Scoring Rubrics and Google Scripts:
A Means to Smoothly Provide Language Learners with Fast Corrective Feedback and Grades

Sylvie Thouësny*

Abstract. Language teachers, as one might expect, are often confronted with the task of assessing and grading students’ assignments, which should ideally be addressed with respect to not only reliability and validity, but also functionality. Based on Knoch’s (2011) taxonomy features with regards to design and development of writing assessments, an analytic approach was devised to assign scores to a certain amount of independent aspects of language learners’ performance through the means of specific rubrics. The rubrics, elaborated with the graders and students in mind, describe the rating of the various tasks intermediate learners of French had to undertake. Following a brief description of one assignment, this short paper highlights the significance of following the scoring grids to maintain a relatively constant grading style across students and teachers alike. Additionally, it illustrates how Google documents and forms, used in conjunction with simple and undemanding scripts, assisted in the process of correcting and providing students with timely feedback.

Keywords: scoring rubrics, Google scripts, corrective feedback, language learning.

1. Introduction

Responding to Yancey’s (1999) call for a “fourth wave” of writing assessment (p. 500), i.e., a call for assessment that moves beyond multiple-choice questions, scored essay tests, and portfolio assessments, Wardle and Roozen (2012) propose an ecological model of writing assessment that considers both the vertical and horizontal dimensions of students’ development. In short, they emphasise the idea that “the breadth of students’ [...] literate experiences – in and out of school – impacts their ability to ‘do’ academic literacy tasks” (Wardle & Roozen, 2012, p. 107, emphasis in original). While the authors’ reflection seems appropriate and full of common sense, adapting their model...

* Contact author: sylvie.thouesny@icall-research.net
of writing assessment would be a rather complicated endeavor as it implies including data from different modules and other non-canonical texts from outside the academic structures.

Assessing learners’ writing is a challenging task for teachers, as it involves serious considerations on how to evaluate the learners’ ability to write. As Behizadeh and Engelhard (2011) reported, “[i]t was much easier to objectively score multiple-choice tests, as compared to complex writing samples” (p. 202). Whichever approach is used, one aim of assessing and evaluating written language is to diagnose learners’ weaknesses and strengths, in other words, to “identify those areas in which a student needs further help” (Alderson, Clapham, & Wall, 1995, p. 12). Knoch (2011) further points out that “integral to diagnostic writing assessment is the rating scale” (p. 81).

Weigle (2002) discusses factors that should be considered when designing rating scales and scoring rubrics, such as who is going to use them and how the scores will be reported (pp. 122-124). Rubrics are generally defined as “systematic scoring guidelines to evaluate students’ performance [...] through the use of a detailed description of performance standards” (Zimmer, 2007, p. 1). Knoch (2011) observes that “no currently available theory can serve by itself as a basis for the design of a rating scale for writing for diagnostic assessment” (p. 90). To remedy, she proposes a taxonomy of features – accuracy, fluency, complexity, mechanics, cohesion, coherence, reader/writer interaction, and content – that could be used as a basis for the design of a rating scale (Knoch, 2011, p. 91).

Based on Zimmer’s (2007) practical steps in designing rubrics to grade a student’s performance, and Knoch’s (2011) taxonomy features, this short paper illustrates how rubrics combined with Google scripts help provide language learners with fast corrective feedback and grades.

2. Educational setting and assignment

Students were enrolled at university level in various Bachelor degrees in which the French language was either an obligatory or a facultative component of their formation. As part of their assignment, they had to experience French autonomously and were asked to write an account of their activities as well as to reflect on their learning outcomes. A minimum of 800 words with at least eight different entries was imposed on students.

3. Elaboration of the scoring rubrics

Learners’ written documents were assessed according to five criteria carrying equal weight: (1) content, (2) vocabulary, (3) conjugation, (4) syntax, and (5) reflection. The first category identified whether the texts provided were adapted to the task and whether the instructions were respected. The second criterion of assessment acknowledged
the lexical diversity, whether the vocabulary was sufficient for explaining the chosen activities. The third criterion identified whether the verbs were correctly conjugated with an appropriate tense. The fourth criterion evaluated the syntax of the sentences, whether the structures were simple or complex. Finally, the last criterion examined the breadth of the learners’ reflection and whether it was thoughtfully considered.

All criteria were rated on a 5-point rating scale in accordance with the university’s regulations*, i.e., Fail (<40%), H3 (40-49%), H2.2 (50-59%), H2.1 (60-69%), and H1 (>=70%), as illustrated below with the “reflection” criterion (Table 1). Note that the original rubrics were in French, but were translated into English for the purpose of this publication.

Table 1. Scoring rubrics

<table>
<thead>
<tr>
<th></th>
<th>Fail (&lt;40%)</th>
<th>H3 (40-49%)</th>
<th>H2.2 (50-59%)</th>
<th>H2.1 (60-69%)</th>
<th>H1 (&gt;=70%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflection</td>
<td>No reflection on the language learning is present. Newfound skills cannot be identified.</td>
<td>Reflection on language learning is inept. Newfound skills can be identified.</td>
<td>Reflection on language learning is still inept, but newfound skills are clearly identifiable.</td>
<td>Good reflection on language learning and newfound skills are clearly identifiable.</td>
<td>Excellent reflection on language learning and newfound skills are clearly identifiable.</td>
</tr>
</tbody>
</table>

4. Template, form and script to report grades

Google documents, forms, and scripts were used to assess the learners’ written account of their learning experience, as well as to provide them with their respective scores. The template, as illustrated below in Figure 1, was created with the same word-processing tool. The shaped document served as a starting point for each student, and variables – preceded with the @ sign – were automatically filled in depending on the teacher’s scoring.

Figure 1. Template

The Google form (Figure 2) enabled the assessment of the written texts in accordance with the scoring grades. Once the data was entered into the form, and that the form was saved, the data was sent to the spreadsheet that was automatically generated by the system.

Figure 2. Form

The aim of the script was then to get the last data entered in the active spreadsheet so as to fill the information into the template. Each variable in the template was replaced with its corresponding data, as illustrated in Figure 3 below*

* A more detailed account on the script can be found on my website: http://blog.icall-research.net/2011/10/23/template-and-script-to-report-grades-of-learners-of-french
The result was as follows (Figure 4):

Each score sheet was accompanied with the score grids so that students knew exactly what the rating on each section implied. This sheet could then have been shared with the student through Google Drive (formerly Google Docs) or sent via email.
5. Discussion and conclusion

This short paper illustrated how scoring grids combined with the use of Google documents, forms, and scripts would assist in assessing and scoring texts written by learners as a response to an open task. Although Kulhavy (1977) reported that “delaying the presentation of feedback for a day or more leads to significant increases in what students remember on a retention test” (p. 214), more recent research has highlighted the significance of timely feedback (e.g., Denton, Madden, Roberts, & Rowe, 2008). By following a scoring grid and using a script that automatically fills in a template, it is manageable to have a rapid and constant grading across students and tasks, which means that variations due to tiredness or subjectivity can be easily avoided. It is nevertheless always pertinent to perform a second rating by another rater, which will indicate fairness to students (see Penny and Johnson (2011) for a classification of resolution procedures in scoring writing assessment).

References