Norway*

[e]*Talkademy, Austria*

*s.bull@bham.ac.uk

Abstract: With the increase of ICT in classrooms comes much data that can be used for evidence-based assessment. We focus on harnessing and interpreting this data to empower teachers in formative assessment. We describe e-assessment of English as a Second Language and illustrate how we move from data collected in classroom activities, through an automated assessment method, to visualising competence levels in an open learner model.

Keywords: second language, evidence-based formative assessment, open learner model

Introduction

Today's classrooms may comprise a range of tools [1] producing much data that can be tapped to support formative assessment. There is a need for methods to capture and present the data so teachers can interpret and transform it to a meaningful form for students and themselves. We are developing such tools and methods for English as a Second Language. We describe moving from classroom data, through an automated assessment method, to an open learner model (OLM) for use by teachers to support their formative assessment work.

The Common European Framework of Reference for languages (CEFR) offers competence-based common reference levels in language learning [2]. These are based on language use and abilities (what students *can do*). CEFR is not detailed enough to design diagnostic testing items or define task difficulty, but is a useful starting point [3]. A similar focus is at the forefront of many current language courses and applications. In Norway, for example, a specified set of learning goals and competences must be integrated into English teaching in schools [4], and teachers plan activities to address the competences. Our OLM provides students and teachers with an overview of current competence levels, enabling better planning of teaching and student recognition of their learning. The approach also offers a way to facilitate teachers' classroom orchestration [5].

In this paper we introduce the OLM as a teacher and learner feedback tool, describe data available to teachers, how they can transform interaction data to include in a learner model, and outline how such data may be displayed to help raise awareness of competencies.

1. Open Learner Models and Classroom Data

A learner model is a representation of a user's skills and abilities, as inferred during their interactions, and enables a system to adapt to the needs of the individual. Increasingly, learner models are being opened to users as a means to help prompt learner reflection, help teacher planning and decision-making, etc. [6]. There are now also strong arguments for placing OLMs in the centre of contexts where there are multiple sources of data available for the learner model [7],[8],[9] since a variety of tools are in use in classrooms [1]. While an OLM can be likened to technology-based student progress and performance reports, rather than reporting progress, it *models* and externalises competences and skills. The problem in technology-rich classrooms is that data is not always available in a form that matches competence descriptors, and is often not able to pass data to a learner modelling service. We therefore offer teachers a means to transform activity data for an OLM.

Usually activity results are stored with scores or qualitative descriptors in an overview. An illustration of a teacher's spreadsheet recording results is given in Figure 1. This allows the teacher to see at a glance, how an individual is progressing in goal-related competences. As time advances and further items are added, we expect to see a shift towards good and excellent - as is indeed happening in this example. We aim to support teachers with an approach that is similar to their self-generated methods (e.g. Figure 1), or methods with which they are already familiar, but providing a focus on overviews of *current competences*. These can be presented through an OLM, so students may more readily recognise the importance of competences (rather than specific activities), and teachers can gain an overview they can act on in the classroom or in later planning.

	Goal 1	Goal 2	Goal 3	Goal 4	Goal 5
Level test	Starting competence	Starting competence	Starting competence	Good competence	Excellent competence
Test politics and democracy	Starting competence	Low competence	Good competence	Low competence	Good competence
Student comments					
Homework test	Excellent competence		Good competence	Low competence	Good competence
group work	Good competence	Good competence	Good competence	Low competence	Excellent competence
Student comments					
Self assessment	Excellent competence	Good competence	Excellent competence	Good competence	Good competence
Oral presentation		Excellent competence	Excellent competence		Excellent competence
5 hour test	Good competence	Excellent competence	Good competence	Excellent competen	Excellent competence

Figure 1: Example of a teacher's record of competencies that combines colour with text

This is in line with education policy in Europe moving from a focus on knowledge to a focus on competence. For example, in Norway, the learning goals and competences cover three areas: communication; language learning; culture, society and literature – each of which comprises sets of competences [4]. For example, two of the "communication" competences are that after four years of English students should be able to "read and understand the main content of texts on familiar topics" and "understand and use common English words and phrases related to daily life, leisure time and interests, both orally and in written form". Teachers plan how to incorporate appropriate activities into their classrooms to enable students to develop the competencies.

We illustrate with a set of activities aimed at 11-12 year-olds, including an electronic reading and listening test; interactions in a virtual world (Second Life); and an electronic self-assessment (from the European Language ePortfolio). Assessment methods, automatic and manual, are applied to data from these activities to determine achievement level for relevant competencies. The first activity, the online listening and reading test, has a mix of item types: multiple choice, click item, click text, click name, click word, move paragraph. Each item is weighted according to difficulty by professional test developers and these weights, along with student answers and other test item information, is used by ProNIFA (an automatic assessment method – see below), to generate competence levels for students taking the test before data is passed to the OLM. The second data set derives from activity within Second Life, and includes chat logs and video recordings of activity in 3D space. For example, from Second Life we get (i) a simple chat log file (time stamp, chatting person/entity, chat text); (ii) a set of competencies (CEFR skills [2] shown below), specified in a text file (number, id, initial probability that students have that skill, short description); and educator-defined (scripted) rules, which vary from very simple such as checking whether a certain entity writes a certain text; to more complicated, such as computing distances travelled in Second Life. ProNIFA parses the log files, checks whether the rules apply and updates the probabilities of the competencies (and the probability distribution over the competence states).

