UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE LETRAS

ALEXANDRE LUÍS KOCH DA SILVA

THE USE OF CAMBRIDGE ESOL PARAMETERS IN ORAL PERFORMANCE ASSESSMENT IN AN ENGLISH LANGUAGE INSTITUTE

Porto Alegre

2012

ALEXANDRE LUÍS KOCH DA SILVA

THE USE OF CAMBRIDGE ESOL PARAMETERS IN ORAL PERFORMANCE ASSESSMENT IN AN ENGLISH LANGUAGE INSTITUTE

Monografia apresentada ao Curso de Letras da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção de Grau de Licenciado em Letras.

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

Porto Alegre

2012

DEDICATION

I dedicate this study to my wife Daniela and my daughter Amanda: everything I do is for you.

ACKNOWLEDGEMENTS

"It matters not how strait the gate,
How charged with punishments the scroll.
I am the master of my fate:
I am the captain of my soul."

William Ernest Henley

RESUMO

Este estudo tem por objetivo investigar a eficácia do uso dos critérios de The University of Cambridge English for Speakers of Other Languages (ESOL) Examinations para avaliação da fala em Língua Inglesa como segunda língua numa escola de inglês localizada no município de Canoas, RS. Os participantes deste estudo foram 4 estudantes e 5 professores da escola referida. Os dados para a realização deste estudo foram coletados através da avaliação pelos professores da fala em inglês de alunos que foram gravados em áudio participando de simulados de provas de proficiência de Cambridge e de questionários respondidos pelos professores participantes após a avaliação. Esta metodologia foi empregada visando a atingir os objetivos específicos propostos pela pesquisa: a) investigar se os professores da escola de inglês reconhecem e usam os critérios para avaliação da fala de Cambridge, que estão em uso na escola; b) examinar se existem discrepâncias entre professores quanto às notas dadas a cada estudante e os critérios usados; c) constatar se o fato de um professor ser ou não um avaliador certificado de Cambridge influencia o processo de avaliação. No que diz respeito aos resultados obtidos, verificou-se que alguns professores utilizam os critérios desejados e outros não. Comprovou-se também que existem algumas diferenças no que diz respeito às notas dadas, e que o fato de ser ou não um avaliador certificado influenciou o resultado da avaliação. Concluiu-se que o uso dos critérios de Cambridge na avaliação da produção oral dos alunos da escola investigada ainda não é eficaz pela falta de conhecimento ou treinamento de alguns professores nestes critérios.

Palavras-chave: Avaliação, Fala em inglês como LE, Cambridge ESOL

ABSTRACT

This study aims to investigate the effectiveness of the use of the University of Cambridge English for Speakers of Other Languages (ESOL) Examinations criteria to assess speaking in English as a second language in an English language institute in Canoas, RS. The participants in this study were 4 students and 5 teachers at this school. The data for this study were collected through teachers' assessment of students' speaking in English, who were recorded on audio taking Cambridge proficiency mock (simulated) tests, and questionnaires answered by the participant teachers after the assessment. This methodology was used in order to achieve the specific objectives proposed by the study: a) to investigate if the teachers in the school recognize the Cambridge criteria used to evaluate speaking, which are in use at school, b) to examine if there are discrepancies between teachers concerning the ratings given to each student and the criteria used, c) to see if the fact that a teacher being or not a certified Cambridge examiner influences the assessment process. Regarding the results obtained, we point out that some teachers use the desired criteria and others do not. It was also verified that there are some differences regarding the given grades, and that being or not a certified examiner influenced the outcome of the assessment process. We concluded that the use of the Cambridge criteria in the assessment of students' oral production at the investigated school is not effective yet due to some teachers' lack of knowledge or training on these criteria.

Keywords: Assessment, Speaking English as a foreign language, Cambridge ESOL

LIST OF TABLES

SUMMARY

1. INTRODUCTION

The aim of any person who starts learning a second language[1] varies exactly according to each person. However, the most common objective is reaching a level of oral fluency which enables them to communicate effectively in any context. Hughes (1994) states that the goal of teaching spoken language is the development of the ability to interact successfully in that language, and he also says that this involves comprehension as well as production. With the advent of globalization, a phenomenon which causes people to do business and make friends all around the globe, among other outcomes, the search for learning additional languages has increasingly grown. Given the economical power of countries which have English as an official language, this language has become a lingua franca for friends, acquaintances and business associates who speak dissimilar languages. Owing to that, nowadays there is a well-structured but competitive market of institutions which offer English language tuition as an extra activity thoroughly apart from regular school education in private language institutes.

The language institute that takes part in this study has the largest number of ESL students in Canoas: more than 900, with branches in the cities of Porto Alegre and Gravataí, both in the state of Rio Grande do Sul. Its staff is composed by teachers who have come from various educational backgrounds and got also different types of experience with the English language. The school has also become a Cambridge ESOL Examination Centre, due to its long tradition and excellence (in every term more than 90% of its students taking Cambridge ESOL examinations succeed in them) in preparing students for the University of Cambridge English proficiency exams. Owing to that, the language institute has revised its assessment procedures regarding all four skills (reading, writing, listening and speaking), adapting them so as to use Cambridge ESOL assessment criteria. Being regarded as the most subjective and most hardly quantified, the assessment of the speaking skill demanded to be further studied, once there were no defined criteria for speaking assessment in the school prior to that. Beforehand, the school trusted its teachers' experience and feeling to assess oral performance.

The study aforementioned was conducted by this researcher and, after presenting the results in a meeting of the school's pedagogical department in February of 2010, the oral assessment procedure was designed. It was agreed that, given the close relationship between

---

[1] The terms "second language" and "foreign language" are used indistinctively in this study.

the school and Cambridge ESOL, the school should adopt procedures similar to the ones being used in the speaking papers of Cambridge ESOL exams. Each level at the participant institute closely corresponds to the proficiency levels of Cambridge ESOL, and therefore each level uses different criteria when it comes to assessing speaking.

The general goal of this study is to analyze the effectiveness of the use of Cambridge ESOL criteria in oral assessment procedure at this English language institute in Canoas, RS, where the researcher currently works, aiming at either reinforcing its use or suggesting improvements if necessary. The researcher intends to record students of the school taking a simulated Speaking Paper of two exams, PET (Preliminary English Test) and FCE (First Certificate in English), due to the fact that these two exams are the most frequently taken by the school's students (around 85%). After the data collection process, these two recordings were submitted to five teachers of the investigated institute, who analyzed students' performance and rate them using their proper criteria. The teachers also answered a short questionnaire about the criteria they used during the analysis immediately after it ends. Through this study, we also expect to establish comparisons among teachers evaluating the same proficiency level and the criteria used, as well as between each teacher individually and the criteria each one used.

The search for an effective method of speaking assessment for the language school in which I learned English and currently work serves as intrinsic motivation. After having studied English in some other schools, I found a place where the process of learning was made easier by the school's teaching methods and constant concern about students' improvement. The school also made me become interested in the English language and the culture of English-speaking countries. Had I not studied at this institute, I would have probably chosen a career other than teaching English. It has been six years since I started working at this school, and nowadays, besides being a teacher, I also accumulate the position of Cambridge Exams Coordinator in the school.

By comparing students' end-of-term grades in speaking throughout these semesters working there, I noticed that many students had continuously been awarded top grades during their English learner trajectory and later on did not show the same level of performance when taking a PET or FCE preparatory course at the school. These discrepancies made me wonder if because of the tougher requirements found in the Cambridge exams students saw their performance levels drop or teachers did not know how to validly and reliably assess speaking. This doubt later gave rise to another question, whose pursuit of answer gave rise the general

objective of this study: after the adoption of Cambridge parameters, how effectively can teachers use Cambridge ESOL criteria in the speaking assessment process?

Taking into consideration this general goal, the specific objectives of this study are:

a) to investigate if the school's teachers acknowledge and use the Cambridge speaking assessment criteria, which are being used in the school;

b) to examine if there are discrepancies among teachers regarding the ratings awarded to each student of each level of proficiency and the criteria used;

c) to see if the fact of being a teacher a certified Cambridge examiner or not influences assessment process.

In order to achieve the objectives listed above, this study aims to respond the following Research Questions:

1) Which criteria are used by Cambridge examiners and non-Cambridge examiners?

2) Are there any differences among the candidates' grades in speaking given by each group of examiners?

3) Are there any differences when it comes to the criteria used for assessing each level of proficiency of relevance to this research (PET and FCE) used by the same assessor?

4) Are the criteria used by Cambridge examiners and non-examiners the same as the ones stated by Cambridge ESOL Examinations?

This study is divided into five chapters: Introduction, Theoretical Background, Method, Data Analysis and Discussion and Closing Remarks. In the present chapter, the objectives and Research Questions which guide the present study were established. In the following chapter, the relevant theoretical basis for the elaboration of this will be presented. The third chapter brings a description of the method, by describing the participants involved in this study, as well as the procedures used in the data collection process. In the fourth chapter, the data collected for this study will be analyzed and discussed. The fifth and last chapter brings final considerations about the study. The references and appendices are available at the end of this paper.

This study is expected to bring substantial input in order to clarify, reinforce and bring improvements to the use of Cambridge criteria in the speaking assessment process at the English language institute. On top of that, this study is also expected to contribute to my

comprehension of the processes involved in speaking assessment and to add to my experience while a teacher and a Cambridge examiner.

2. THEORETICAL BACKGROUND

This chapter presents the theoretical discussion without which it would be virtually impossible to conduct this research. The first section briefly tells the history of language testing. Section 2.2 contains a short discussion on what some authors have published about speaking skills regarding the development of the speaking ability. The following section discusses the University of Cambridge ESOL Examinations, bringing some concepts of relevance according to Cambridge ESOL's understanding of speaking skills. After that, there are subsections on Cambridge ESOL Assessment criteria and the two exams of special interest to this study: PET and FCE. The last section brings a summary of the chapter.

2.1. A brief history of assessment in second language teaching

According to Brindley (2001), the term *assessment* refers to "a variety of ways of collecting information on a learner's language ability or achievement." This concept is restated by Brown (1993) when he says that a test is a way of measuring one's ability or knowledge in a given area. Brown adds that assessment takes place every day, whenever and virtually wherever we learn something. Hence, assessment seems to have had a righteous place in the process of acquiring a second language. Notwithstanding, it has not always been in that way. As it will be evidenced, the issue of assessment has been seen in many different ways and phases, many of which do not reflect the aspect aforementioned.

2.1.1. Pre-scientific stage

Spolsky (1995) defines the period before the 1950's, in which researchers seldom studied the area of second language assessment, as "pre-scientific stage". At that time, since the language tuition was considered a dependent part of what was called Human Science, language assessment was not specifically studied.

2.1.2. American structuralism

After the 50's and with the advent of the American structuralist theories the second language teaching field started to be seen as a separate discipline. As a consequence of this interest, many studies in this area began to flourish. In this sense, Schoffen (2003) argues that the structuralist school was "the first theoretical stream that scientifically worried about language studies and also the first to think scientifically of language tests and to seek a definition for proficiency."

