Oral-based rubrics design: A case study with undergraduate Spanish students in ESP settings

Carolina Girón-García,
Claudia Llopis-Moreno
Universitat Jaume I, Castellón de la Plana, Spain

Abstract

During several decades, considerable attention has been paid to speaking assessment procedures. Probably, the reason for this relevance lies in the difficulties in dealing with oral assessment and in determining the adequate type of assessment (Campbell, Mothersbaugh, Brammer & Taylor, 2001; Stoynoff, 2013; Schwartz & Arena, 2013). The main aim of this article is to analyze to what extent students' oral competences (such as their English competence and fluency) affect their peers' English oral production. More precisely, we intend to study effective procedures to assess oral production in ESP contexts. In order to assess our students, we have designed a speaking-based rubric as the main instrument (called 'Speaking Diagnostic Test'), based on previous research (Wilson, 2006; Spandel, 2006). A total number of 10 participants were selected in the 1st year course at Universitat Jaume I (Spain). In order to carry out our investigation two different instruments were used for data collection: (1) A 'Student Questionnaire' to reflect students' preferences towards taking an oral exam with their peers; and a (2) 'Speaking Diagnostic Test' so as to assess students' oral production in English concerning the following variables: fluency, vocabulary, grammar, pronunciation, coherence, and communicative ability. The results obtained from both questionnaires will help us identify to what extent students' preferences in working with other classmates influence their English oral production and therefore to what extent these results could lead to the reformulation and modification of the assessment methods.

Keywords: ESP contexts, rubric, 'Speaking Diagnostic Test', oral production

Resum

Durant varies dècades s'ha prestat molta atenció als procediments d'avaluació d'exàmens orals. Possiblement, la raó estigi deguda a les dificultats que existeixen al tractar el tema de l'avaluació d'exàmens orals i el tipus d'avaluació apropiat. (Campbell et al., 2001; Stoynoff, 2013; Schwartz & Arena, 2013). El principal objectiu d'aquest article és analitzar fins a quin punt les competències orals del estudiants (com la seua competència en llengua anglesa i la fluïdesa) afecten a la producció oral en anglès del seus companys. Més concretament, el que pretenem és

estudiar procediments efectius per avaluar la producció oral en contextos ESP. Per avaluar els nostres estudiants, hem dissenyat una rúbrica per avaluar competències orals com instrument principal (anomenada "Test de diagnòstic d'orals"), basada en investigacions prèvies (Wilson, 2006; Spandel, 2006). Un total de 10 participants van ser seleccionats en el primer curs acadèmic de la "Universitat Jaume I" (Espanya). Per portar a terme la nostra investigació dos instruments foren utilitzats: (1) Un "Qüestionari del Estudiant" per reflectir les preferències dels estudiants cap a la realització d'un examen oral amb els seus companys, i (2) Un "Test de Diagnòstic Oral" per avaluar la producció oral en anglès dels estudiants en referència a les següents variables: fluïdesa, vocabulari, gramàtica, pronunciació, coherència, i habilitat comunicativa. Els resultats obtinguts d'ambdós qüestionaris ens ajudaran a identificar fins a quin punt les preferències dels estudiants respecte de treballar amb altres companys influeixen la seua producció oral en anglès i per tant fins a quin punt aquests resultats podrien suposar una reformulació i modificació dels mètodes d'avaluació.

Paraules clau: Contextos ESP, rúbrica, "Test de Diagnòstic Oral", producció oral

Resumen
Durante varias décadas se ha prestado mucha atención a los procedimientos de evaluación de exámenes orales. Posiblemente, la razón de este énfasis radique en las dificultades que existen al tratar el tema de la evaluación de exámenes orales y en el tipo apropiado de evaluación (Campbell et al., 2001; Stoynoff, 2013; Schwartz & Arena, 2013). El principal objetivo de este artículo es analizar hasta qué punto las competencias orales de los estudiantes (tales como su competencia en lengua inglesa y su fluidez) afectan la producción oral en inglés de sus compañeros. Más concretamente, pretendemos estudiar procedimientos efectivos para evaluar la producción oral en contextos de ESP. Para evaluar a nuestros estudiantes, hemos diseñado una rúbrica para evaluar competencias orales como instrumento principal (llamada "Test de diagnóstico de orales"), basada en investigaciones previas (Wilson, 2006; Spandel, 2006). Un total de 10 participantes fueron seleccionados en el primer curso académico de la "Universitat Jaume I" (España). Para llevar a cabo nuestra investigación se utilizaron dos instrumentos distintos para la recogida de datos: (1) Un "Cuestionario del Estudiante" para reflejar las preferencias de los estudiantes hacia la realización de un examen oral con sus compañeros, y (2) Un "Test de Diagnóstico Oral" para evaluar la producción oral en inglés de los estudiantes en referencia a las siguientes variables: fluidez, vocabulario, gramática, pronunciación, coherencia, y habilidad comunicativa. Los resultados obtenidos de ambos cuestionarios nos ayudará a identificar hasta qué punto las preferencias de los estudiantes hacia trabajar con otros compañeros influencian su producción oral en inglés y por tanto hasta qué punto estos resultados podrían conducirnos a una reformulación y modificación de los métodos de evaluación.