*(i) [07:21 UTC] <i>Teacher</i>Well done, Svein.
*
(ii) 001 CEFR#094 0,5 Listening A1
(iii) [Rule1] Who=*Teacher What=Well done, <NAME>. ASkills=1;2 AUpdate=0,2 LSkills=3 LUpdate=0,1*
***NB:** If the teacher says "Well done" and a name, the probabilities of skills 1 and 2 for learner <NAME> are increased by 0.2; and for skill 3, decreased by 0.1.*

The third data set is produced by student self-assessments. The European Language ePortfolio self-assessment grid was used to elicit self-assessment of speaking, listening and reading skills. Questions relate to various "can do's", e.g. "I can understand simple, short greetings and expressions, such as hello, thank you or you are welcome" and students assess themselves between "I can do this a bit / quite well / very well". The teacher interprets these data sets and the results are manually entered directly into the OLM – i.e. not all data needs to be transformed using ProNIFA.

As explained above, not all data is immediately available in competence form, and needs to be assessed either automatically or manually. ProNIFA (probabilistic non-invasive formative assessment) is a tool to support teachers in the assessment process. It establishes a user interface for data aggregation and analysis services and functions. Conceptually, the functions are based on Competence-based Knowledge Space Theory (CbKST), originally established by Doignon and Falmagne [10], a well-elaborated set-theoretic framework for addressing the relations amongst problems (e.g. test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. While the original idea considered performance (behaviour, e.g. solving a test item), extensions introduced a separation of observable performance and latent, unobservable competencies, which determine the performance [11]. CbKST assumes a finite set of more or less atomic competencies (in the sense of some well-defined, small scale descriptions of some sort of aptitude, ability, knowledge, or skill) and a prerequisite relation between those competencies. A prerequisite relation states that competency a is a prerequisite to acquire another competency b. If a person has competency b, we can assume they also have competency a. Because more than one set of competences can be a prerequisite for another (e.g., competency a or b are a prerequisite for acquiring competency c), prerequisite functions have been introduced, relying on and/or type relations. A person's competence state is described by a subset of competencies. Due to the prerequisite relations between competencies, not all subsets are admissible competence states. Using interpretation and representation functions, the latent competencies are mapped to a set of tasks (or test items) covering a domain: mastering a task correctly is linked to a set of necessary competencies; not mastering a task is linked to a set of lacking competencies. This assignment induces a performance structure: the collection of all possible performance states. Recent versions of the conceptual framework are based on probabilistic mapping of competencies and performance indicators, accounting for lucky guesses or careless errors. This means, mastering a task correctly provides evidence for certain competencies and competence states, with a certain probability.

ProNIFA retrieves performance data and updates the probabilities of competencies and competence states in a domain. When a task is mastered, all associated competencies are increased in their probability, and failing in a task decreases the probabilities of associated competencies. A distinct feature in formative assessment is the multi-source approach. ProNIFA allows connecting the analysis features to a range of evidence sources (such as the listening and reading test or activity in a virtual world). The interpretation of the sources of evidence depends on a-priori specified and defined conditions, heuristics and rules, which associate sets of available and lacking competencies to achievements exhibited in the evidence. The idea is to define certain conditions or states in a given environment, for example: the direction and speed a learner is moving, following instructions in English in an adventure game, or a combination of correctly and incorrectly ticked multiple choice tasks in a regular online test. The specification of such states can occur in multiple forms, ranging from simply listing test items and the correctness of the items, to complex heuristics such as the degree to which an activity reduced the 'distance' to the solution in a problem solving process (technically this can be achieved by pseudo code scripting). The next step of this kind of planning/authoring is to assign a set of competencies that can be assumed available and also lacking when a certain state occurs. This assumption can be weighted with strength of the probability updates. In essence, this approach equals the conceptual framework of micro adaptivity (e.g. [12]). Figure 2 shows ProNIFA-analysed data from a Second Life activity (see Section 1). The resulting model built around atomic competencies and related probability distribution, is passed to an OLM platform as a next step to support teacher appraisal efforts (Figure 3).