At that moment in history, according to Bachman and Cohen (1998) languages were seen as a set of separate features (such as grammar and vocabulary) and skills (reading, writing, listening and speaking). In order to acquire a language, a learner also had to acquire habits which later would develop into proficiency through automation. For Vollmer (1983), to be proficient in a given language meant knowing its elements and automate various forms to put these elements into practice. For the structuralists, learning a second language meant having the command of its structural elements.

Since the language was understood as formed by these different elements, Lado (1957, 1961) proposed that each element should be measured separately. In that way, grammar, vocabulary and pronunciation knowledge as well as reading, writing, listening and speaking skills were tested one at a time, so that tests did not have any concern regarding context and skills dependence on one another.

Due to the structuralist assessment model and its practicability and reliability an entire tradition of objective tests was created. In these tests, candidates need to classify alternatives as either correct or incorrect, without any necessity for creating new answers. This method is still in use in many proficiency tests such as KET, PET, TOEFL and TOEIC, among others.

2.1.3. Linguistic competence and communicative competence

In the mid-1960's, the notion of "linguistic competence" was introduced by Noam Chomsky. Linguistic competence is the system of linguistic knowledge which is possessed by native speakers, and it contrasts to the concept of linguistic performance, or the way the language system is used in communication. The concept of linguistic competence is part of

the foundations of Chomsky's Generative Grammar and, according to him, "is concerned with an ideal speaker-listener, in a complete homogeneous speech-community, who knows its (the speech community's) language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of this language in actual performance" (CHOMSKY, 1965, p. 75). The author went on to say that every native speaker of a given language has a generative grammar within, or a system of rules that allows new combinations to form new sentences and permits these sentences semantic integration (CHOMSKY, 1973, p. 32).

Nevertheless, Chomsky's ideas were contested when Dell Hymes coined the concept of "communicative competence". According to Hymes, there are rules of usage whose inexistence would turn grammar rules useless (HYMES, 1972, p. 278). As Schoffen (2003) explains, Hymes believes that the language has to abide by rigid and important rules which deal with the conditions of production of utterances as well as the purpose and social place they express. Hymes's linguistics is therefore centered in the use of language, whilst Chomsky's is not. The former copes with the real, tangible evidence of the latter, the abstract knowledge of the system of rules, parameters and principles or mind configurations a language needs.

According to Schoffen (2003), from that point on and from the perspective of the communicative performance, a strong stream of studies of pragmatic and sociolinguistics starts to take place in the field of Linguistics, and researchers start realizing that it is impossible to be proficient in a language without being proficient in its use. The language use component began to appear in tests in more contextualized tasks, allowing candidates to provide a wider range of answers.

When the use of language came into the limelight, researchers such as Oller started criticizing Lado's notion of separate skills (isolated items), arguing that linguistic competence was composed by a range of interactional abilities that cannot be separated. Here the era of "integrative-sociolinguistic" assessment began: the communicative competence was considered too broad to be assessed through distinct test of grammar, reading, vocabulary, etc. For Oller (1979), if isolated items separate linguistic skills, the integrative tests unite them again. While isolated items try to test language knowledge one bit at a time, the integrative tests try to assess a learner's capacity to use all language components at the same time. This rationale led to the coining of the "unitary trait hypothesis", which proposed proficiency as

indivisible. According to this hypothesis, vocabulary, grammar, phonology as well as the four skills (reading, writing, reading, and speaking) could not be taken separately. Oller (1979) defended this hypothesis, arguing that there was a global linguistic proficiency which could not get to be reached by simply adding all isolated items.

2.1.4. Communicative proficiency

A new concept of proficiency based on Chomsky's competence (1965) and Hymes's communicative competence (1972) arose. This concept, which takes into account communicative, socio-cultural and situational factors, was described by Bachman (1990) as the knowledge (or competence) and capacity to implement or execute this competence in an appropriate and contextualized communicative use of language.

Since the beginning of the studies called *communicative* in the late 70's, several models of communicative competence have been proposed. One of the most widely accepted is the one introduced by Canale and Swain (1980) and Canale (1983), which claims that the communicative competence cannot be reduced to a mere global dimension of linguistic proficiency. For this model, communicative competence must be divided into at least three different dimensions: grammatical competence (which includes lexicon, morphology, syntax and phonology), sociolinguistic competence (which consists of socio-cultural and discursive laws that regulate the use of language) and strategic competence (verbal and non-verbal strategies that are used to compensate for any possible communication problems).

According to Widdowson (1991), to be proficient in a language means more than just being able to read, write, listen and speak it. It also means to use these skills to produce utterances in order to achieve a desired communicative effect. According to this author, "it is generally required that we use our knowledge of the linguistic system with the objective of obtaining some sort of communicative effect. This means to say that it is generally required that we produce samples of the use of language" (WIDDOWSON, 1991, p.16). In other words, being proficient in a language requires not only a great knowledge of the linguistic system of it, but also a great knowledge of its use in a determined situation.

These studies and others, besides leading to language teaching methods which were based on communication, have also influenced testing practice. If before "the majority of tests

had an objective character (isolated items), now they bring the use of language in communicative situations[2]" (SCHOFFEN, 2003, p. 21).

According to Brown (1993), a communicative test must demand that the candidate use the language within an appropriate context. Furthermore, Brown believes that a communicative test must assess the learner's performance in a variety of language functions.

Following this trend, the area of foreign language assessment started making use of the notion of use of language in order to define proficiency. For Schlatter, Garcez and Scaramucci (2004, p. 356) "based on the concept of use of language as a joint action of participants with a social purpose, the concept of linguistic proficiency/success changes from meta-linguistic knowledge and system control to appropriate language use to perform actions in the world". This means that proficiency testing should take into consideration the context of language use, the participants involved in the process and the participants' goals.

It is clear that the concept of linguistic proficiency has been agreed and disagreed upon many times during the history of language acquisition. Besides, according to Schoffen (2003, p. 22), "it still lacks clearer definitions". According to Vollmer (1993, p. 152), "proficiency is what proficiency tests measure". The author means that proficiency is what people who elaborate a determined proficiency test want them to assess, based on both theories they believe in and language definition they carry.

That is why Bachman (1990) states that language tests must be designed in accordance with a clear definition of linguistic abilities. When a test is devised, it is essential to define very clearly what means to know a language, because the essence of our comprehension of the skills will define the nature of the test.

2.2. Speaking Skills[3]: development

According to Ur (1996), of all the four skills (listening, speaking, reading and writing), speaking seems intuitively the most important. Still according to Ur, one clear example of this intuition is that the "people who know a language are referred to as 'speakers' of that language, as if speaking included all other kinds of knowing; and many if not most foreign

---

[2] The translations in this study are of my responsibility.
[3] This section presents a general discussion about the issue. The skills considered by Cambridge ESOL Examinations in its speaking assessment are presented in section 2.4.

language learners are primarily interested in learning to speak" (UR, 1996, p. 120). For Lazaraton (2001, p. 103), most people believe that "to speak a language is synonymous with knowing that language since speech is the most basic means of human communication".

The rising popular interest in speaking a foreign language, more evidently English, combined with the impact of communicative competence theory on second language teaching made researchers study what features of "the communicative approach" can be considered relevant to teaching and assessing oral skills. Lazaraton (2001) states that one of the most obvious advantages which was brought by this theory is that it is no longer acceptable to focus only on the development of our students' grammatical competence. The author argues that teachers are expected to balance a focus on *accuracy* (or grammatical competence) with a focus on *fluency* as well. The term "fluency", for Lazaraton, is understood as a 'natural language use', which is likely to take place when speaking activities focus on meaning and its negotiation, when speaking strategies are used, and when overt correction is minimized.

Bailey (2003) claims that in order to communicate well in another language, we must make ourselves understood by the people we are speaking with, which means that there is some need to be accurate in speaking the target language. This author comprehends accuracy as being the extent to which students' speech matches what people actually say when they use the target language. A definition of fluency is also provided by the author: it is "the extent to which speakers use the language quickly and confidently, with few hesitations and natural pauses, false starts, word searches, etc".

Apart from accuracy and fluency, Hinkel (2006) adds that speaking in an L2 requires a sufficient lexicogrammatical repertoire for meaningful communication to take place. Due to the fact that English has increasingly become a language which is used by speakers of other languages, clarity is essential. As a consequence, the author proposes that overall intelligibility is mostly desirable, and defends the teaching of *pronunciation* (articulation of specific sounds, word stress, etc.) and *intonation* (prosody, rhythm). Also, as an outcome of the growing global mobility and internationalization of English, Hinkel argues that learners should be instructed on sociocultural features, which entails teaching *discourse organization and structuring*, as well as speech acts, such as requests, refusals, compliments, or clarification questions. On top of that, the author claims that, owing to the variety of contexts in which English has been used, learners must be instructed not only on written but also spoken register (or *appropriacy* of language).

What can be understood from this discussion is that in order to speak a second language, one cannot rely only on an intensive course on its grammatical rules and repeating phrases or sentences to gain fluency. The ability to speak a language is composed by many different but relevant aspects, such as intelligibility, appropriacy, discourse management, sociocultural knowledge and, of course, accuracy and fluency as well. In this sense, an assessor needs to have an accurate sensitivity and a great amount of awareness of all these aspects, and proficiency tests must provide opportunities for candidates to demonstrate their level of knowledge concerning the same aspects.

As far as the Cambridge examinations are concerned, they consist of broad tests which check candidates' performance in all aspects of communicative competence including the speaking ability. The Cambridge exams can be observed in more detail in the following subsection.

2.3. The University of Cambridge ESOL Examinations

All around the world, institutions of all sorts feel the necessity of classificatory tools with views to provide data about a wide range of areas, which are used in various purposes such as allowing the admission of a foreign student into an MBA course or deciding on what candidate best fits a key position in a company. In this context, language examinations for students whose mother tongue differs from the official system in a country of interest have been designed. Among the exams which deal with the level of knowledge in the English language TOEFL, TOEIC, IELTS, and Cambridge ESOL (this one being of particular relevance in this paper) are the best-known.

University of Cambridge ESOL Examinations (or Cambridge ESOL) has provided examinations in English for speakers of other languages since 1913, whatever age and cultural background they may be. These examinations cover all four language skills – reading, writing, listening and speaking - , and include a wide range of tasks designed to assess the use of English in a variety of contexts. All in all, "Cambridge ESOL examinations assess the ability to communicate effectively in English" (UNIVERSITY OF CAMBRIDGE, 2007a, p. 3).

The University of Cambridge offers five exams which encompass different levels of proficiency in English, following the regulations of the Association of Language Testers in

Europe (ALTE)[4] and the Council of Europe[5] and its Common European Framework of Reference for Languages (CEFR)[6]. This framework consists of "a table that shows different levels of proficiency by listing which linguistic tasks learners can do at each level, and is used worldwide by a number of institutions" (COUNCIL OF EUROPE, 2001, p. 5). The Cambridge ESOL exams are: KET (Key English Test), PET (Preliminary English Test), FCE (First Certificate in English), CAE (Certificate in Advanced English) and CPE (Certificate of Proficiency in English), and correspond to the CEFR levels according to the following table.