Palabras clave: Contextos ESP, rúbrica, "Test de Diagnóstico Oral", producción oral.

Introduction

In recent years, teachers of English as a Second Language (ESL) have paid particular attention to the relevance of communicative competence as an essential aspect in their English language courses. Canale and Swain (1980) and Canale (1983) defined the term of *communicative competence* as a combination of language knowledge and social knowledge for communication. The former was referred to as knowledge on how to use the language in terms of syntax, morphology, and philosophy; whereas the latter implied the use of such knowledge for actual communication. As a result, ESL teachers have generally agreed on the idea of teaching students how to be communicatively competent in different social contexts and situations as the main goal of their English courses. Moreover, it is important to highlight students' perception on an existing gap between the relevance given to oral performance and the limited time devoted to develop it in the university context (Kavanagh & Drennan, 2008).

In this study, we will be paying special attention to the use of rubrics as an effective oral assessment tool in English for Specific Purposes (ESP) contexts since they have been commonly used during recent years in Higher Education English courses. Particularly, we will be analyzing to what extent students' oral competences affect their peers' English oral production. With peer interaction in mind, in our research we intend to study effective procedures to assess oral production in ESP contexts.

This article begins by outlining the development of Communicative Approach (CA) and how it has influenced not only the teaching of English as a Foreign Language but also the way in which language competences are assessed. Then the article describes the design of the rubric which served as both assessment tool and as object of study concerning its efficacy. The next section is devoted to the study. This includes the context and participants who took part in the experiment; and procedure and data collection, and instruments employed. Next, the results and interpretation of the data obtained in this analysis are discussed.

Oral-based rubrics: Criteria to evaluate communicative competence

Communicative approaches have become highly significant, trespassing the barriers of rigid learning structures to a conception geared towards the promotion of communication in the classroom. At this point, we feel it is useful to look back into the development of the communicative approach in order to better understand how this approach is often seen going down the path of eclecticism, with ill-defined borders.

According to Richards & Rodgers' (1986, p. 72) principles in the Communicative Approach represented a fundamental contribution in the language learning field, since they support learning by means of introducing real-life tasks that involve communication and supply the learners' educational needs. These principles are the following:

1. The principle of communication (activities involving real communication that promote learning).

2. The principle of tasks (those activities that involve using the language to carry out significant tasks that promote learning)

3. The principle of significant relevance (language becomes meaningful for learners in order to positively support their learning process).

As for the first principle, discursive competence was one of the dimensions taken into consideration as an optimal tool to promote communicative competence by the first authors who launched the notion of Communicative Competence (Halliday, 1970; Widdowson, 1978; Canale & Swain, 1980).

With regards to the second principle, the notion of tasks became central in the Communicative Approach. The materials used for task design were taken from real-life resources and not manipulated at all for pedagogical purposes. The introduction of realia would become an important aspect to be taken into account, since realia are believed to be the most convenient kind of material to promote language learning due to their discursive and cultural context.

Concerning the third principle, the Communicative Approach places the learner at the centre of the learning process. Therefore, learners' needs and objectives are the basis to construct a flexible and dynamic syllabus. By employing a Communicative Approach, an interactive and meaningful relationship among learners' needs,

objectives and the syllabus can be achieved. This multiple relationship can increase learners' involvement in their own learning process.

Within this context and in order to fulfil the communicative standards provided by the Common European Framework of Reference for Languages (CEFRL), an assessment tool (i.e. rubrics) is needed. For this reason, 'rubrics' have become very popular in the field of language education and have caused a great impact among teachers and students as a new tool to evaluate oral production in English.

Despite the vast amount of definitions for the word 'rubric' found, for the purpose of this research we use the term 'rubric' in the same way that researchers Allen and Tanner (2006) already pointed out to refer to "a type of matrix that provides scaled levels of achievement or understanding for a set of criteria or dimensions of quality for a given type of performance…" (p. 197). Along these lines and concerning our study, our intention is to assess undergraduate students' oral performances with the design of a speaking-based rubric and then to analyze the efficacy of said rubrics.

In order to use the adequate type of rubric, we need to focus on a specific kind that establishes a link between a particular content and the objectives that account for a given subject matter. At this respect, Allen and Tanner's (2006) description of analytical and holistic rubrics may lead us to a quite open view of how our rubric might be categorized. Thus, according to these authors, analytical rubrics "use discrete criteria to set forth more than one measure of the levels of an

accomplishment for a particular task", whereas holistic rubrics are defined as those that "provide more general, uncategorized […] descriptions of overall dimensions of quality for different levels of mastery" (p.198).

Many educators have mentioned and suggested relevant communication features so as to assess speaking in general speaking courses (Jones, 1994; Rubin, Graham & Mignerey, 1990). Moreover, students' learning assessment has traditionally and basically been focused on written exams; in contrast to non-traditional assessment procedures, such as portfolios and oral assessments. In fact, many oral assessment procedures require the use of a rubric; it is probably for that reason that non-traditional assessment is on the rise of many educators concerns at all educational levels, especially at universities.

