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

Sylvie Thouësny*

Independent Researcher, Dublin, Ireland

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

In L. Bradley & S. Thouësny (Eds.), *CALL: Using, Learning, Knowing, EUROCALL Conference, Gothenburg, Sweden, 22-25 August 2012, Proceedings* (pp. 286-291).

of writing assessment would be a rather complicated endeavour 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" (Zimmaro, 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 Zimmaro'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

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

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

2011/2012
Barème de notation et résultats
Please note that these results are provisional and will only be confirmed after the examination board meeting.

Date de la correction : @date

Nom du correcteur : @corrector	**Note sur 100 :** @percent
Nom de l'étudiant : @student	**Note sur 30 :** @relative

Critères	Entrée 1	Entrée 2	Entrée 3	Entrée 4	Entrée 5	Entrée 6	Entrée 7	Entrée 8	Entrée 9	Entrée 10
Contenu et clarté	@q11	@q21	@q31	@q41	@q51	@q61	@q71	@q81	@q91	@q101
Vocabulaire	@q12	@q22	@q32	@q42	@q52	@q62	@q72	@q82	@q92	@q102
Grammaire - Verbes	@q13	@q23	@q33	@q43	@q53	@q63	@q73	@q83	@q93	@q103
Grammaire - Autres	@q14	@q24	@q34	@q44	@q54	@q64	@q74	@q84	@q94	@q104
Réflexion	@q15	@q25	@q35	@q45	@q55	@q65	@q75	@q85	@q95	@q105

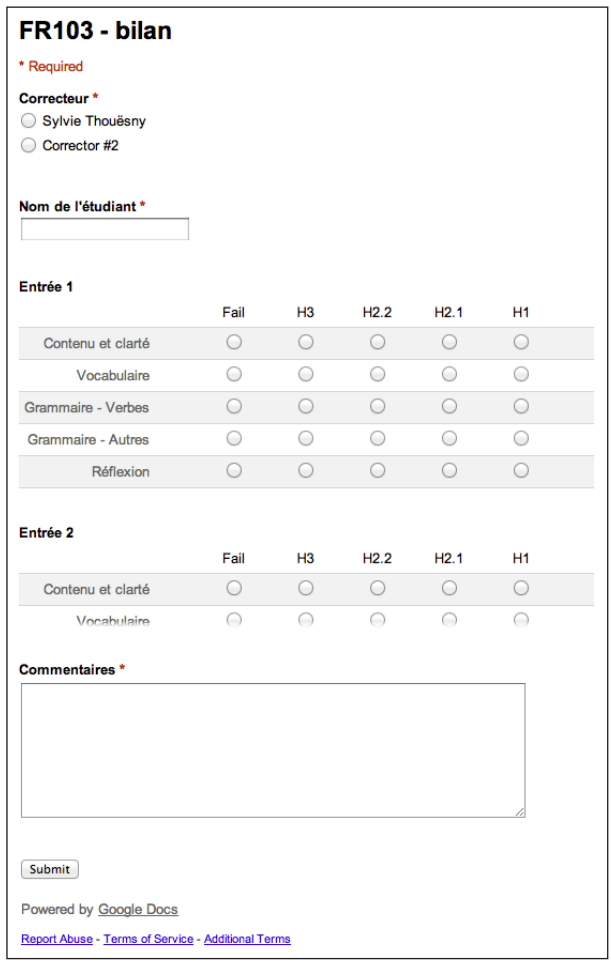
Commentaire :

@comments

* Marks and Standards. (2012). Retrieved from: http://www4.dcu.ie/sites/default/files/registry/pdfs/M-S_version-2012.1.0.pdf

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



FR103 - bilan

* Required

Correcteur *

- Sylvie Thouësny
- Corrector #2

Nom de l'étudiant *

Entrée 1

	Fail	H3	H2.2	H2.1	H1
Contenu et clarté					
Vocabulaire					
Grammaire - Verbes					
Grammaire - Autres					
Réflexion					

Entrée 2

	Fail	H3	H2.2	H2.1	H1
Contenu et clarté					
Vocabulaire					

Commentaires *

Submit

Powered by Google Docs

Report Abuse - Terms of Service - Additional Terms

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

Figure 3. Script

```
//insert information into template from spreadsheet
copy_body.replaceText("@comments", comments);
copy_body.replaceText("@corrector", corrector);
copy_body.replaceText("@student", student);

//retrieve results per question
var sum = 0;//to compute average
var count = 0;//to compute average
var question_number = 1;
for (var i = 0; i <= index_comment; i++) {
  if (data[0][i] in levels || data[0][i] == ""){
    var index = 0;
    for (var j = 1; j <= amount_of_controls; j++) {
      //replace text in template by result e.g. "Q11" => "H2.2" or 0 if question not answered
      if(data[0][i+index] == "" && question_number<=entries_required){
        copy_body.replaceText("@q"+question_number+j, 0);
        count++;
      }
      else if(data[0][i+index] == "" && question_number>entries_required) {
        copy_body.replaceText("@q"+question_number+j, "-");
      }
      else {
        copy_body.replaceText("@q"+question_number+j, data[0][i+index]);
        sum = sum + levels[data[0][i+index]];
        count++;
      }
      index++;
    }
    question_number++;
    i=i+index-1;
  }
}
```