Table 1: CEFR Levels and Cambridge ESOL Main Suite

	CEFR Levels	Cambridge Main Suite	Command of the language
Proficient user	C2 Mastery	Certificate of Proficiency in English	Fully operational
Proficient user	C1 Advanced	Certificate in Advanced English	Good operational
Independent user	B2 Vantage	First Certificate in English	Generally effective
Independent user	B1 Threshold	Preliminary English Test	Limited but effective
Basic user	A2 Waystage	Key English Test	Basic
Basic user	A1		

The Cambridge ESOL approach to speaking is grounded in communicative competence models, according to Hymes (1972), including Bachman's (1990) Communicative Language Ability, which was built on the work of Canale & Swain (1980) and Canale (1983), and the work of other researchers working in the field of task-based learning and assessment, such as Skehan (2001), Weir (1990, 2005). As Taylor (2003) notes

---

[4] The Association of Language Testers in Europe (ALTE), which is an association of language exam providers, was founded in 1989 and has now 33 members. More at http://www.alte.org/

[5] The Council of Europe is an international organization which promotes co-operation between all countries of Europe in the areas of legal standards, human rights, democratic development and cultural co-operation. More at http://hub.coe.int/

[6] The Common European Framework of Reference for Languages (CEFR) is a guideline used to describe achievements of learners of foreign languages across Europe and in other countries. More at http://www.cambridgeesol.org/about/standards/cefr.html

in her discussion of the Cambridge approach to speaking assessment, Cambridge ESOL tests have always reflected a view of speaking ability which involves multiple competencies (e.g. lexical/grammatical knowledge, phonological control, pragmatic awareness, etc), to which has been added a more cognitive component which sees speaking ability as involving both a knowledge and a processing factor. The knowledge factor relates to a wide range of lexis and grammar which allow flexible, appropriate and precise construction of utterances in real time. The processing factor involves, for Levelt (1989), a set of procedures for pronunciation, vocabulary/grammar and established phrasal *chunks* of language which enable the candidate to conceive, formulate and articulate relevant responses with on-line planning reduced to acceptable amounts and timing. What is more, spoken language production is seen as a situated social practice which involves reciprocal interaction with others, as being purposeful and goal-oriented within specific context.

The features of the Cambridge ESOL speaking exams reflect the underlying construct of speaking. One of the main features is the use of direct tests of speaking, which aims at ensuring that speech elicited by the test engages the same processes as speaking in the world beyond the test and reflects the view that speaking has not just a cognitive, but a socio-cognitive dimension. Pairing candidates where possible is a further feature of Cambridge ESOL tests, which allows for a more varied sample of interaction, i.e. candidate-candidate as well as candidate-examiner. Similarly, the use of a multi-part test format allows for different patterns of spoken interaction, i.e. question and answer, uninterrupted long turn, discussion. The inclusion of a variety of task and response types is supported by numerous authors such as Bygate (1988), Chalhoub-Deville (2001), Fulcher (1996), Shohamy (2000) and Skehan (2001), who have stated that multiple-task tests allow for a wider range of language aspects to be used and so provide more evidence of the underlying abilities tested, and contribute to the exam's fairness.

Another feature of the Cambridge ESOL speaking tests, according to Bachman (1990), is the authenticity of test content and tasks, as well as authenticity of the candidate's interaction with that content. Still according to the author and also Saville (2003) and Spolsky (1995), a concern for authenticity in Cambridge ESOL exams can be seen in the fact that particular attention is given during the design stage to using tasks which reflect real-world usage and are relevant to the contexts and purposes for use of the candidates.

As well as informing speaking test format and task design, the underlying construct of spoken language ability also shapes the choice and definition of assessment criteria, which

cover Grammar/Vocabulary, Discourse Management, Pronunciation, and Interactive Communication.

PET and FCE are of special interest for this study because the number of the school's students enrolled in these two exams exceeds by far other exams, and therefore their oral assessment parameters were chosen to guide this research.

2.3.1. Cambridge ESOL Terminology on Oral Assessment

In this section, the concepts which are used throughout this study are explained according to the PET Handbook for Teachers (UNIVERSITY OF CAMBRIDGE, 2007b, p. 52) and the FCE Handbook for Teachers (UNIVERSITY OF CAMBRIDGE, 2007a, p. 67). It is essential to understand that, according to Cambridge, the students' speaking performances should demonstrate resourcefulness in what these concepts profess. Unfortunately, the University of Cambridge does not divulge the researchers on which these concepts are based.

Accuracy: this term refers to the ability to produce grammatically correct sentences/utterances, both in form and syntax.

Appropriacy: it refers to whether a word is suitable for the context it is being used in.

Coherence: it is the logical arrangement of utterances to form spoken discourse and to develop arguments or themes.

Extent: it refers to the appropriate length of individual contributions (long or short) to develop the discourse.

Hesitation: the ability to participate in the development of the interaction without excessive hesitation or silence.

Individual sounds: the effective articulation of individual sounds to facilitate understanding.

Initiating and responding (flexibility): the ability to participate in a range of situations and to develop the interaction by taking the initiative in a conversation and adapting to new topics or change of direction.

Intonation: this term refers to the use of a sufficiently wide pitch range and the appropriate use of it to convey intended meanings.

Range: the active use of various grammatical forms as vocabulary.

Relevance: the extent in terms of importance of the speaker's contribution, related to the tasks and to preceding contributions in the discourse.

Stress and rhythm: the appropriate use of strong and weak syllables in words and connected speech, the linking of words, and the effective highlighting of information-bearing words in utterances.

Turn-taking: this refers to the sensitivity to listen, speak, and allow others to speak, as appropriate.

These concepts are the elements of which each of Cambridge's assessment criteria are formed. In the next subsection, the assessment criteria will be explained and detailed.

2.3.2. Cambridge ESOL Oral Assessment

The assessment criteria are the same in both PET and FCE examinations. Candidates are assessed on their individual performance, and not in relation to each other. Both examiners assess them according to what is expected from a PET-level or a FCE-level learner. The interlocutor awards a mark for global achievement, and the assessor awards marks according to four criteria: Grammar and Vocabulary, Discourse Management, Pronunciation, and Interactive Communication. Each of these criteria is worth a mark which can range from 0 to 5, with a minimal difference of 0.5 between marks, totalizing 25 marks. Cambridge understands that, following the ideas of Hymes's communicative competence, all criteria are absolutely essential to effective communication, and consequently each criterion weighs the same.

2.3.2.1. Grammar and Vocabulary

This criterion deals with how well a candidate uses grammatical and lexical forms and syntax effectively and appropriately (*accuracy* and *appropriacy*) when approaching the tasks. It also includes the *range* of both grammatical forms and vocabulary. It is essential to say that some errors are acceptable, as long as they do not impede communication.

2.3.2.2. Discourse Management

This scale refers to the *coherence*, *extent* and *relevance* of each candidate's contribution. Candidates are expected to maintain a coherent flow of language, either within a single utterance or over a string of utterances. Also, he/she is assessed on how important his/her utterances are in relation to what has been said.

2.3.2.3. Pronunciation

The ability to produce understandable utterances in order to fulfill the tasks is assessed here. Pronunciation includes *stress and rhythm*, *intonation*, as well as *individual sounds*. With views to that, the examiners put themselves in a non-language-specialist point of view and assess the overall impact of pronunciation and the amount of effort one has to make to comprehend the speaker. The use of different varieties of English is acceptable, provided that they are used through the whole test.

2.3.2.4. Interactive Communication

It assesses how well the language is used to achieve meaningful communication, and it includes *initiating and responding* without excessive *hesitation* as well as using strategies to keep or repair communication. In addition, examiners take in consideration the candidate's sensitivity regarding the norms of *turn-taking*.

2.3.2.5. Global Achievement Scale

The interlocutor awards each candidate a mark taking into consideration his performance and overall effectiveness throughout the entire Speaking Paper. The global mark is an independent impression which reflects only the interlocutor's perspective.

As we can notice, Cambridge´s criteria follow the communicative competence concepts of communicative, sociocultural and situational factors. Recollecting what was previously mentioned in this study, according to Bachman (1990) a speaker must have the knowledge and capacity to implement or execute the speaking competence in an appropriate and contextualized communicative use of language. Widdowson (1991) states that being proficient in a language is to be able to produce utterances in order to achieve a desired communicative effect. Brown (1993) talks about the use of language within an appropriate context. It means that being proficient in a language is to make good use of it in different contexts. It is to be able to modify the language to achieve different goals.

In the next two sections, in order to reach a full understanding of the entire process, the two tests chosen to make part of this study will be explicated in more detail.

2.3.3. Preliminary English Test (PET)[7]

As far as the speaking skill is concerned, PET is mainly focused in analyzing whether an ESL student can communicate at an intermediate level in the target language. They are expected to perform simple communicative tasks, such as:

- get all the information needed from a tourist information center;
- understand the main point of a commentary and ask questions in order to get more information;
- deal with most situations likely to arise when making travel arrangements or when travelling;
- take part in discussion involving the exchange of factual information or receiving instructions, at an academic or professional level;
- receive and pass on telephone call information;
- express opinions, emotions, reactions, etc.

It is clearly stated that a candidate is allowed to pronounce words using all varieties of English, as long as he does not change the pronunciation of the same word at different parts of his performance.

---

[7] The information shown in this section was taken from PET Handbook for Teachers (UNIVERSITY OF CAMBRIDGE, 2007b, p. 4-15).

2.3.3.1. PET Speaking Paper: general description

The Speaking Paper is the third part of the exam and is usually the last paper a learner takes. The entire paper takes about 12-14 minutes and is taken by a pair (or exceptionally a trio) of candidates simultaneously, and is conducted by two examiners: one who acts as both assessor and interlocutor, managing the interaction by asking questions and proposing tasks; and another who acts only as an assessor and does not join the conversation. It is divided in four parts:

- Part 1 (Interview): each candidate talks to the interlocutor, who asks them questions in turn, using pre-defined questions, in about 2-3 minutes. It focuses on giving factual personal information about the present, past and future.
- Part 2 (Collaborative task): candidates interact with each other for about 2-3 minutes. The interlocutor sets up the activity, which is a simulated situation. Visual aid is provided (e.g. pictures or photographs), and candidates are expected to make and respond to suggestions, discuss alternatives and negotiate agreement.
- Part 3 (Long turn): each candidate talks on his own for about 1 minute about a color photograph, aiming at describing it and managing discourse, using appropriate vocabulary. Both photos relate to the same issue, and the whole part takes about 3 minutes.
- Part 4 (Discussion): general conversation, in which candidates interact with each other about the theme established in Part 3 in about 3 minutes. The interlocutor proposes the activity using a standardized rubric. This part targets the use of language to express opinions, likes/dislikes, preferences, experiences, habits, etc.