Evaluators spend most of their time listening to students' speeches, and then discussing their oral assessment regarding each of the competencies reflected in a speaking rubric. In this respect, evaluators usually share the same points of view concerning what constitutes the features of communicative competence, but on some occasions they do not agree on the same students' score. For this reason, the use of rubrics has become a very influential tool for assessment procedures and results with regards to maintaining consistency among teachers (Dunbar, Brooks, & Kubicka-Miller, 2006).

Designing a Rubric for ESP

When designing the rubric, we first asked ourselves the following question: "How should educators use a rubric in ESP contexts with undergraduate students?" With this question in mind, we began to design our *'Speaking Diagnostic Test'* in ESP contexts (see Appendix 1 for the full rubric). With regards to the required level (B1) for the subject under investigation ('Scientific English: Modern Language', code EX 1005) and as the Council of Europe (2001: 24) describes in the CEFRL, the following descriptors have were used as a basis for our rubric design.

- Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc.
- Can deal with most situations likely to arise while traveling in an area where the language is spoken.
- Can produce simple connected text on topics that are familiar or of personal interest.
- Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.

These descriptors were then developed into more detailed components of the language dimensions we aimed at evaluating, including: Fluency, Vocabulary, Grammar, Pronunciation, Coherence, and Communicative Ability (how these are used in the rubric is illustrated in the extract taken from the rubric in figure 1). This last

component (communicative ability) has been taken into consideration to foster communicative competence among students, since we are interested in enhancing communication in the classroom. Next, the horizontal section of the rubric included the descriptors organized according to their rate: (1) Weak, (2) Good, and (3) Excellent (figure 1).

	1	2	3	COMMENTS
COHERENCE	The message is incoherent and difficult to understand. Does not use connectors and linkers.	The message is mostly coherent. Uses few linkers and connectors.	The message is coherent. Uses suitable linkers and connectors. Correct content organisation.	
	1	**2**	**3**	**COMMENTS**
COMMUNICATIVE ABILITY	Does not adapt to other speakers' levels. Does not use turn taking and interrupting techniques. Sometimes, he /she remains silent until the other speaker finishes.	Mostly tries to adapt to other speakers' levels. Few attempts and mistakes when using turn taking and interrupting techniques.	Flexibility to speakers of different levels. Uses turn taking and interrupting techniques effectively. Helps other students when they are stuck in the conversation.	

Figure 1. Sample of part of the rubric

The rubric was then applied to the following activities: students were asked to perform both a monologue and two dialogues. In the monologue students performed they had to develop the topic *"Why do you think studying English is important?"*. This was followed by a performance two dialogues (role-plays) with two different classmates (regarding their preferences in working with others, as they stated in the 'Student Questionnaire') on the topic of *"The language of*

socializing: A night at the opera". The rubric was implemented and modified so as to include specific grammar and vocabulary in relation to these topics. Both communicative tasks were designed bearing in mind the importance of performing real-life situations

Finally, it is important to emphasise that the same 'Speaking Diagnostic Test' was employed in the three tasks (i.e. monologue and two dialogues). Accordingly, by using the same rubric we aimed to obtain relevant data from students' performances in different tasks while evaluating the same components, descriptors, and rates. In other words, we were interested in establishing a contrast between students' results in the two different communicative tasks under the same assessment criteria, but under different circumstances. The results obtained point at discerning the validity of this rubric as an effective tool for assessment.

The study

Context and participants

The aim of this study consists of the analysis of students' features (such as their English competence and fluency) in order to find out to what extent those features affect their peers' English oral production. The study was carried out in four stages:

1. Elaboration and design of a 'Student Questionnaire' designed to find out a classmate they would like to take an oral exam with and another classmate they would not like to.

2. Elaboration and design of a 'Speaking Diagnostic Test'

3. Realization of a monologue on the topic *"Why do you think studying English is important?"*

4. Realization of two dialogues (role-plays) on the topic *"The language of socializing: A night at the opera".*

4.1.   Students perform the role-play with a classmate who they would like to take an oral exam with their 'Ideal Partner', (shortened as 'IDP').

4.2.   Students perform the role-play with a classmate who they would not like to take an oral exam with ('Unwanted Partner', shortened as 'UNP').

Concerning the students' participation, 10 subjects were selected in the 1st year course at 'Universitat Jaume I' (Spain). Although they gave their permission to use their data, an individual identification code was provided in order to safeguard their privacy.

Procedure and data collection

The study took place during the students' regular class time in the first semester of the 2013-2014 academic course. The module is compulsory for all the students; and their participation in this study, although not mandatory, was presented as one of the complementary activities to be given partial credit in addition to the final mark at the end of the semester. In order for results to be available, students had to fill out a questionnaire ('Student Questionnaire') to be able to observe what exactly the students had answered.

