

Author: Liliann Byman Frisé  
assessment

TITLE: Teachers' conceptualizations of  
of L2 English oral proficiency reflected in  
note-taking practices of a high-stakes,  
national test

Your full name: Liliann Byman Frisé

Affiliation: Karlstad University, Sweden

Current position: Phd-student in English linguistics

Title of your paper:

**Teachers' conceptualizations of assessment of L2 English oral proficiency reflected in note-taking practices of a high-stakes, national test**

**Abstract** (300 words)

In Sweden, all students in years 6 and 9 (ages 12/13 and 15/16) take a national test in English as a second language (L2), where oral proficiency, written proficiency and receptive skills are tested. It is a high-stakes test assessed by students' own teachers, where the Swedish Agency for Education provide teachers with guidelines and criteria in focus for assessment. Despite the fact that assessment guidelines are available, some teachers create and use their own scoring rubric when operationalizing assessment of the oral proficiency part of the test. The aim of the study is to gain new knowledge about what role teachers' scoring rubrics play in the assessment situation, to reveal teachers' conceptualizations of L2 English oral proficiency assessment. Data consists of interviews ( $N= 13$ ) with teachers of English that all act as raters of the speaking part of the national test in L2 English (grade 6 and/or 9). Research questions focus how teachers use their scoring rubric for note-taking in the assessment situation, in what way they draw upon their notes when deciding the grade, and reasons behind the creation and use of a scoring rubric. Interviews were recorded and transcribed. Method of analysis is qualitative thematic analysis, and theoretical framework is the *Anthropological Theory of Didactics*. Preliminary results indicate that teachers' scoring rubrics created an overview of students' attainment of criteria. When coming to a decision about individual students' grades, teachers expressed that scoring rubrics were used as mnemonic devices in a "negotiation" of strengths and weaknesses shown by students in relation to criteria. Scoring rubrics also facilitated note-taking in situ, as teachers "ticked off boxes" instead of writing extended notes, something that eased students' test-anxiety. Moreover, scoring rubrics facilitated formative assessment, indicating that *accountability* might be part of the discourse behind their use in a summative test.

**Extended summary** (993 words, excluding reference list)

**Introduction**

National tests in English are administered in almost all European countries, and in a majority of countries they carry high-stakes for students (European Commission, EACEA, & Eurydice, 2015). In Sweden, the Swedish National Agency for Education administers every year a national test in English as a second language (L2) for all students in year 6 (ages 12-13) and year 9 (ages 15-16) that includes oral proficiency, receptive skills and written proficiency. Part A of the test, the National English Speaking test (NEST), tests students' L2 oral proficiency. It is a high-stakes, summative proficiency test. Contrary to common practise for high-stakes tests internationally, where external raters are used, NEST is typically administered and assessed by the students' own English teachers (Crisp,

2013). Students are divided into pairs and are instructed to react to and discuss topics from the test material.

Despite extensive assessment guidelines from the Swedish National Agency of Education, some teachers create and use their own scoring rubric when operationalizing assessment of the NEST (Byman Friséén, Sundqvist & Sandlund, 2021). The aim of the present study is to gain new knowledge about what role teachers' scoring rubrics play in the assessment situation, to reveal teachers' conceptualizations of L2 English oral proficiency assessment. The following research questions guided the study:

- RQ1: In what way(s) do teachers use scoring rubrics for note-taking in the assessment situation?
- RQ2: In what way(s) are scoring rubrics used when deciding the grade?
- RQ3: What underlying reasons are reported for the creation and use of self-made scoring rubrics for assessment of the NEST?

## Theoretical background

The theoretical framework for the study is the *Anthropological Theory of Didactics* (ATD) (Chevallard, 2007). According to ATD, content to be taught, assessed and learnt is co-constructed on a hierarchy of levels and therefore subject to *didactic transpositions* (Achiam & Marandino, 2014). Knowledge is thus a changing reality, formed and affected by the institution and participants within which it exists. Scholarly knowledge of a subject might be different from teachers' knowledge, which in turn might be different from learners' knowledge, and this is referred to as *the institutional relativity of knowledge* (Achiam and Marandino, 2014). The idea of *praxeologies*, which consists of *praxis* and *logos*, is part of the ATD framework. *Praxis* is a type of *task* as well as the *technique* used to carry out the *task*, whereas *logos* is the logic behind using that particular *technique* for that particular *task*. Viewing scoring rubrics as the *technique* used to carry out the *task* of assessing L2 English oral proficiency (OP), the ATD framework is used in this study to analyze the *logos* behind the *technique* – i.e., the discourse of why scoring rubrics are beneficial for carrying out the *task*. As such, teachers' institutional knowledge of assessment of L2 English OP will be revealed.