2.3.4. First Certificate in English (FCE)[8]

A student at this level can be assumed to have sufficient ability to operate effectively in English in many contexts, e.g. social, tourist, work and study. However, the FCE aims at

---

[8] This section presents information based on FCE Handbook for Teachers (UNIVERSITY OF CAMBRIDGE, 2007b, 4-16).

the use of English in a more professional scope, for instance secretarial and managerial positions. When it comes to speaking, a learner is expected to be able to:

- follow a talk on a familiar topic;
- keep up a conversation on a fairly wide range of topics;
- ask for clarification and further explanation and be very likely to understand the answer;
- express his/her own opinion, and present arguments to a limited extent;
- answer predictable or factual questions;
- check that all instructions are understood.

As well as in PET, speakers at FCE level are allowed to use any particular variety of English providing that they keep using the same variety with some consistency.

2.3.4.1. FCE Speaking Paper: general description

When it comes to FCE, the Speaking Paper is the 5th in the order of papers, and it lasts about 14 minutes. Two candidates (or exceptionally three) take the test at the same time, being assessed by two examiners. Again, one of the examiners works as both interlocutor and assessor and manages the interaction either by asking questions or providing cues for candidates. The other examiner acts as an assessor and does not take part in the conversation. The Speaking Paper consists of 4 parts:

- Part 1 (Interview): it is a conversation between the interlocutor and each candidate, who must present his ability by talking about himself and giving information about past experiences, present circumstances and future plans. This part takes about 3 minutes.
- Part 2 (Long Turn): the interlocutor delegates an individual task to each candidate by giving a pair of photographs for him to talk about. The candidate must organize a larger unit of discourse by comparing and contrasting the photos, giving information and expressing opinions. Each candidate has 1 minute to do that.
- Part 3 (Collaborative task): a two-way conversation between the candidates, who are given visual and spoken prompts, which are used in a decision-making task. In this part, learners are supposed to show they are able to exchange information,

express and justify opinions, agree and/or disagree, suggest, speculate and reach a decision within about 3 minutes.

- Part 4 (Discussion): the interlocutor here leads a 4-minute discussion on topics related to the collaborative task to explore them a bit further. Again, candidates are expected to exchange information, express and justify opinions, agree and/or disagree, etc.

2.4. Summary of the chapter

Throughout this chapter, we presented the necessity many companies and academic institutions have to make use of some sort of proficiency testing to classify potential employees and students. These tests are supplied by various organizations, among those is the University of Cambridge and its Cambridge ESOL Exams, a testing system which consists of a range of exams aiming at certifying different levels of English knowledge. The most important theoretical guideline onto which Cambridge bases his proficiency exams is Hymes's communicative competence (1972), among other fellow researchers in this same stream and others.

Hymes and the communicative studies do not represent the only stream in this regard, but nothing less than one in a vast list of studies in language proficiency. This fact justifies a brief history of assessment in language proficiency, starting before the 1950's with the pre-scientific stage, passing through the American structuralism, the concepts of linguistic and communicative competence, finally landing on the multifaceted communicative proficiency.

After that, the chapter starts narrowing the discussion again by introducing some authors' thoughts on what skills are necessary to be learnt in order to learn a second language. There are many different views, although most authors cited in this study believe that accuracy and fluency are the most desirable skills.

At this point of the chapter, the discussion and presentation of Cambridge ESOL Exams demands a clear definition of the terms currently in use by the international English proficiency exam venue and its speaking assessment criteria is felt. The Cambridge ESOL oral performance assessment criteria and aspects used in each of them are:

- Grammar and Vocabulary: accuracy, appropriacy and range;

- Discourse Management: coherence, extent and relevance;
- Pronunciation: rhythm, intonation and individual sounds;
- Interactive Communication: initiating and responding, hesitation and turn-taking.

It is also relevant to say that there is a fifth criterion called Global Achievement, which is an expression of each candidate's general performance, awarded by the interlocutor examiner.

The last part or the chapter brings a separate general overview of the tests used in this study (PET and FCE) and their parts.

3. METHOD

Along this chapter, the procedures taken throughout this research study are defined and justified, under the guidance of my Research Questions. As a means to achieve this, the chapter is divided in 3 sections: the first of them copes with a necessary reiteration of the Research Questions; in the second section, the reader will find information on the subjects who participated in this study; the third section presents the routines adopted towards a reliable data collection; last but not least, the fourth section briefly recapitulates the main points of the chapter.

3.1. Research Questions

This study aims to verify the effectiveness of the use of Cambridge ESOL criteria and parameters in the oral assessment performed by the teaching staff at the language institute in this study. As explained in the Introduction, the school has adopted Cambridge speaking assessment parameters, since the institution is a certified Cambridge ESOL Examination Centre and applies the exams with students who have reached certain levels of proficiency in English. Also, it is essential to say that the number of students taking PET and FCE each term exceed by far the number of students taking other exams, generally because of job market and academic demands.

Therefore, in order to achieve the aforementioned goal, I restate the Research Questions this study seeks to answer:

1) Which criteria are used by Cambridge examiners and non-Cambridge examiners?

2) Are there any differences among the candidates' marks (grades) in speaking attributed by each group of examiners?

3) Are there any differences when it comes to the criteria used for assessing each level of proficiency of relevance to this research (PET and FCE) used by the same assessor?

4) Are the criteria used by Cambridge examiners and non-examiners the same as the ones stated by Cambridge ESOL Examinations?

The result of this study is expected to contribute to a better understanding and application of the parameters in use by Cambridge ESOL Examinations by both the researcher and the teaching staff at the language school.

3.2. Participants

This section, which aims to provide more information about the study participants, will be organized in order to include, firstly, the students who took the oral test, and then the teachers who served as examiners in this investigation. Since the purpose of this analysis is merely academic, every participant in this experiment, candidates and assessors, received a codename with which they were identified during the study. This is to prevent any form of influence based on previous knowledge of the candidates on the assessors' awarded grades, as well as keep the assessors' real names under secrecy, in order to detach them from any professional or personal judgment that may arise. It must be highlighted that each participant in this study has signed a Term of Informed Consent. Both terms of consent to candidates and assessors can be found attached to this report, in Appendixes A and B respectively.

3.2.1. Candidates

Four candidates were invited to take part in this study: two PET-level learners and two FCE-level learners. Each learner presents one pertinent variable to the investigation, as Table 2 below shows.

The choice for students in PET and FCE levels of proficiency in English is explained by the fact that the students in the school who take these exams are much more numerous than the ones who take other Cambridge exams. What is more, PET and FCE criteria for oral production assessment are equal, which leaves only one aspect that differentiates them: in PET, students are able to talk about familiar issues and are not expected to make distinction in register, that is, in the use of different levels of formality, whereas FCE students are more aware of the subtleties of using English in a wider scope of situations. This difference makes FCE students contribute with a wider range of structures and vocabulary, as well as producing larger utterances.

Table 2: Candidates

Candidate's codename	Age	Gender	City	Native language	Other languages	Variable
PET1	16	Male	Porto Alegre	Portuguese	English	Under preparatory PET classes
PET2	18	Female	Canoas	Portuguese	English, learning Spanish	Recently certified in PET – grade PASS
FCE1	38	Female	Canoas	Portuguese	English	Under preparatory FCE classes
FCE2	21	Female	Canoas	Portuguese	English, German	Recently certified in FCE – grade C

Another point to be clarified is the choice for students who present the particular variables shown in Table 2. Each candidate shows a dissimilar level of proficiency with the objective of performing differently in the oral test so that assessors also grade them differently, as assessors are expected. Theoretically, a student preparing to take PET will perform slightly worse than one who has been recently certified in the level with a grade PASS, which means, between 70 and 79% (70% represents the minimal passing percentage). The same rationale was used concerning the choice of students for the FCE speaking test (grade C means a percentage between 60 and 74%; borderline is 60%). If the chosen certified students had gotten the highest grades (PASS WITH DISTINCTION for the PET; A for the FCE), the difference in performance in relation to the uncertified students would have been so greater that assessment would be obvious and not test the assessors' command of the use of the Cambridge ESOL parameters when assessing subtle differences in proficiency, as intended.

The chosen candidates are speakers of English as a second language whose mother tongue is Brazilian Portuguese. All candidates were selected among the researcher's students, and study English at the investigated institute, in Canoas.

3.2.2. Assessors

Five assessors were invited to take part in this study and each one was chosen due to the fact that they represent, on account of presenting some relevant variables, the totality of teachers working at the school. All of them make part of the English school teaching staff in Canoas. Below, Table 3 shows assessors' codenames and whether he is or not a Cambridge examiner, as well as other relevant information.

Table 3: Assessors

Assessor's codename	Age	Experience as a teacher	Teaches PET/FCE?	Experience PET/FCE levels	Cambridge examiner?	Experience as examiner
A	19	1 year	Yes, PET and FCE	1 year	No	-
B	35	12 years	Yes, PET	5 years	Yes, PET only	6 months
C	23	4 years	No	-	No	-
D	31	7 years	Yes, PET and FCE	3 years	Yes, PET and FCE	1 year
E	33	14 years	Yes, PET and FCE	5 years	Yes, PET and FCE	3 years

Firstly, assessor A is a 19-year-old female teacher who has just started her career as a teacher. She was a student at the language school and has plenty of experience with Cambridge ESOL exams as she has taken all of them (PET, FCE, CAE and CPE). She has been also helping with the preparation of PET and FCE students of the school in speaking mock tests, by acting as an interlocutor. Assessor A was chosen to participate in the non-examiner group for she represents some teachers at the language institute who present the same characteristics.

Second of all, assessor B is a very experienced 35-year-old female teacher who has recently returned to work at the school after a 7-year leave, in which she dedicated to other personal and professional projects. Since her return, she has resumed her work with PET and FCE groups and more recently has become a certified Cambridge oral examiner, although she is only allowed to assess PET speaking at first. She was chosen to integrate the Cambridge examiner group in this study and represent the share of inexperienced certified examiners working in the school as teachers.

Third, the choice for assessor C represents a chance for a member of the largest group of teachers who compose the teaching staff in the school. She is a 23-yeart-old teacher with a certain experience teaching English, although she has never taught any of the Cambridge groups of the school. She has, however, participated in PET and FCE speaking mocks playing the role of interlocutor. In this study, she takes part in the non-examiner group.

Last but not least, assessors D and E represent the teachers who have experience both as teachers in PET and FCE levels and as certified Cambridge examiners, E more than D in this sense though. Assessor D, a 31-year-old female, got her certification as an examiner about one year ago, and since then has participated in a number of speaking tests for Cambridge ESOL. Assessor E, who is 33 years old and female, is also the most experienced Cambridge examiner working at the moment at the investigated school. Assessors D and E integrate, together with assessor B, the group of Cambridge examiner in this study.

3.3. Data Collection Procedures

This section brings a presentation of the instruments used in order to collect the necessary data to this study. The first sub-section (3.3.1) explains the use of recordings of the Speaking Mock Tests, followed by section 3.3.2 which shows the Assessment of these recordings by teachers working at the English language institute. The third section introduces a questionnaire to which the assessors were submitted, which aimed to provide an explanation about the criteria used during the evaluation process and thus help the rationale behind the marks awarded.