Instruments

The main instrument for our study has been (a) a speaking-based rubric, also called 'Speaking Diagnostic Test' (see appendix 1). Nevertheless, we have used other instruments endorsed for data collection: (b) 'Student Questionnaire', and (c) 'Voice Recorder'.

'Student Questionnaire'

In order to determine the students' profile and information about their attitude towards learning English, a questionnaire was used, namely, 'Student Questionnaire' Appendix 2). This questionnaire elicits students' name and surname, age, gender, and two simple questions addressed to the students in order to get information about their preferences towards taking an oral exam with their classmates giving proper justifications.

Question 1 (Q1): Asks students to give a name and a reason for choosing a classmate they would like to take an oral exam with ('Ideal Partner').

Question 2 (Q2): Asks students to give a name and a reason for choosing a classmate they would not like to take an oral exam with ('Unwanted Partner').

These questions will help us classify students in pairs regarding their preferences towards taking an oral exam with their peers.

'Speaking Diagnostic Test'

Students' oral production was measured by means of a *'Speaking Diagnostic Test'* that was previously designed considering the following categories: Fluency (Fl.), Vocabulary (Voc.), Grammar

(Gr.), Pronunciation (Pr.), Coherence (Co.), and Communicative ability (Comm. ability). The score for each category could vary from 1 to 3, 3 being the highest score that could be obtained in each category. Thus, students could have a maximum of 18 points in this rubric, in case that their oral production was perfectly performed in the task.

All the variables included in this test are considered as relevant (as in conventional rubrics) in order to assign a certain score in each of them to every student.

The 'Speaking Diagnostic Test' was used in two different stages of our study. Firstly, the rubric was used in order to assess a speaking monologue task in which students were asked the following question: *'Why do you think studying English is important?'*. Secondly, it was used in a different task in which students had to perform a role-play that consisted of inviting a peer to go to the opera. Hence, the rubric was used in two different tasks with the aim to analyse to what extend students' oral production may be positively or negatively affected by their oral interaction with other peers. In order to analyse such a possible existing influence, we compared students' scores obtained in the rubric in their individual task (monologue) and in the interaction tasks (role-plays).

'Voice Recorder'

Students' oral production in the different tasks of our analysis was recorded by means of a 'voice recorder', in which all students' monologues and dialogues performances were kept for the purpose of

data collection in our study. In addition, this tool helped us to review students' performance in terms of vocabulary, grammar, communicative ability, pronunciation, and so forth. Therefore, this instrument has been useful in order to check and compare the rubric's scores obtained by each student, so as to support their assessment result.

Results and discussion: Students' scores

The results reveal the different elements or variables included in that rubric, taking into account in both the monologue and dialogues performances on behalf of the students. The 'Speaking Diagnostic Test' shows the different variables that we considered in order assigning a certain score from 1 to 3 in each of them. The total score regarding all these variables is 18 points. Therefore, this test is our main instrument in order to get relevant data concerning students' results in the two tasks they performed: (a) Monologue employed as a diagnostic placement test; and (b) Two dialogues (Role-plays) with classmates who wanted to work together ('Ideal Partners') and classmates who did not want to (Unwanted Partners').

Students' scores in the monologue

As for the students' scores obtained in the monologue, the table below (Table 1) reflects the total amount of students who participated in the experiment, as well as the scores obtained in each of the variables and final scores.

Table 1: Students' scores in the monologue

| STUDENTS | MONOLOGUE | | | | | | |
| | VARIABLES & SCORES | | | | | FINAL SCORE |
	Fl.	Voc.	Gr.	Pr.	Co.	Comm. ability	
A	1	1	1	1	1	1	6/18
B	1	1	1	1	1	3	8/18
C	1	1	1	1	1	1	6/18
D	2	1	2	1	1	2	9/18
E	1	1	1	1	1	1	6/18
F	2	1	1	1	1	2	8/18
G	1	1	1	1	1	1	6/18
H	2	1	1	1	1	2	8/18
I	1	1	1	1	2	3	9/18
J	1	1	1	1	1	1	6/18

Due to research purposes, we are not going to analyse in full detail the results concerning the students' scores in the monologue, but we will proceed to comment on them in general terms since these results are the product of an initial proficiency level task ('Speaking Diagnostic Test').

This table provides indicative results, because of 10 students that are indicated in the table; only 2 students achieved an average result (9/18). These results seem to confirm that the students' initial English proficiency level did not reach our desirable expectations.

Students' scores in the dialogues

Concerning the two dialogues (role-plays), it is worth highlighting that in order to obtain relevant results, students indicated their preferences towards taking an oral exam with their peers in a 'Student Questionnaire' that was designed for the purpose of this research. Thus, according to their preferences ('IDP'/'UNP'), the following pairs were selected:

- 'Ideal Partner' ('IDP'): 'Student A' and 'Student E', 'Student B' and 'Student C', 'Student C' and 'Student F', 'Student D' and 'Student I', 'Student E' and 'Student D', 'Student F' and 'Student C', Student G' and 'Student B', 'Student H' and 'Student D', 'Student I' and 'Student E', and 'Student J' and 'Student D'.