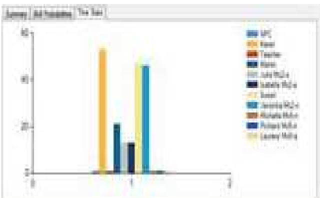

Figure 2: Screenshot of ProNIFA

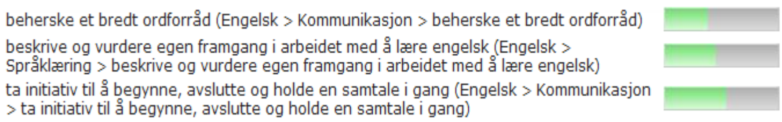

Figure 3: OLM skill meters

2. Competence Visualisation using an Open Learner Model

Using the easy-to-interpret ProNIFA display, teachers can add competency information to the OLM, as shown in Figure 4. They provide a numerical value for the model (by clicking on the stars) and may also include additional (non-modelled) feedback. The example shows competences in English according to the required learning goals and competences [4]. So, for example, if ProNIFA-analysis of recent Second Life logs indicates increased competence in some aspect of a student's learning, the teacher can easily update the OLM accordingly. This can happen alongside other, possibly automated input to the learner model, self-assessments, etc., if other activities are also ongoing. Thus, both teachers and students can flexibly use the OLM for formative assessment support.

Figure 4: Teacher updates to the OLM

As stated previously, information at this broad level of granularity is intended primarily to help gain a quick overview of students' competences which can, for example, be highly useful in classrooms where teachers are trying to manage classroom activities, give formative feedback, or update their teaching plan. In addition to the simple skill meters (Figure 3), student rankings by competence, and a table overview are available. Work is underway on word clouds – providing another way for teachers to quickly identify where to focus their attention [13]; and treemaps, which will allow drill-down to more detail, supporting more reflective formative assessment. These (and possibly other) learner model views will help teachers easily interpret the kind of information they already collect (e.g., Figure 1), but in a more immediately usable format (or, in the case of the planned treemaps, in a way that facilitates access to detail). Student use of the OLM, as well as promoting awareness of their learning [6], will help focus students on thinking in terms of competences (for English [4]), rather than activity-specific results (as in the example in Figure 1).

3. Summary

This paper has introduced a way to help teachers take the range of data now available about students, and transform it into a form that can be used in an OLM. This can help students note the importance of language competences, and help teachers' classroom orchestration.

Acknowledgements

The project is supported by the European Community (EC) under the Information Society Technology priority of the 7th Framework Programme for R&D, contract no 258114 NEXT-TELL. This document does not represent the opinion of the EC and the EC is not responsible for any use that might be made of its content.

References

[1] Richardson, W. (2009). *Blogs, Wikis, Podcasts, and Other Powerful Web Tools for Classrooms,* California: Corwin Press.
[2] Council of Europe (nd). The *Common European Framework of Reference for Language*s, http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf. Accessed 1 June 2012.
[3] Huhta, A. & Figueras, N. (2004). Using the CEF to Promote Language Learning through Diagnostic Testing, In K. Morrow (ed), *Insights from the Common European Framework*, Oxford: OUP, 65-76.
[4] Utdanningsdirektoratet. Læreplan i Engelsk, http://www.udir.no/Lareplaner/Grep/Modul/?gmid=0&gmi=166919 . Accessed 1 June 2012.
[5] Dillenbourg, P. & Jermann, P. (2010) Technology for Classroom Orchestration, In M. S. Khine & I. M. Saleh (eds). *New Science of Learning: Cognition, Computers and Collaboration in Education,* Berlin: Springer Verlag, 525-552.
[6] Bull, S. & Kay, J. (2007). Student Models that Invite the Learner In: The SMILI Open Learner Modelling Framework, *International Journal of Artificial Intelligence in Education* 17(2), 89-120.
[7] Morales, R., Van Labeke, N., Brna, P. & Chan, M.E. (2009). Open Learner Modelling as the Keystone of the Next Generation of Adaptive Learning Environments. In C. Mourlas & P. Germanakos (eds), *Intelligent User Interfaces*, Information Science Reference, 288-312, London: ICI Global.
[8] Mazzola, L. & Mazza, R. (2010). GVIS: A Facility for Adaptively Mashing Up and Representing Open Learner Models, In M. Wolpers et al (eds), *EC-TEL 2010*, Berlin: Springer Verlag, 554-559.
[9] Reimann, P., Bull, S., Halb, W. & Johnson, M. (2011). Design of a Computer-Assisted Assessment System for Classroom Formative Assessment, *CAF11*, IEEE.
[10] Doignon, J., & Falmagne, J. (1999). Knowledge Spaces. Berlin: Springer Verlag.
[11] Korossy, K. (1999). Modelling knowledge as competence and performance. In D. Albert & J. Lukas (Eds.), *Knowledge spaces: Theories, empirical research, and applications* Mahwah, NJ: LEA, 103–132
[12] Kickmeier-Rust, M. D., & Albert, D. (2011). Micro adaptivity: Protecting immersion in didactically adaptive digital educational games. *Journal of Computer Assisted Learning*, 26, 95-105.
[13] Reimann, P., Bull, S. & Ganesan, P. (in press). Supporting the Development of 21st Century Skills: Student Facilitation of Meetings and Data for Teachers, *TAPTA Workshop Proceedings*, EC-TEL 2012.