3.3.1. Speaking mock tests

Aiming at the recording of a large set of utterances to be assessed, it was decided that the best way to do so was to submit each pair of candidates (PET1/PET2; FCE1/FCE2) at the same time (due to the interactional nature of the test) to a speaking mock test of their respective tests. Mock tests are previously applied tests that mimic the actual test, in order to prepare candidates for it, and are taken from preparatory books published to help students in their preparatory process. The PET and FCE speaking mock tests used in this study were taken respectively from the books *PET Practice Tests* (2008) and *FCE Practice Tests* (2009), published by Cambridge University Press, and are available in Appendixes C and D, at the end of this monograph. A general description of what both PET and FCE Speaking Papers consist of is provided in chapter 2 of this study.

The simulation was recorded in audio, and it only differed from a real test in one aspect: there was not any assessor. Apart from the candidates taking the test, there was only one person in the room: the interlocutor, whose role was played by this researcher. The PET mock test lasted 12 minutes and 37 seconds, while the FCE mock test took lasted 13 minutes and 44 seconds. The audio was recorded with the use of a Dell Inspiron 15R laptop computer and 3 microphones (one for the interlocutor, one for each candidate). The computer was located on the left side of the table used in the mock test.

3.3.2. Assessment

After having recorded both PET and FCE speaking mock tests, the recordings were submitted to the assessors' appreciation. All assessors listened to the two recordings together, in the presence of the researcher, and could not speak to or look at each other during the whole time. The experiment took place in the Language Lab of the school, which provides 20 individual stations divided by partitions and equipped with headphones and computers. The assessors should award each of the four students a grade between 0 and 5 marks, within intervals of 0.5 marks. They were not instructed on which criteria they should follow to assess the candidates, or on which group of the school each candidate was in. What was said was that they were going to listen to two recorded speaking mocks, the first a PET and the second an FCE mock. The mock test was digitalized and shown to the assessor during the hearing of

the recordings, so that they could follow the structure of questions asked and see the pictures shown during the test. Each assessor received mark sheets to help them assess each candidate. The mark sheets used in this study are shown in Appendix E, and differ from the ones which are normally used in speaking mocks at the language institute in one aspect: the former does not show fields with the criteria to be used, while the latter does. Appendix F shows the mark sheets used at the school.

3.3.3. Questionnaire to assessors

In this part of the experiment, immediately after having graded all students, the assessors received a written questionnaire with questions to be answered which would help this researcher analyze the rationale behind the given grades. The questionnaire was shown on the computer screen in front of them and each assessor answered the questions on a paper sheet, still in the presence of the researcher. All assessors answered the questions in English, except for Assessor C, who used both English and Portuguese.

The questionnaire consists of five questions which target at reaching the objective of examining how teachers at the school evaluate their students' speaking and if and how they use the Cambridge ESOL assessment criteria. The questions to assessors were the ones that follow:

1) Did you use the same criteria to assess the PET and FCE students in this study?

2) Which criteria did you use in order to assess the students in the recordings in each level (PET and FCE)?

3) Define conceptually each criterion used to assess the students in the recordings.

4) How much is each criterion used worth?

5) Are the criteria used in this study the same as the ones used by you with students in these levels during a regular term?

In case any answer was considered unsatisfactory or unclear, the researcher could ask any assessor for clarification in form of a conversation. However, this procedure was not necessary. A copy of the written questionnaire is available in Appendix G.

3.4. Summary of the chapter

In this chapter, we presented the procedures taken during the execution of this research study.

The beginning of the chapter restates the Research Questions, or questions which guided the procedures taken and lead this study to a reliable outcome.

Right after that, the subjects who participated in this study were introduced: the candidates, who had been chosen according to some variables which are pertinent to the investigation, took a speaking mock test and consented about being recorded doing so; and assessors, who had also been chosen for showing characteristics which in a way represent the totality of the pedagogical staff at the language institute, awarded grades to the candidates' speaking performance and provided thorough explanation on the criteria used in this assessment.

The third section presented the instruments which were used to collect the data for this study. These instruments consist of PET and FCE speaking mock tests, which were recorded and submitted to the assessors' appreciation. These subjects assessed and graded the students' oral performances and, immediately after that, answered a written questionnaire to explain the criteria used in their evaluation.

4. DATA ANALYSIS AND DISCUSSION

In this chapter, the data collected in this experiment will be presented and discussed. In the first and second sections, the grades awarded by the assessors and their answers to the questionnaire will be shown and confronted. Following this, the third section will bring the analysis and discussion of the set of information gathered, which will be developed with aim to provide answers to the Research Questions which have guided this study.

4.1. Analysis of Candidates' Mark Sheets

This section presents the grades provided by the five assessors who analyzed the four ESL students who were recorded taking either a PET or FCE Speaking Mock Test. The grades, as well as a mathematical average of them, are shown on Table 4 below, aiming at a better understanding of the data.

Table 4: Grades and Mathematical Averages

Candidate	Assessor A	Assessor B	Assessor C	Assessor D	Assessor E	Average
PET1	4.0	5.0	5.0	5.0	3.5	**4.5**
PET2	3.5	4.5	4.0	4.0	3.5	**3.9**
FCE1	4.0	4.0	3.5	4.0	3.0	**3.7**
FCE2	4.5	4.0	3.0	5.0	3.5	**4.0**

The results clearly show that, according to the majority of assessors, the students taking the exams performed slightly differently in the mock tests, taking the mathematical average into consideration. In general, it seems that candidate PET1 did a bit better than PET2 as far as the PET mock test is concerned. Yet, assessor E apparently considers that both candidates performed equally. When it comes to the FCE speaking mock test, candidate FCE2 appears to have done a bit better than FCE1 generally. Even so, a deviation can be seen in the marks awarded by assessor C: for him, candidate FCE1 did slightly better than FCE2. What

must be remembered is that assessor C is not a certified Cambridge ESOL examiner and has never taught any PET or FCE levels at the language institute.

The table also shows that there were not many divergences among the assessors in terms of the numeric grade awarded. The University of Cambridge admits a standard deviation of up to 10% higher or lower than the mathematical average of marks given by assessors taking the examiner training course (2007a). In case of a higher percentage, the assessor is sent to be retrained. However, we can noticeably perceive that some of the existing divergences were somewhat larger than that. This can be noticed as we take, for instance, the grades of candidate PET1: he was given three marks 5.0, a 4.0 and a 3.5. The assessors who gave marks 5,0 probably considered that the candidate performed flawlessly in that level or so, and did not make any mistake/error which could have impeded or misled communication. On the other hand, the other two assessors did not consider his performance likewise. Conversely, there is a disagreement that varies between 20-30% from the others. This shows that either situations might have happened: 1) assessors B, C and D did not consider or hear some limitations; or 2) assessors A and E judged the candidate too demandingly, reckoning his limitations crucial to a small failure in communication to a certain extent. It is also important to notice that, according to the admitted 10% standard deviation rule, assessor E would have fallen into retraining.

On top of that, the same sort of dissimilarity seems to have occurred in the assessment of candidate FCE2. From the lowest to the highest mark there was a difference of 30%. Again, assessor D awarded the top mark, 5.0. This time she was alone in her judgment though. According to the standard deviation rule, she would be sent to retraining. And one more time the bottom mark was given by assessor E, which makes us think if her standards are not excessively high or that the other assessors should be more demanding. It is essential to remember, however, that assessors D and E are certified examiners for PET and FCE. Nevertheless, one proviso must be added: assessor E has been a Cambridge examiner for 5 years, whilst assessor D has just begun examining PET and FCE.

4.2. Analysis of Questionnaire to Assessors

Throughout this section, the assessors' answers to the questionnaire will be presented and discussed. Each sub-section corresponds to a question, and all the answers to each question will be analyzed within these separate parts.

4.2.1. Did you use the same criteria to assess the PET and FCE students in this study?

Basically, some examiners use the same criteria to assess both PET and FCE speaking, only differing in terms of quality of language produced. Some other assessors did not. FCE candidates are supposed to use wider ranges of vocabulary and grammatical structures, be more accurate and participate more confidently in a wider variety of contexts. Thus, the correct answer to this question would be invariably a positive one.

Therefore, it was not at all surprising that the assessors who were not examiners (A and C) denied having used the same criteria, although assessor A later mentioned that, due to her having participated in many speaking mock tests. she knew about the correct criteria and that she reckons that in FCE "students are supposed to have a better fluency, accuracy, wider range of vocabulary and grammar structures".

Conversely, the assessors who were Cambridge ESOL examiners (B, D and E) confirmed having used the same criteria, adding that they took into consideration students' levels.

These results mean that although all teachers of the school were trained to follow Cambridge ESOL directives, not all of them are still aware of what they say. I believe that a second training session must be scheduled to reinforce parameters and directives. It is necessary to point out that Cambridge examiners did know that there are no differences in the criteria regarding PET and FCE.

4.2.2. Which criteria did you use to assess the students in the recordings in each level (PET and FCE)?

An excellent answer to this question would be: a) Grammar and Vocabulary, b) Discourse Management, c) Pronunciation and d) Interactive Communication, which are the standard assessment criteria used in PET and FCE exams, following the ideas of communicative competence. It is important to remember that all teachers at the school received instruction on these criteria.

The use of *Global Achievement Scale*, as it was mentioned in chapter 2, is considered irrelevant here since the assessors are working on their own during this experiment, and not in pairs as they would be in a real speaking test.

Assessors A, B and E mentioned having used the expected criteria. Assessor E also adds that "the level of English (PET or FCE) changes the kind of structures which are analyzed", and assessor A reckons that she emphasizes the criterion of Interactive Communication to her students because, due to her experience taking Cambridge exams, she believes "this is the one assessors take most into consideration". This rationale seems to agree with Canale and Swain's (1980) and Canale's (1983) idea of strategic competence, that means, the use of verbal and non-verbal strategies that are used to compensate for any possible communication problems.

To assessor C, levels of fluency and accuracy are of extreme importance. She also highlights the candidates' "engagement to the conversation and interaction". Discourse Management, Range of Vocabulary and Interaction are the criteria observed by assessor D.

These results clearly show the there is still room for improvement regarding teachers' training on Cambridge ESOL parameters. It seems that the teachers still have not achieved full understanding of the rules and procedures, as well as what is relevant or not.

4.2.3. Define conceptually each criterion used to assess the students in the recordings.

The criteria used in the exams are thoroughly defined in the Theoretical Background (Chapter 2) of this report. Still, it is convenient to restate and redefine them briefly with a view to better clarification.

a) *Grammar and Vocabulary:* the accurate and appropriate use of a given range of grammatical and lexical forms.

b) *Discourse Management:* this includes the candidate's levels of coherence, extent and relevance when performing orally.

c) *Pronunciation:* the ability to produce intelligible utterances, making use of proper stress and rhythm, intonation and individual sounds.

d) *Interactive Communication:* how the language is used to reach meaningful communication, through sensitive respect to rules of initiating, responding and turn-taking, and showing no undue hesitation.