- 'Unwanted Partner' ('UNP'): 'Student A' and 'Student I', 'Student B' and 'Student A', 'Student C' and 'Student A', 'Student D' and 'Student A', 'Student E' and 'Student F', 'Student F' and 'Student E', 'Student G' and 'Student H', 'Student H' and 'Student G', 'Student I' and 'Student D', and 'Student J' and 'Student A'.

Regarding Table 2 (Students' scores in the role-plays – 'Ideal Partners'), the 10 pairs who wanted to take an oral exam together ('IDP') are shown, providing their score in each of the variables as well as their final score.

Table 2: Students' scores in the role-plays – 'Ideal partners'

DIALOGUES (ROLE-PLAYS)		
	IDEAL PARTNERS	FINAL
PAIRS	VARIABLES & SCORES	SCORE

(Students)	Fl.	Voc.	Gr.	Pr.	Co.	Comm. ability	
A / E	1/2	1/2	1/2	1/2	1/2	1/2	A (6/18) E (12/18)
B / C	1/2	1/2	1/2	2/2	1/2	1/2	B (7/18) C (12/18)
C / F	2/1	2/1	2/1	2/1	2/1	3/3	C (13/18) F (8/18)
D / I	2/2	2/1	2/2	2/1	2/2	3/2	D (13/18) I (10/18)
E / I	2/2	2/1	2/1	2/1	2/1	2/2	E (12/18) I (8/18)
F / C	1/2	1/2	1/2	1/2	1/2	3/3	F (8/18) C (13/18)
G / B	2/2	1/2	1/2	1/1	1/1	3/3	G (9/18) B (11/18)
H / D	2/2	2/2	2/2	2/2	2/2	3/3	H (13/18) D (13/18)
I / E	2/2	1/2	1/2	1/2	1/2	2/2	I (8/18) E (12/18)
J / D	1/2	1/2	1/2	1/2	2/2	1/3	J (7/18) D (13/18)

From these results we may say that only 5 individuals did not reach the average score (9/18) with regards to the final score (18 points). Thus, some examples such as students 'A', 'B', 'F', 'I', and 'J' do not reach the average class level. Accordingly, we will now comment these students in full detail:

'Student A' does not show fluent speech, which is reflected in long pauses and breaks. Vocabulary expressions and grammar use is rather poor. Furthermore, he does not employ specific vocabulary demanded in the subject. He pronounces words incorrectly, but he tries to be coherent. However, he does not adapt to his partner's level,

which shows a great communicative ability in English. Finally, this student obtains 6/18 points in his final score.

'Student B' does not speak fluently. He uses long pauses, hesitations and long breaks. Furthermore, his vocabulary and grammar is very poor and does not use the specific topic vocabulary. The message he conveys in incoherent and difficult to understand, because he does not use linkers and connectors. Finally, his final score (7/18) is justified.

'Student F' does not speak fluently either and long pauses and hesitations are present during his speech. The vocabulary and expressions used are very poor, as well as poor grammatical structures with significant mistakes. Regarding pronunciation, most of the words he uses are incorrectly pronounced. Furthermore, the message he tries to transmit is difficult to understand, because of his lack in the use of connectors. Taking into account all these comments, we could certify that he obtained 8/18 in the final score.

'Student I' speaks mostly fluently, although he uses poor vocabulary and grammar. He does not try to employ the specific technical vocabulary and expressions required in this subject. In addition, he pronounces most words incorrectly with non-existing intonation or word stress. Finally, the message he conveys is completely incoherent due to the lack of connectors; however, he tries to adapt to his partner's level through the use of turn taking and interrupting techniques. These data confirm his final score obtained (8/18).

'Student J' does not show fluent speech, which is reflected in long pauses, hesitations, and long breaks. Furthermore, vocabulary expressions and grammar use is poor. He does not employ the specific vocabulary demanded in the subject. Although he pronounces words incorrectly, he makes an effort in transmitting a coherent discourse and be understandable. However, he does not adapt to his partner's level, which shows a great communicative ability in English. Finally, for these reasons this student obtains 7/18 points in his final score.

On the other hand, students 'C' (13/18), 'D' (13/18), 'E' (12/18), 'G' (9/18), and 'H' (13/18) were perceived as having obtained the highest scores in this role-play task, in which they could express their preferences when choosing their pairs. The five of them showed better oral skills than the rest of the subjects taking part in the same task. They spoke mostly fluently with a good command of vocabulary (using appropriate topic vocabulary); they used basic grammatical structures, although they still made some key mistakes. Furthermore, they were able to pronounce mostly all words correctly with correct intonation and word stress, so as their messages were mostly coherent but with few linking words. Finally, their communication ability was excellent, as these four students used turn taking and interrupting techniques very effectively.

As table 3 illustrates, the students who obtained lower scores in this case were 'A' (6/18), 'B' (6/18), 'C' (8/18), 'F' (8/18), 'G' (6/18), and 'J' (6/18). The six of them were below the average (9/18) when performing the task, as they were not able to speak fluently, because

they needed time to plan what to say using long pauses and hesitations. The vocabulary and grammatical structures they employed were very poor and contained many significant mistakes. In addition, they made many pronunciation mistakes with non-existing intonation and word stress in their production. In general, their messages were incomplete and incoherent in most cases, as they did not use connectors. And finally, their communicative ability was very poor too, since they were not able to use turn taking or interrupting techniques.