## Methods

Data consists of interviews with teachers of English in Sweden ( $N = 13$ ) who all act as raters of the NEST, for grades 6 or 9. Data were retrieved in two phases; a first phase in 2017 and a second phase in 2019.

Participants for the first phase ( $N = 5$ ) were selected in connection with a previous project (Byman Friséén, Sundqvist & Sandlund, 2021). The first phase interviews generated new questions about teachers' assessment practices when using scoring rubrics for assessment of L2 English OP, which resulted in a revision of the interview guide and led to phase two, where more in-depth questions to interviewees were asked.

Interviewees for the second phase of data collection ( $N = 8$ ) were selected from professional networks of teachers in grades 6 and 9. A criterion for being considered for the interviews was that teachers were to rate the NEST in years 6 or 9 in the autumn of 2019.

The 13 interviews were audio recorded and transcribed orthographically. The method of analysis of data was qualitative thematic analysis (Braun and Clarke, 2006), for which the software program

Author: Liliann Byman Friséén  
assessment

TITLE: Teachers' conceptualizations of  
of L2 English oral proficiency reflected in  
note-taking practices of a high-stakes,  
national test

NVivo 12 was used. The analysis was guided by the research questions for the study as well as the ATD framework (see Achiam & Marandino, 2014; Chevallard, 2007).

### Preliminary findings

Preliminary results indicate that teachers' scoring rubrics created an overview of students' attainment of criteria. When reaching a decision about individual students' grades, a majority of the interviewed teachers expressed that the scoring rubric is used as a mnemonic device in a "negotiation" of strengths and weaknesses shown by students in relation to given test criteria. One aspect that appears to be imperative for which grade to award students' test performances was their engagement in the test situation. Possibly because of this, several of the interviewed teachers stressed the importance of reducing students' anxiety in the test situation. One way to do this was to write as little as possible while listening to the students. Scoring rubrics with pre-printed level descriptors facilitated note-taking in situ, as teachers "ticked off boxes" instead of writing more extended notes. In addition, preliminary results indicate that despite the summative aim of the test, scoring rubrics are used for formative assessment to students and/or parents, indicating that *accountability* might be part of the discourse behind this *technique*.

### Theoretical and educational significance

Results from the study might lead to new knowledge when it comes to raters' tacit assessment knowledge as well as their assessment processes. The *institutional relativity of knowledge* (Achiam and Marandino, 2014) that emerged through the study can be used to compare and contrast knowledge of assessment of L2 English OP on other levels in the educational system (such as assessment guidelines from the Swedish National Agency for Education).

### Relevance to the QUINT ambition

Extensive assessment guidelines to teachers that act as raters of a national test can constitute one way to enhance inter-rater reliability. However, in order to make guidelines operable, adjustments must be made, and therefore, didactic transpositions of the test construct will most likely occur. This study sheds light on teachers' assessment knowledge, as well as what assessment quality entails in this particular situation.

### Reference list

- Achiam, M., and Marandino, M. (2014). A framework for understanding the conditions of science representation and dissemination in museums. *Museum management and curatorship*, 29(1), 66-82.
- Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101
- Byman Friséén L, Sundqvist P, Sandlund E. Policy in Practice: Teachers' Conceptualizations of L2 English Oral Proficiency as Operationalized in High-Stakes Test Assessment. *Languages*, 2021; 6(4), 1-23. <https://doi.org/10.3390/languages6040204>
- Chevallard, Y. (2007). Readjusting Didactics to a Changing Epistemology. *European Educational Research Journal*, 6(2), 131-134.
- Crisp, V. (2013). Criteria, comparison and past experiences: how do teachers make judgements when marking coursework? *Assessment in Education*, 20(1), 127-144.

Author: Liliann Byman Frisé  
assessment

TITLE: Teachers' conceptualizations of  
of L2 English oral proficiency reflected in  
note-taking practices of a high-stakes,  
national test

European Commission, EACEA, & Eurydice. (2015). Languages in Secondary Education: An Overview of National Tests in Europe – 2014/15. Retrieved from <https://op.europa.eu/en/publication-detail/-/publication/62ac43c3-dac4-11e5-8fea-01aa75ed71a1/language-en>

Author: Liliann Byman Frisé  
assessment

TITLE: Teachers' conceptualizations of  
of L2 English oral proficiency reflected in  
note-taking practices of a high-stakes,  
national test