As far as the assessors' answers are concerned, three of them (A, B and E) answered according to Cambridge standards. Assessor E fully described each criterion in detail, showing a great level of mastery in Cambridge assessment, whereas assessor A, besides mentioning the four criteria, added Global Achievement and its notion. Assessors B and E used words such as "language control" to refer to accuracy and did not mention conceptual words such as hesitation or turn-taking; however, they mentioned all other concepts properly. Assessor A used some few concepts, e.g. appropriacy, stress and intonation, and paraphrased others such as: "using the structures and words in a correct way" to refer to accuracy, "create valid arguments" to refer to relevance, and "asking questions to call the other participant's interaction" to make reference to initiating. She failed to mention however other important concepts such as individual sounds, extent, hesitation and turn-taking.

On the other hand, assessor C defined her own criteria (fluency, accuracy and engagement and interaction). According to her, fluency is "the ability to express yourself in a confident way and showing certain 'intimacy' to the language"[9]; accuracy is "the exactness in the use of rules"[10]; and engagement and interaction is "how well they (candidates) could get involved in the conversation, how natural they seemed and how much they respected the time and partner's speech"[11]. Despite the fact that assessor C apparently lacks knowing exactly how to define some standard concepts, many of them seem to be hidden in her answer, noticeably accuracy, appropriacy, stress, intonation, initiating and responding and turn-taking, them. Of course, this must be related to her not having taken any training as an oral examiner

---

[9] Translated from Portuguese: "a habilidade de expressar-se de uma maneira confiante e mostrando uma certa 'intimidade' com a língua."
[10] Translated from Portuguese: "a precisão no uso das regras."
[11] Translated from Portuguese: "quão bem eles conseguiram se envolver na conversação. quão natural eles pareceram e o quanto eles respeitaram o tempo e a participação do parceiro."

and not being directly involved with students' PET and FCE preparatory classes. However, she was present when the teachers were trained in Cambridge standards.

When it comes to assessor D's answers, she conceptualized her criteria as it follows: Discourse Management is "the capacity (…) of keeping their (candidates') speeches flow, without being helped and managing to convey their messages effectively"; Range of Vocabulary is "the appropriacy and extent of vocabulary used"; and Interaction as being "the ability of using language to exchange information and opinions effectively". The only concepts apparently present in assessor's D answer are appropriacy and range of vocabulary. There are no tentative reference to grammar or pronunciation whatsoever, not to mention that her idea of Discourse Management and Interactive Communication appear to be consistently faulty.

Once again, the necessity of retraining the teaching staff at the English institute appears. At this point, even a teacher (Assessor D) who is supposed to know the Cambridge standards shows lack of knowledge in this regard. Some teachers seem to remember some important concepts from the first training session, but not all teachers and not all concepts. We can conclude that these teachers are not using the Cambridge parameters with their groups.

4.2.4. How much is each criterion used worth?

Since all criteria are considered significant to communication, Cambridge ESOL states that each criterion is worth the same, or in other words, each criterion has the same number of total possible marks to be distributed.

This time, the assessors were unanimous to say that all criteria are worth the same and therefore have the same level of importance. To illustrate this, here is an excerpt of assessor D's answer: "each one of them (the criteria) are worth the same because they consider individually what students can do, and also help teachers measure different parts of a conversation". To add to that, assessor C argues that "each criterion is equally worth as everything together makes communication flow in a natural way". These answers seem to indicate that some traces of the teacher standardization training still remain.

One proviso must be made regarding assessors A's answer. As she mentioned in question 2, she claims to put an emphasis on Interactive Communication because in her words "people may show better their speaking skills when interacting". Still, she reckons that all four criteria must be equally weighed. Once again, assessor A's ideas seem to match what was stated by researchers regarding the communicative competence, which doubtlessly serve as a guide for the Cambridge examinations. They defend that a learner's proficiency can only be fully assessed when meaning is negotiated through pair or group interactive tasks. Here, another remnant of the first training on Cambridge standards seems to appear, although not quite the expected one.

4.2.5. Are the criteria used in this study the same as the ones used by you with students on these levels during a regular term?

Here, some distinct answers are expected, for not all assessors work with PET and/or FCE levels at the school. Therefore, these assessors are not expected to use the same criteria during regular term. Also, teachers working at this school are instructed to take into consideration some other features such as students' feelings and emotions or students' overall progress during a given period of time. As these concepts are not the main focus of this study, they are not going to be taken into account.

Assessors B, C, D and E confirmed using the same criteria in classroom assessment. B argues that she even makes sure her students know what they are expected to show in terms of speaking. D adds that teachers should also consider students' feelings and emotions and students' progress during the term. E points out that she uses the same criteria, but just in higher levels of the school.

The exception lies with assessor A. She honestly admits that the criteria used in a speaking mock differ from the ones she uses in class. She also says that teachers must take into account what they produce in class and if students match what is expected from them at a given level. In addition, she states that PET students are supposed to "communicate clearly, even making some mistakes of grammar and vocabulary, and be able to speak fluently, even hesitating a little." She says she assesses "how hardworking they (students) are, how much effort they put on trying to speak more, how much they speak in class, and not if they are

successful or not". Nonetheless, she recognizes she should be more severe in order to fully prepare students to Cambridge ESOL exams.

Since the necessity of retraining was already detected, these other results can serve to make us think about what other measures could be taken in order to minimize a deviation from Cambridge ESOL standards and assure that teachers will use the expected criteria according to the students' levels. One of the feasible actions is to write a document such as a reference book containing the criteria to be used by teachers in each level, from the kids' groups to the CPE groups. If teachers feel in doubt, they can easily access the document in the institute's computer network or even in a printed version.

4.3. Discussion Guided by the Research Questions

In this section, by resuming the main aspects brought by the data gathered in the mark sheets and in the questionnaire, I aim to provide answers to the Research Questions propounded to the present study. In 4.3.1, the first question will be discussed. The second question will be analyzed in 4.3.2, and sub-section 4.3.3 will, at last, compile questions 3 and 4.

4.3.1. Criteria used by examiners and non-examiners

The first Research Question was formulated in order to verify if the two groups of assessors (Cambridge examiners and non-Cambridge examiners) would present any differences between themselves. What was found, however, was a disagreement within members of the same groups.

To start with, the Cambridge examiners (B, D and E) showed some similarities in the criteria used. They also presented some differences though. While assessors B and E indeed used the criteria stated by Cambridge ESOL Examinations (Grammar and Vocabulary, Discourse Management, Pronunciation and Interactive Communication), assessor D presented a faulty acknowledgement of them, which may result in a problematic assessment process for both Cambridge and the language school.

What is more, the group composed by non-Cambridge examiners (assessors A and C) also delivered contradictory results. On the one hand, assessor A apparently has knowledge or at least a good awareness of which criteria must be used when assessing oral performance in these two exams (PET and FCE), showing she has studied and prepared herself to be responsible for Cambridge preparatory levels. On the other hand, assessor C revealed lack of knowledge or awareness of the elected criteria, which means that this teacher needs to be better prepared until she is able to teach the preparatory levels of the school.

4.3.2. Differences among examiners

The numeric indexes, also known as marks, attributed by each group of assessors certainly differed, yet again much more within the groups.

Taking the non-examiners in to consideration, they never agreed in any of the candidates' marks. Assessor A was more severe in the PET speaking mock, and assessor C in the FCE mock. What may have affected their assessment was, again, A's lack of examiner training, but her somewhat experience dealing with preparatory levels, and C's lack of training and also experience with the mentioned levels.

It seems that assessors B and D have agreed more regarding the marks. Two of their marks match, and one differs in only 0,5. Conversely, the numeric difference in comparison to assessor E reaches 1,5 for PET1 (concerning B and D) and FCE2 (concerning only D). These numbers give rise to a further question about the examiners' performance: once again, is assessor E simply more demanding than B and D, or her bigger experience as a Cambridge examiner gives her other useful tools such as maybe a better perception of what is relevant or irrelevant in the process of assessment?

4.3.3. Criteria used by level and assessor

These questions have been answered previously in this report, in sub-section 4.2.1. All assessors used the same criteria for both levels, PET and FCE, only varying the level of exigency. And all of them used the criteria preached by Cambridge ESOL Examinations,

except for assessor C, which can be explained by her lack of training as an oral examiner and experience with PET and FCE levels.

At the end of this study, the results found gave rise to another Research Question: in what ways can the results of this study assist the next steps in the implementation of a standardized oral assessment system in this language institute? The answer to this new question can be found in the conclusion of this study, in the next chapter.

5. CLOSING REMARKS

When this study comes to an end, it is imperative to recognize that there are discrepancies among the investigated English language institute teachers' ideas regarding speaking assessment. Some of the teachers who participated in this study do acknowledge and use the expected criteria. However, other teachers have a feeble notion of what Cambridge say about these parameters, and others seem never to have gotten in contact with what the school preaches as standard procedure, although we know they did receive instruction on that. Needless to say, the results found in this investigation corroborate the assumption that being a trained Cambridge oral examiner would help teachers' understanding of what should be consider in the speaking assessment process.

The purpose of this study was proven to be pertinent due to the factuality of the obtained results. The use of Cambridge ESOL parameters and criteria in the oral assessment process at the language school is not effective as it was expected to be, mostly because a considerable part of its teaching staff needs a more intensive instruction on the same procedures. It is believed that once teachers are reinstructed, maybe in another meeting of the pedagogical department, the Cambridge parameters start being used more effectively in the school. Also, the idea of the writing of a reference book for teachers' use whenever in doubt about which criteria they should use in a given level seems valid and absolutely accomplishable.

Regarding the limitations of the present study, it is necessary to recognize that the use of recordings of the speaking mock tests does not achieve the purpose of simulating adequately a real exam situation. In the real speaking paper, PET or FCE, the assessors can count on visual cues in order to interpret speech pauses and hesitation, how natural a candidate uses the language and everything which concerns interactive communication. In that sense, the use of video recordings would provide better material for the assessment process. Also, instead of having assessors answer a questionnaire, an interview with each of them could have been recorded in video or audio, providing explanations that could clarify their choices in terms of ratings and criteria used. Despite this, and taking into consideration all limitations, the study gave input and supplied relevant inferences taken from the data collected and its analysis.

To conclude, this study represents an invitation to future investigations about the effectiveness of the criteria used in the assessment of any sort of language testing, aiming to provide valid and reliable feedback to language learners in their journey in the path of proficiency. The study can also provide a consistent starting point for researchers who seek to continue monitoring the oral performance assessment at the English language institute and in any other language school.

REFERENCES

BACHMAN, L.F. *Fundamental considerations in language testing.* Oxford: Oxford University Press, 1990.

_____; COHEN, Andrew D. Language testing – SLA interfaces: An update. In: BACHMAN, Lyle F.; COHEN, Andrew D. *Interfaces Between Second Language Acquisition and Language Testing Research.* Cambridge: Cambridge University Press, 1998.

BAILEY, Kathleen M. Speaking. In: NUNAN, D. (ed.) *Practical English Language Teaching.* New York: McGraw-Hill/Contemporary, 2003.