Table 3: Students' scores in the role-plays – 'Unwanted Partners'

DIALOGUES (ROLE-PLAYS)							
UNWANTED PARTNERS							**FINAL SCORE**
VARIABLES & SCORES							
PAIRS (Students)	**Fl.**	**Voc.**	**Gr.**	**Pr.**	**Co.**	**Comm. ability**	
A / I	1/2	1/1	1/2	1/1	1/2	1/2	A (6/18) I (10/18)
B / A	1/1	1/1	1/1	1/1	1/1	1/1	B (6/18) A (6/18)
C / A	1/1	1/1	1/1	2/1	1/1	2/1	C (8/18) A (6/18)
D / A	2/1	2/1	2/1	2/1	2/1	1/1	D (11/18) A (6/18)
E / F	3/1	3/1	3/1	2/2	2/1	2/2	E (15/18) F (8/18)
F / E	1/3	1/3	1/3	2/2	1/2	2/2	F (8/18) E (15/18)
G / H	1/2	1/2	1/1	1/1	1/2	1/3	G (6/18) H (11/18)
H / G	2/1	2/1	1/1	1/1	2/1	3/1	H (11/18) G (6/18)
I / D	2/2	1/2	2/2	1/2	2/2	2/3	I (10/18)

						D (13/18)
J / A	1/1	1/1	1/1	1/1	1/1	1/1	J (6/18) A (6/18)

In the light of all the results shown at the beginning of this section, further discussion is presented taking the following Research Question as the starting point: *(RQ) "To what extent students' preferences in working with other classmates influence their English oral production?"*

The analysis of the data and the information obtained in tables 2 and 3 allows us to compare the scores obtained in both role-plays ('IDP' and 'UNP'). In addition, these data could lead us to suggest the following concepts: 'Higher Performance', 'Invariable Performance', and 'Lower Performance', since these terms could present a certain degree of ambiguity if we take into consideration the variability of assessment criteria. Let us recall that the rubric employed for the assessment of both role-plays included a total score of 18 points. Thus, we refer to 'Higher Performance' in the case of students who improved their performance results when interacting with a classmate they chose in the 'Student Questionnaire'. We understand as 'Invariable Performance' in the case of students whose oral performance did not imply any change in their final score. And 'Lower Performance', in the case of students whose final scores dropped.

Comparison of Data Sets

Figure 2, below, illustrates the final percentages obtained with regards to the comparison between students who wanted to take an oral exam together ('IDP') and those who did not want to interact with certain classmates ('UNP'), to what their preferences in the 'Student Questionnaire' are concerned.
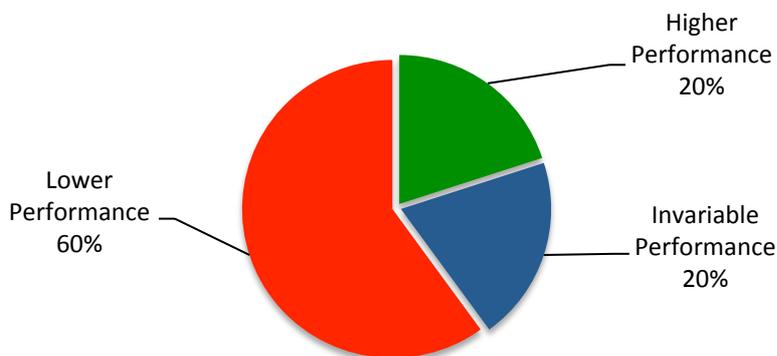

Figure 2: Percentage of results with 'Unwanted Partners'

These percentages have been calculated comparing the students' final scores in the rubric (up to 18 points) in tables 2 and 3. Our initial hypothesis was that students who preferred working together would get higher results than those who did not want to. For this reason, we have only focused on the evolution of students comparing their results obtained in the 'IDP' role-play with regards to their results in the 'UNP' role-play.

　　　Thus, taking into account the previous classification ('Higher', 'Invariable', and 'Lower Performance') and the percentages shown in Figure 3, we can observe the following:

- 'Higher Performance': Students 'E' and 'I' obtained higher scores in the 'UNP' role-play than their scores obtained in the 'IDP' role-play. Student 'E' got 12/18 points ('IDP' role-play) and 15/18 points ('UNP' role-play). Student 'I' got 8/18 ('IDP' role-play) and 10/18 ('UNP' role-play).

- 'Invariable Performance': Students 'A' and 'F' remained the same in both role-plays results. Student 'A' got 6/18 and student 'F' got 8/18.

