

## Report on the ECLM workshop in Graz, 24-26 November 2010

“Class RelEx!: Training in Classroom Assessment related to the Common European Framework of Languages (CEFR)”

Participant: Dr. Jessica Allen Hanssen, University of Nordland, Bodø

The European Centre for Modern Languages (ECML) of the Council of Europe, which represents a diverse and wide-ranging group of European language professionals, is an important resource for the planning, development, execution, distribution, and evaluation of projects related to the use of and instruction in languages in the European Union and its partner nations.

The workshop I attended was on the topic of testing and evaluation in the classroom as they relate to the stated goals of the CEFR. This workshop focused upon disseminating the research of the ClassRelEx! medium-term project; its members led the workshop. The ClassRelEx! project itself concerns the value of relating language examinations to the CEFR, and as such requires some familiarization with the CEFR guidelines. Therefore, after the project team and its working methods were introduced (and with a lovely little “Hollywood” icebreaker for everyone!), workshop participants were given an introduction to linking phases with an emphasis on CEFR familiarization. Participants worked on understanding the importance of important self-evaluation in the ranking of one’s L2 abilities, from the points of view of both the teacher and the student. We were asked to consider the types of assumptions that teachers can and should make when it comes to self-evaluation, and to notice the subtle differences between understanding a language’s vocabulary in and out of context. It was emphasized that the CEFR was not meant as an evaluative tool for one’s L1, which is a mistake that is frequently made by even the most well-intended CEFR users.

We then moved into various exercises on oral production and oral interaction, and we worked through an assessment gap exercise which I found to be very relevant and useful, in which we were asked to make group judgments on the CEFR levels of numerous examples based on descriptors for oral production and oral interaction. My group included members from Finland, Cyprus, and Montenegro, and we worked very efficiently and had a fruitful dialogue about our classroom experiences and practices as they relate to language training. We concluded the session with a round-table discussion of our groups’ findings and their implications with respect to self-evaluation. In the evening, workshop members enjoyed a fantastic “multicultural” cocktail party, in which we came together to share food, drinks, and entertainment from our various countries.

The second day of the workshop moved beyond self-evaluation and into assessment, which was the primary target area for the ClassRelEx! project. The day began with a highly informative presentation on specification and identifying valid items and tests. We learned about the varying degrees of quality of tests in a general sense, and then focused the discussion onto analyzing “bad items” in groups. The “bad items” were generally of poor quality due to their lack of relevance, difficulty, or ambiguous context, making it difficult to assess the test-taker’s true L2 abilities. We learned about, for example, the problems of using “true-false” and “matching-type” questions to make an accurate ability assessment, since there is a high degree of chance to be reckoned with. The strengths and weaknesses of various national and locally-used tests were compared, and I am pleased to report that the Nordic countries were applauded for generally presenting tests which were considered to be of good quality. The so-called “Dutch Grid” for considering the validity of test questions was

introduced and applied to good and bad test items for comparison purposes, again in the group context. There was a lively conversation about the role of testing in the classroom, which included the pitfalls of “piloting” a test or pretesting. After the day’s work, workshop participants were treated to an excellent city tour of Graz, a city of many contrasts between its tradition and modernity, and a delicious farewell dinner at a local restaurant, the Schiefe Laterne or “Leaning Lantern”.

The third day of the workshop was by far the most content-heavy and informative session. We began with a presentation entitled “A Taste of Benchmarking and Standard Setting,” which introduced and explained the various terminology used when deciding which test questions were best for demonstrating performance. The importance of group work during the benchmarking phase of test writing was emphasized, as was the importance of testing a particular item enough times to produce a valid statistical result before “going live” with it as a functional test item. The presentation also examined the process of standard setting, which was defined as the process of arriving at a cut-off score for a given examination. Again, the value of group work was emphasized, and the key words were “authenticity,” “relevance,” and “objectivity.” After the presentation, we were given the opportunity to put the ideas into practice in our work groups by producing statistical samples of test results.

We then considered the issue of benchmarking with respect to the first day’s emphasis on oral production; this time, however, the emphasis was on oral testing. We watched DVDs of real students’ oral performances, and were asked to evaluate them in our groups, which was a very useful and fun exercise. It is possible to replicate our experience by looking at the following website: [http://www.ciep.fr/en/publi\\_evalcert/dvd-productions-orales-cecrl/videos/english.php](http://www.ciep.fr/en/publi_evalcert/dvd-productions-orales-cecrl/videos/english.php). One can use this website to see and evaluate the oral examinations displayed on this website based on established CEFR criteria. We followed up on the idea of test-building with a discussion of validation, which is determining whether a given test indeed measure what it intends to measure. The importance of psychometric data was emphasized, as was the importance of eliminating barriers to effective test taking: a valid test does not, for example, test one’s ability to comprehend the directions for the given test instead of actually testing the learner’s ability to demonstrate his or her language skills. It was also stressed that too many commercial test manufacturers link their tests to the CEFR without proper evidence that their tests indeed correctly use the CEFR as a point of reference.

The workshop concluded with dissemination of the ClassRelEx! draft manual, which workshop members were asked to provide commentary on before it goes to publication. The manual, which is now available at the EC ML in Graz is designed to provide language teachers with a clear reference point for related their language examinations to the CEFR in three stages: learning, teaching, and assessment. We all received certificates for our participation in the workshop, and were urged to use the knowledge we gained to continue the process of encouraging the CEFR as a part of language instruction in our home countries.

The workshop was highly informative, relevant, and well-organized. I look forward to continuing the excellent work of the ClassRelEx! project by allowing it to influence the way I train my teacher students in Nordland County in evaluative strategies.

March 2011, Bodø