BRINDLEY, G. Assessment. In: CARTER, R.; NUNAN, D. (eds.) *Teaching English to Speakers of Other Languages.* Cambridge: Cambridge University Press, 2001.

BROWN, H. Douglas. *Principles of Language Learning and Teaching.* 3 ed. Englewood Cliffs, WJ: Prentice Hall Regents, 1993.

BYGATE, M. (1988). *Speaking.* Oxford: Oxford University Press, 1988.

CANALE, M. On some dimensions of language proficiency. In: OLLER JR., J. W. (ed.) *Issues in language testing research.* Rowley, MA: Newbury House, 1983.

_____; SWAIN, M. Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, vol. 1, p. 1-47, 1980.

CHALHOUB-DEVILLE, M. Task-based assessments: Characteristics and validity evidence. In: BYGATE, M.; SKEHAN, P.; SWAIN, M. (eds.) *Researching pedagogic tasks.* London: Longman, 2001.

CHOMSKY, Noam. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press, 1965.

_____. Linguistic Theory. In: OLLER JR., J. W. (ed.) *Focus on the learner: pragmatic perspectives for the language teacher.* Rowley, Massachusetts: Newbury House, 1973.

COUNCIL OF EUROPE. *Common European Framework of References for Languages: Learning, Teaching, Assessment.* Cambridge. Cambridge University Press, 2001.

FULCHER, G. (1996). Testing tasks: Issues in task design and the group oral. *Language Testing*, ano 1, n. 13, p. 23-51, 1996.

HINKEL, Eli. Current perspectives on teaching the four skills. *TESOL Quarterly*, ano 1, n. 40, p. 109-131, 2006

HYMES, D. H. On communicative competence. In: PRIDE, J. B.; HOLMES, J. (orgs.) *Sociolinguistics: selected readings.* Harmondsworth: Penguin, 1972.

LADO, R. *Language Testing.* New York: McGraw-Hill, 1961.

_____. *Linguistics across cultures.* Ann Arbor, Michigan: University of Michigan Press, 1957.

LAZARATON, Anne. Teaching Oral Skills. In: CELCE-MURCIA, M. (ed.) *Teaching English as a Second or Foreign Language.* 3 ed. Boston, MA: Heinle & Heinle, 2001.

LEVELT, W. J. M. *Speaking: from intention to articulation.* Cambridge, MA: MIT Press, 1989.

OLLER, John W. *Language Tests at School: A pragmatic approach.* London: Longman Group Ltda., 1979

SAVILLE, N. The Process of Test Development and Revision within UCLES EFL. In MILANOVIC, M.; WEIR, C. (eds.) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002.* Cambridge: Cambridge University Press, 2003.

SCHLATTER, M.; GARCEZ, P. M.; SCARAMUCCI, M. V. R. O papel da interação na pesquisa sobre aquisição e uso da língua estrangeira: implicações para o ensino e para a avaliação. *Letras de Hoje,* ano 3, n. 39, p. 345-378, 2004.

SCHOFFEN, Juliana R. *Avaliação de proficiência oral em língua estrangeira: descrição dos níveis de candidatos falantes de espanhol no exame Celpe-Bras.* Dissertação de mestrado. Porto Alegre: UFRGS, 2003

SHOHAMY, E. Assessment. In: CELCE-MURCIA, M.; OLSHTAIN, E. (eds.) *Discourse and Context in Language Teaching.* Cambridge: Cambridge University Press, 2000.

SKEHAN, P. Tasks and language performance assessment. In: BYGATE, M.; SKEHAN, P.; SWAIN, M. (eds.) *Researching pedagogic tasks.* London: Longman, 2001.

SPOLSKY, B. *Measured Words.* Oxford: Oxford University Press, 1995.

TAYLOR, L. (2003). The Cambridge approach to speaking assessment. *Cambridge ESOL Research Notes*, 13, 2-4, 2003.

UNIVERSITY OF CAMBRIDGE. *FCE Handbook for Teachers.* Cambridge: Cambridge University Press, 2007a.

UNIVERSITY OF CAMBRIDGE. *PET Handbook for Teachers.* Cambridge: Cambridge University Press, 2007b.

UR, Penny. *A Course in Language Teaching: Practice and Theory.* Cambridge: Cambridge University Press, 1996.

VOLLMER, Helmut J. The structure of foreign language competence. In: HUGHES, Arthur; PORTER, Don. *Current Developments in Language Testing.* London: Academic Press Inc., 1983.

WIDDOWSON, H. G. *O Ensino de Línguas para a Comunicação.* Campinas: Pontes, 1991.

APPENDIX A: Term of Consent to Candidates

FORMULÁRIO DE CONSENTIMENTO

Prezado participante:

Por favor, leia o texto a seguir. Ele apresenta informações importantes a respeito do estudo do qual você fará parte. Após ter lido o texto, assine o documento, indicando que você entende a natureza desta pesquisa e que você consente em participar dela.

DADOS DA PESQUISA:

Título: *The use of Cambridge ESOL parameters in oral performance assessment in an English language institute*

Pesquisador: Alexandre Luís Koch da Silva, graduando em Letras

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

PROPÓSITO E BENEFÍCIOS:

Este estudo tem por objetivo investigar os parâmetros de avaliação oral em alunos de um curso livre de idiomas em níveis correspondentes ao PET (Preliminary English Test) e FCE (First Certificate in English) da Universidade de Cambridge. A pesquisa visa contribuir para a definição de um padrão de avaliação oral no mesmo curso livre, e é requisito parcial para a obtenção do título referente ao curso Licenciatura em Letras – Língua Inglesa e Literaturas de Língua Inglesa da Universidade Federal do Rio Grande do Sul.

PROCEDIMENTOS:

O pesquisador gravará em áudio sua participação em Tum teste oral simulado de PET ou FCE, de acordo com seu nível de proficiência em língua inglesa. A gravação será posteriormente ouvida por avaliadores que darão uma nota de desempenho oral de 0 a 5 a cada participante.

OUTRAS INFORMAÇÕES:

A participação neste estudo é de caráter voluntário. Todos os participantes têm liberdade para cancelar sua participação a qualquer momento. As gravações ficarão de posse do pesquisador, e os dados coletados poderão ser utilizados em estudos posteriores. Os nomes dos participantes e avaliadores serão mantidos em sigilo. Para isso, cada participante e avaliador receberá um codinome.

DECLARAÇÃO:

Declaro que li e compreendi as informações acima e que consinto participar deste estudo.

_____   _____

Nome do Avaliador   Assinatura

_____

Data

APPENDIX B: Term of Consent to Assessors

FORMULÁRIO DE CONSENTIMENTO

Prezado participante:

Por favor, leia o texto a seguir. Ele apresenta informações importantes a respeito do estudo do qual você fará parte. Após ter lido o texto, assine o documento, indicando que você entende a natureza desta pesquisa e que você consente em participar dela.

DADOS DA PESQUISA:

Título: *The use of Cambridge ESOL parameters in oral performance assessment in an English language institute*

Pesquisador: Alexandre Luís Koch da Silva, graduando em Letras

Orientador: Prof. Dr. Ubiratã Kickhöfel Alves

PROPÓSITO E BENEFÍCIOS:

Este estudo tem por objetivo investigar os parâmetros de avaliação oral em alunos de um curso livre de idiomas em níveis correspondentes ao PET (Preliminary English Test) e FCE (First Certificate in English) da Universidade de Cambridge. A pesquisa visa contribuir para a definição de um padrão de avaliação oral no mesmo curso livre, e é requisito parcial para a obtenção do título referente ao curso Licenciatura em Letras – Língua Inglesa e Literaturas de Língua Inglesa da Universidade Federal do Rio Grande do Sul.

PROCEDIMENTOS:

Os avaliadores ouvirão duas gravações: a primeira de um simulado oral de PET, e o segundo de FCE, devendo então avaliar e dar uma nota total a cada um dos participantes. Em seguida, os avaliadores responderão um questionário, formulado pelo pesquisador, sobre os procedimentos tomados ao avaliar os participantes das gravações.

OUTRAS INFORMAÇÕES:

A participação neste estudo é de caráter voluntário. Todos os participantes têm liberdade para cancelar sua participação a qualquer momento. As gravações ficarão de posse do pesquisador, e os dados coletados poderão ser utilizados em estudos posteriores. Os nomes dos participantes e avaliadores serão mantidos em sigilo. Para isso, cada participante e avaliador receberá um codinome.

DECLARAÇÃO:

Declaro que li e compreendi as informações acima e que consinto participar deste estudo.

_____ _____

Nome do Avaliador   Assinatura

_____

Data

APPENDIX C: PET Speaking Mock Test

3

PAPER 3: SPEAKING
Part 1

Part 1 (2-3 minutes)

Phase 1
Interlocutor

A/B Good morning / afternoon / evening.
Can I have your mark sheets, please?

(Hand over the mark sheets to the Assessor.)

A/B I'm .... and this is ....
He / she is just going to listen to us.

A Now, what's your name?
Thank you.

B And, what's your name?
Thank you.

B Candidate B, what's your surname?
How do you spell it?

Thank you.

A And Candidate A, what's your surname?
How do you spell it?

Thank you.

(Ask the following questions. Ask Candidate A first.)

Where do you live / come from?

Adult students
Do you work or are you a student in .....?
What do you do / study?

School-age students
Do you study English at school?
Do you like it?

Thank you.

(Repeat for Candidate B.)

Back-up prompts

How do you write your family / second name?

Do you live in ...?

Have you got a job?
What job do you do? / What subject(s) do you study?

Do you have English lessons?

Phase 2
Interlocutor

(Select one or more questions from the list to ask each candidate. Use candidates' names throughout. Ask Candidate B first.)

Do you enjoy studying English? Why (not)?

Do you think that English will be useful for you in the future?

What did you do yesterday evening / last weekend?

What do you enjoy doing in your free time?

Thank you.

(Introduction to Part 2)
In the next part, you are going to talk to each other.
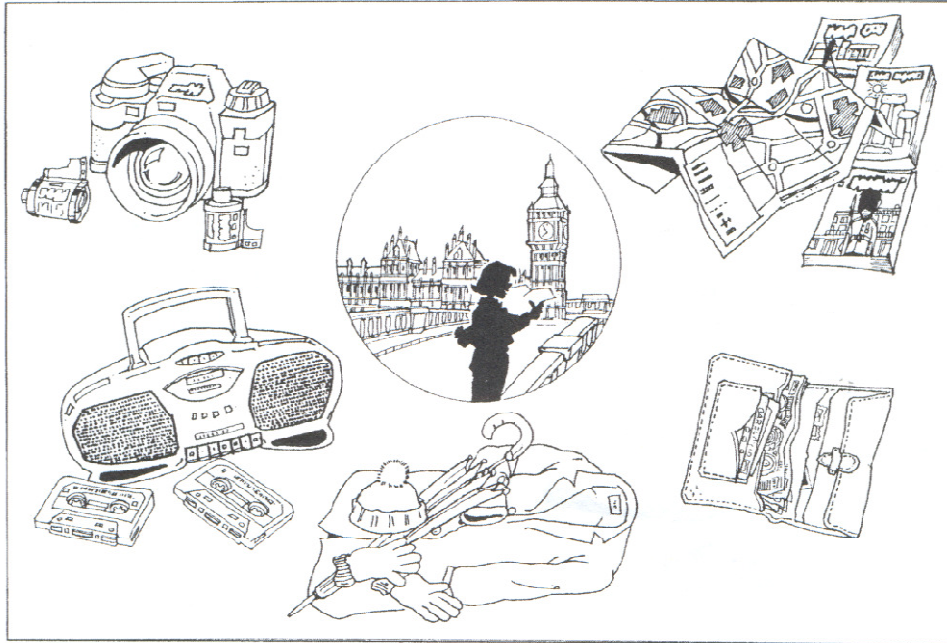
Part 2 (2-3 minutes)

Trip to England

Interlocutor
Say to both candidates:

I'm going to describe a situation to you.