- 'Lower Performance': Students 'B', 'C', 'D', 'G', 'H', and 'J', obtained lower scores. Student 'B' got 7/18 ('IDP' role-play) and 6/18 ('UNP' role-play). Student 'C' got 13/18 ('IDP' role-play) and 8/18 ('UNP' role-play). Student 'D' got 13/18 ('IDP' role-play) and 11/18 ('UNP' role-play). Student 'G' got 9/18 ('IDP' role-play) and 6/18 ('UNP' role-play). Student 'H' got 13/18 ('IDP' role-play) and 11/18 ('UNP' role-play). Finally, student 'J' got 7/18 ('IDP' role-play) and 6/18 ('UNP' role-play).

All these results considered are illustrated in the following table:

Table 4: Students' scores in the role-plays – 'Ideal Partners' and 'Unwanted Partners'

ROLE-PLAYS		
STUDENTS	'IDEAL PARTNERS'	'UNWANTED PARTNERS'
A	6/18	6/18
B	7/18	6/18

C	13/18	8/18
D	13/18	11/18
E	12/18	15/18
F	8/18	8/18
G	9/18	6/18
H	13/18	11/18
I	8/18	10/18
J	7/18	6/18

To conclude, looking at the percentages of the 10 students who participated in this study concerning learners who performed the tasks with someone they did not initially want to interact with, we can observe that 60% had a lower performance while 20% had an invariable performance and another 20% had a higher performance.

Conclusion

The present work focuses on analysing to what extent students' oral competences (pronunciation, fluency, grammar, etc.) are affected by interaction with their peers, in relation to their peer's perceived level of English oral production. Furthermore, we intended to study effective procedures to assess oral production in ESP contexts.

We have seen that our study was based on the design of a rubric ('Speaking Diagnostic Test') that was employed to test students' initial performances in English. This rubric has been useful to assess students' performances in the two role-plays. This instrument (i.e. rubric) has also been relevant for several principles that imply

learners' communication in order to supply their educational needs: (a) Activities that involve real communication to promote oral skills in English (Principle of communication); (b) Activities that imply employing the English language to develop tasks that are considered as relevant to promote learning (Principle of tasks); and (c) Learning process support through meaningful language (Principle of significant relevance). According to the CEFRL, all these principles have been proved to be necessary within the field of language education, and thus, rubrics have become very popular as an instrument for oral production assessment.

Comparing the scores obtained in both role-plays we could conclude that our initial hypothesis has been supported by such scores: when students interacted with classmates they had positively chosen in the 'Student Questionnaire', they performed significantly better than when they interacted with those ones they did not want to.

Our data have shown that there is a tendency towards a 'Lower Performance' (60%) when students perform a role-play with 'Unwanted Partners'. In contrast, we have observed that our initial hypothesis about a 'Higher Performance' (20%) is reflected in students who wanted to work together. Nevertheless, some students' oral performances remained the same ('Invariable Performance') final score without being affected by their partners (20%).

While the present study has attempted to examine to what extent students' preferences towards taking an oral exam with their peers influence their oral production, its results and further

conclusions must be taken into consideration in order to reformulate and modify the assessment methods. However, this study has some limitations, and the recognition of these should help refine future research efforts. Firstly, it is important to consider modifications and/or implementation of the different variables used in the rubric in order to focus more in depth on students' weaknesses. This fact may lead to reconsider students' needs in English ESP courses. Secondly, regarding the 'Student Questionnaire', we could include more specific questions and statements so as to obtain more precise information about the students' preferences. We could include a list of reasons why students would (not) choose a classmate, as we have found that students felt reluctant to provide such information. Thirdly, we concede that there have been a limited number of subjects in our study. Bigger groups of students were not available at the time, and with more students we could have more likely found more nuances. Finally, we think it is necessary to investigate not only on the use of rubrics on the part of the teacher ('Teacher assessment'), but also their use on behalf of students ('Peer assessment'). This last idea could contribute to the students' awareness on a series of limitations and aspects to consider when taking part in an oral exam in English for ESP contexts.

This study has only been a first step towards the design of more complex rubrics with bigger quantities of students as subjects. In the light of the results of our study, we view a quite open field to further explore the design and use of rubrics in ESP contexts.

References

Allen, D., & Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE—Life Sciences Education*, 5, 197–203.

Campbell, K. S., Mothersbaugh, D. L., Brammer, C., & Taylor, T. (2001). Peer versus self-assessment of oral business presentation performance. *Business Communication Quarterly, 64*(3), 23-42. http://dx.doi.org/10.1177/108056990106400303

Canale, M. (1983). From communicative competence to communicative language pedagogy. In Richards, J. C., & Schmidt, R. W. (Eds.), *Language and communication* (pp. 2-27). London: Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47. http://dx.doi.org/10.1093/applin/i.1.1

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Strasbourg: Language Policy Unit. Retrieved 8 September 2015 from http://www.coe.int/t/dg4/linguistic/source/framework_en.pdf

Dunbar, N.E., Brooks, C.F., & Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31 (2), 115-128. http://dx.doi.org/10.1007/s10755-006-9012-x

Halliday, M. A. K. (1970). Language structure and language function. In Lyons, J. (ed.) *New horizons in linguistics* (pp. 140-165). Harmondsworth: Penguin.