The result was as follows (Figure 4):

Figure 4. Result

2011/2012
Barème de notation et résultats
Please note that these results are provisional and will only be confirmed after the examination board meeting.

Date de la correction : 21 Dec 2011

Nom du correcteur : Sylvie Thouësny	**Note sur 100 :** 59
Nom de l'étudiant : Jane Doe	**Note sur 30 :** 17.7

Critères	Entrée 1	Entrée 2	Entrée 3	Entrée 4	Entrée 5	Entrée 6	Entrée 7	Entrée 8	Entrée 9	Entrée 10
Contenu et clarté	H2.1	H1	H1	H2.1	H2.1	H1	H1	H2.1	-	-
Vocabulaire	H2.2	H1	H2.1	H2.2	H2.2	H2.1	H2.2	H2.2	-	-
Grammaire - Verbes	H2.2	H2.1	H2.1	H2.2	H3	H2.1	H2.2	H2.2	-	-
Grammaire - Autres	H2.2	H2.1	H2.1	H2.2	H2.2	H2.1	H2.2	H2.2	-	-
Réflexion	H2.2	H3	H2.2	Fail	H2.2	H3	H2.2	H3	-	-

Commentaire :

La présentation est agréable et les références sont mentionnées. Quelques fautes d'accord et de genre. Le vocabulaire est généralement bien choisi. Les activités sont intéressantes. La réflexion sur l'apprentissage de la langue en autonomie est parfois maladroite et superficielle. Travail satisfaisant.

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

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.

Behizadeh, N., & Engelhard, J. G. (2011). Historical view of the influences of measurement and writing theories on the practice of writing assessment in the United States. A*ssessing Writing, 16*(3), 189-211. doi:10.1016/j.asw.2011.03.001

Denton, P., Madden, J., Roberts, M., & Rowe, P. (2008). Students' response to traditional and computer-assisted formative feedback: A comparative case study. *British Journal of Educational Technology, 39*(3), 486-500. doi:10.1111/j.1467-8535.2007.00745.x

Knoch, U. (2011). Rating scales for diagnostic assessment of writing: what should they look like and where should the criteria come from? *Assessing Writing, 16*(2), 81-96. doi:10.1016/j.asw.2011.02.003

Kulhavy, R. W. (1977). Feedback in written instruction. *Review of Educational Research, 47*(2), 211-232. Retrieved from http://www.jstor.org/stable/1170128

Penny, J. A., & Johnson, R. L. (2011). The accuracy of performance task scores after resolution of rater disagreement: a Monte Carlo study. *Assessing Writing, 16*(4), 221-236. doi:10.1016/j.asw.2011.06.001

Wardle, E., & Roozen, K. (2012). Addressing the complexity of writing development: toward an ecological model of assessment. *Assessing Writing, 17*(2), 106-119. doi:10.1016/j.asw.2012.01.001

Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

Yancey, K. B. (1999). Looking back as we look forward: historicizing writing assessment. *College Composition and Communication, 50*(3), 483-503. Retrieved from http://www.jstor.org/stable/358862

Zimmaro, D. M. (2007). *Using rubrics to grade student performance*. Austin: University of Texas at Austin, Centre for teaching and learning. Retrieved from http://ctl.utexas.edu/assets/Evaluation--Assessment/Using-Rubrics-to-Grade-Student-Performance-10-15-07.pdf

Published by Research-publishing.net
Dublin, Ireland; Voillans, France
info@research-publishing.net

CALL: Using, Learning, Knowing
EUROCALL Conference, Gothenburg, Sweden
22-25 August 2012, Proceedings
Edited by Linda Bradley and Sylvie Thouësny

Typeset by Research-publishing.net
Cover design: © Raphaël Savina (raphael@savina.net)
Aquarelle reproduced with kind permission from the illustrator: © Sylvi Vigmo (sylvi.vigmo@ped.gu.se)
Fonts used are licensed under a SIL Open Font License

ISBN13: 978-1-908416-03-2 (paperback)
Print on demand (lulu.com)

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Bibliothèque Nationale de France - Dépôt légal: décembre 2012.