A friend of yours is planning to spend 6 months in **England** to improve her English. Talk together about the things she will **need** in England, and decide which are the most important things to **take / bring** with her.

Here is a picture with some ideas to help you.

*Place **Part 2 booklet**, open at the **Task**, in front of candidates.*

Pause

I'll say that again.

A friend of yours is planning to spend 6 months in **England** to improve her English. Talk together about the things she will **need** in England, and decide which are the most important things to **take / bring** with her.

All right? Talk together.

Allow the candidates enough time to complete the task without intervention. Prompt only if necessary.

Thank you. (Can I have the booklet please?)

Retrieve Part 2 booklet.

Time   About **2-3 minutes** (including time to assimilate the information)

3

PAPER 3: SPEAKING
Part 3

People reading and writing

Part 3 (3 minutes)

Interlocutor
Say to both candidates:

Now, I'd like each of you to talk on your own about something. I'm going to give each of you a photograph of people **reading and writing**

Candidate A, here is your photograph. *(Place **Part 3 booklet**, open at **Task 1A**, in front of Candidate A.)* Please show it to Candidate B, but I'd like you to talk about it. Candidate B, you just listen. I'll give you your photograph in a moment.

Candidate A, please tell us what you can see in your photograph.

(Candidate A)

Approximately one minute

If there is a need to intervene, prompts rather than direct questions should be used.

Thank you. (Can I have the booklet please?)

Retrieve Part 3 booklet from Candidate A.

Interlocutor

Now, Candidate B, here is your photograph. It also shows someone **reading** and **writing**. *(Place **Part 3 booklet**, open at **Task 1B**, in front of Candidate B.)* Please show it to Candidate A and tell us what you can see in the photograph.

(Candidate B)

Approximately one minute

Thank you. (Can I have the booklet please?)

Retrieve Part 3 booklet from Candidate B.

Part 4 (3 minutes)

Interlocutor
Say to both candidates:

Your photographs showed people **reading** and **writing**. Now, I'd like you to talk together about the different kinds of reading and writing **you** did when you were **younger**, and the kinds you do **now**.

Allow the candidates enough time to complete the task without intervention. Prompt only if necessary.

Thank you. That's the end of the test.

Time   Parts 3 & 4 should take about **6 minutes** together

APPENDIX D: FCE Speaking Mock Test

PAPER 5 : SPEAKING
Part 1

For Oral Examiners' Use Only

Education and Work

- Can you remember your first English lessons? ..... What were they like?
- What do you think were the most important things you learned at primary / elementary school?
- Do you plan to study anything in the future?
- Would you prefer to work for a big or a small company? ..... Why?
- What do you think would be the most interesting job to do?

Leisure and Entertainment

- Do you like reading books? ..... What sort of books do you enjoy reading most?
- What sports do people play most in your country? ..... And what do people enjoy watching?
- Is it easy to meet new people where you live?
- Do you normally go out with family or friends?
- What do you enjoy doing with your friends?
- Where's the best place to spend a free afternoon around here/in your town?
- How expensive is it to go out in the evening where you live?
- Do you like going to the cinema?
- Tell me about your favourite filmstar.
- What are you going to do this weekend?

The Media

- How do you find out what's happening in the world?
- Do you ever listen to the radio? ..... What programmes do you like?
- How important is TV to you?
- Do you like the same TV programmes as your parents?
- What's the difference between reading the news in the newspaper and watching it on TV?
- Do you think computers will replace newspapers and TV in the future?

Travel

- How do you prefer to travel, by train or plane? ..... Why's that?
- What's the longest journey you've ever been on?
- What's public transport like in your country?
- Where did you spend your last holiday? ..... What did you do?

©UCLES 2002

For Oral Examiners' Use Only

First Certificate in English
Speaking Test

2002 Test Materials

Interlocutor

Good morning / afternoon / evening. My name is .... and this is my colleague .... He / she is just going to listen to us.

And your names are .... ?

Could I have your mark sheets, please?

Thank you.

(Hand over the mark sheets to the Assessor.)

First of all we'd like to know something about you, so I'm going to ask you some questions about yourselves.

either: (non UK based candidates)   *or: (UK based candidates)*

(Candidate A), do you live in ......?   Where are you from *(Candidate A)*?
(name of town where examination is being held)

And you *(Candidate B)*?   And you *(Candidate B)*?

- What do you like about living *(here / name of candidate's home town)*?
- And what about you *(Candidate A/B)*?

(Select one or more questions from any of the following categories as appropriate.)

Homelife

- Could you tell me something about the area where you grew up? ..... What did you like about living there?
- How much time do you spend at home nowadays?
- What do you most enjoy doing when you're at home?
- Could you describe your family home to me?

Personal Experiences

- Who are the most important people in your life?
- Do you and your friends share the same ideas?
- Tell me about your best friend.
- What's the most exciting thing you've ever done?
- Is there anything you'd love to be able to do in the future?
- What's your favourite day of the week? ..... Why?

©UCLES 2002

PAPER 5 : SPEAKING
Part 2

17

18

**First Certificate in English
Speaking Test**

17 Special moments
18 Leisure activities

Part 2 (4 minutes)

2002 Test Materials

|	
Interlocutor [17]	Now, I'd like each of you to talk on your own for about a minute.
	I'm going to give each of you two different photographs and I'd like you to talk about them. (*Candidate A*), here are your two photographs. **They show people enjoying special moments in their lives.**
	[*Hand over picture sheet 17 to (Candidate A).*]
	Please let (*Candidate B*) see them.
	(*Candidate B*), I'll give you your photographs in a minute.
	(*Candidate A*), I'd like you to compare and contrast these photographs, and **say which of the people you think will remember this moment the longest.**
	Remember, you have only about a minute for this, so don't worry if I interrupt you. All right?
Candidate A	....
⏲ 1 minute	
Interlocutor	Thank you. [*Retrieve photographs*]
	(*Candidate B*), **do you like doing dangerous things?**
Candidate B	....
⏲ 20 seconds	
Interlocutor	Thank you.
[18]	Now, (*Candidate B*), here are your two photographs. **They show people doing different activities in their free time.** Please let (*Candidate A*) have a look at them.
	[*Hand over picture sheet 18 to (Candidate B).*]
	I'd like you to compare and contrast these photographs, and **say how much you would enjoy doing activities like these.**
	Remember (*Candidate B*), you have only about a minute for this, so don't worry if I interrupt you. All right?
Candidate B	....
⏲ 1 minute	
Interlocutor	Thank you. [*Retrieve photographs*]
	(*Candidate A*), **do you ever do activities like these?**
Candidate A	....
⏲ 20 seconds	
Interlocutor	Thank you.

PAPER 5 : SPEAKING
Parts 3 and 4
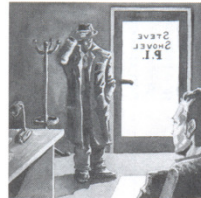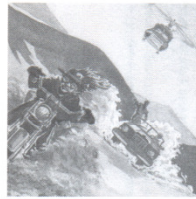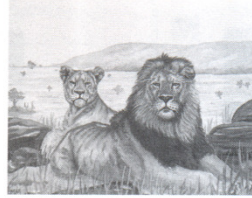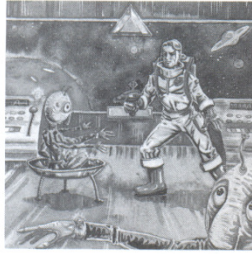
30

First Certificate in English
Speaking Test

30 Film club

Parts 3 & 4

2002 Test Materials

Part 3 (3 minutes)

Interlocutor Now, I'd like you to talk about something together for about three minutes. I'm just going to listen.

The film club at your school has asked you to choose two films which would be interesting for the students to watch and then discuss. Here are the films they are considering.

[Place picture sheet 30 in front of the candidates]

First, talk to each other about how interesting these different types of film would be. Then decide which two would be the best for students to discuss.

You have only about three minutes for this. So, once again, don't worry if I stop you, and please speak so that we can hear you. All right?

Candidates
⏱ 3 minutes
....

Interlocutor Thank you.

[Retrieve picture sheet 30]

Part 4 (4 minutes)

Interlocutor *[Select any of the following questions as appropriate]*

* How popular do you think a club like this would be?

* What sort of films do you never watch? ..... Why?

* Are there any films that you'd like to see again? ..... Why (not)?

* Would you prefer to be in a film or behind the camera?

* How important do you think it is to watch films in English?

* What can you learn about a country's culture by watching films from that country?

Thank you. That is the end of the test.

[Check that all test materials have been replaced in the file]

APPENDIX E: Speaking Assessment Mark Sheet used in this study

CANDIDATES' MARK SHEETS

Teacher's codename: _____

…………………………………………………………………………………………

PET Speaking: Listen to 2 candidates taking a PET Speaking Paper and assess each of them according to the following table, by ticking the box which corresponds to the grade awarded. You will be asked to list and define the criteria used afterwards.

Candidate	5,0	4,5	4,0	3,5	3,0	2,5	2,0	1,5	1,0	0,5	0,0
PET1											
PET2											

…………………………………………………………………………………………

FCE Speaking: Listen to 2 candidates taking an FCE Speaking Paper and assess each of them according to the following table, by ticking the box which corresponds to the grade awarded. Once again, you will be asked to list and define the criteria used afterwards.

Candidate	5,0	4,5	4,0	3,5	3,0	2,5	2,0	1,5	1,0	0,5	0,0
FCE1											
FCE2											

APPENDIX F: Speaking Assessment Mark Sheet used at the language institute

Candidate name _____

Candidate Number _____

Used material _____

	5,0	4,5	4,0	3,5	3,0	2,5	2,0	1,5	1,0	0,5	0,0
Grammar and Vocabulary											
Discourse Management											
Pronunciation											
Interaction											
Global Achievement											
Total Mark											

APPENDIX G: Questionnaire to Assessors

 QUESTIONNAIRE TO ASSESSORS

1. Did you use the same criteria to assess the PET and FCE students in this study?

2. Which criteria did you use to assess the students in the recordings in each level (PET and FCE)?

3. Define conceptually each criterion used to assess the students in the recordings.

4. How much is each criterion used worth?

5. Are the criteria used in this study the same as the ones used by you with students in these levels during a regular term?