

Wisniewski Katrin, Abel Andrea & Vettori Chiara

The influence of raters' linguistic background on their interpretation of CEF scales

The focus of this contribution is the discussion of problematic aspects in the behaviour of raters belonging to two different language groups when using CEFR rating scales in a writing test. The test is part of the KOLIPSI project the aim of which is to evaluate L2 competences of a representative sample of high school pupils in multilingual South Tyrol. It has parallel versions in Italian and German in order to facilitate detailed in-depth comparisons between these two language groups, and it has been related to the *Common European Framework of Reference (CEFR)*, following its theoretical premises and guidelines in designing tasks and a rating grid.

CEFR scales are designed to work independently of the language being assessed. Yet, KOLIPSI benchmarking meetings showed that the cultural/linguistic background seemed to play a considerable role in the behaviour of German/Italian speaking raters. If the interpretation of CEFR scales even after a rater training and several follow-up meetings is dependent on raters' cultural/linguistic background, this suggests that maybe the vague formulations of many CEFR scale descriptors endanger cross-language validity, leaving too much space for - possibly even incompatible - interpretations.

It's possible that even if raters rank candidates consistently, according to their linguistic background they might focus on very different aspects of rating scales. To find out which aspects of performance raters of each language group actually pay attention to, appropriately designed rater questionnaires have been used, the analysis of which constitutes the data basis of this contribution.