Jones, E. (1994). *Essential skills in writing, speech and listening, and critical thinking for college graduates: Perspectives of faculty, employers, and policymakers*. University Park, PA: National Center for post-secondary Teaching, Learning, and Assessment.

Kavanagh, M., & Drennan, L. (2008). What skills and attributes does an accounting graduate need? Evidence from student perceptions and employer expectations. *Accounting and Finance*, 48 (2), 279-300. http://dx.doi.org/10.1111/j.1467-629x.2007.00245.x

Richards, J. C., & Rodgers, T.S. (1986). *Approaches and methods in language teaching: A description and analysis*. Cambridge: Cambridge University Press.

Rubin, R., Graham, E., & Mignerey, J. (1990). A longitudinal study of college students' communication competence. *Communication Education*, 39, 1-14. http://dx.doi.org/10.1080/03634529009378783

Schwartz, D., & Arena, D. (2013). *Measuring what matters most. Choice-based assessments for the digital age.* Cambridge, Massachusetts. London, The MIT Press.

Spandel, V. (2006). Speaking my mind. In defence of rubrics. *English Journal*, 96(1), 19-22. http://dx.doi.org/10.2307/30046656

Stoynoff, S. (2013). Looking backward and forward at classroom-based language assessment. *ELT Journal, 66*(4), 523-532. Oxford University Press. http://dx.doi.org/10.1093/elt/ccs041

Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.

Wilson, M. (2006). *Rethinking rubrics in writing assessment*. Portsmouth, NH: Heinemann

Acknowledgements:

Appendix 1

SPEAKING DIAGNOSTIC TEST – EX 1005
(ROLE-PLAY INTERACTION)

NAME _____

GROUP _____

DATE _____ SCORE _____ /18

	1	2	3	COMMENTS

	1	2	3	
FLUENCY	Does not speak fluently. Uses long pauses, hesitations and long breaks	Speaks mostly fluently. Some pauses and hesitations.	Speaks very fluently. Few or non-existing hesitations and pauses.	
	1	**2**	**3**	**COMMENTS**
VOCABULARY	Uses poor vocabulary and expressions. Does not use or has problems using specific topic vocabulary	Uses basic vocabulary and expressions. Mostly uses some appropriate topic vocabulary	Uses an appropriate wide variety of specific vocabulary and expressions for the topic of the conversation	
	1	**2**	**3**	**COMMENTS**
GRAMMAR	Uses poor grammatical structures. Many significant mistakes.	Uses basic grammatical structures. Several key mistakes.	Uses accurate and appropriate grammatical structures. Very limited mistakes	
	1	**2**	**3**	**COMMENTS**
PRONUNCIATION	Pronounces most words incorrectly. Incorrect use or non-existing intonation and word stress.	Pronounces almost all words correctly. Mostly uses correct intonation and word stress.	Pronounces all words correctly. Uses correct intonation and word stress.	
	1	**2**	**3**	**COMMENTS**
COHERENCE	The message is incoherent and difficult to understand. Does not use connectors and linkers.	The message is mostly coherent. Uses few linkers and connectors.	The message is coherent. Uses suitable linkers and connectors. Correct content organisation.	
	1	**2**	**3**	**COMMENTS**
	Does not adapt to other speakers'	Mostly tries to adapt to other	Flexibility to speakers of different	

COMMUNICATIVE ABILITY	levels. Does not use turn taking and interrupting techniques. Sometimes, he /she remains silent until the other speaker finishes.	speakers' levels. Few attempts and mistakes when using turn taking and interrupting techniques.	levels. Uses turn taking and interrupting techniques effectively. Helps other students when they are stuck in the conversation.	

Appendix 2

Student Questionnaire

PERSONAL INFORMATION

Student's   name   and   surname: _____
Age: _____
Gender: M / F

MORE INFORMATION

Read the following questions carefully and answer them with honesty.
You can write as many names as you consider necessary.

1. What classmate(s) would you personally like to take an oral exam with? Justify your answer.	
Name and Surname	Reason
2. What classmate(s) would you not like to take an oral exam with? Justify your answer.	
Name and Surname	Reason

Author information:
Carolina Girón García is a Doctor in English Philology (2013). She has been working at Universitat Jaume since 2007. Her main lines of research are: language

teaching and learning, autonomous language learning using the new technologies, and more specifically Cybertask design. Finally, she is an active member of the 'Contrastive and Applied Semantics Research Group' (GRESCA), and participates in two research projects with the main objective of improving teaching methodologies at 'Universitat Jaume I' (in Castellón).

Claudia Llopis-Moreno holds an English Philology Degree and an ELIT Master (English Language for International Trade). After gaining some professional experience in secondary education, she has been teaching in Universitat Jaume I (UJI) since 2010 as an associate teacher. Her teaching experience has been devoted to ESP, teaching Technical English to undergraduate students in different engineering degrees as well as teaching different subjects (Phonetics, Literature, Writing and Psycholinguistics) in English Studies Degree. Her main research field is Discourse Analysis, and professional and academic English (ESP).

Email (principal author): gironc@uji.es