

## ORAL LANGUAGE EXAMINATIONS AND ROMANIAN EFL STUDENTS ~ Case Study ~ B2 First for schools – Cambridge English Qualifications

**Andreea Paula LAURENȚIU (KOLARIK)**  
**PhD student<sup>1</sup>, “1 Decembrie 1918” University of Alba Iulia**

*Abstract: Norman Fairclough’s belief is that discourses represent the world as it is perceived by the individuals, personalised according to their own life experience, social or cultural backgrounds. (Fairclough 2003, pp. 124-125).*

*Discourse, one of the most powerful tools of the 21<sup>st</sup> century society, is the focal point of my investigation. Consequently, the element under investigation in my research is the discourse produced by Romanian students in two specific social contexts: that of Oral Language Examinations (OLE) and that of a Public Speaking Program (PSP) Ted-style talk to prove that the latter can positively influence the outcomes of the first.*

*Towards that end, this research, even though in its initial stage, has as core focus the analysis of the language used by three groups of Romanian high school students (15-19 of age) while sitting Cambridge, IELTS, TOEFL and Baccalaureate OLE before and after attending a TED-style (PSP).*

*As part of this ample investigation this article aims to establish through a discursive analysis whether the validity of one of the aforementioned OLEs - B2 First for schools or previous oral Cambridge English: First (FCE) examination, can be and/or should be questioned. The arguments brought by this case study show that two candidates who have never taken such OLEs before, who have not attended a PSP yet and who have been previously tested but have demonstrated an upper-intermediate level of English might not score the same in the oral B2 First examination unless they are systematically prepared and trained on solving each and every part of the exam format.*

*After the recording of the two students while taking the speaking test and after the analysis of the resulting transcript we hereby open a Pandora’s box and demonstrate that a maximum score cannot be obtained unless serious training is conducted beforehand and parents’ money is invested in both tutors and in training materials. The linguistic demonstration has been conducted by comparing the transcript resulted from the two students who were not familiar with the exam format with the transcript resulted from one of the video samples provided online by the test developer.*

*It also needs to be taken under advisement the possibility that preparation for such examinations, essential in students’ academic endeavour, should be incorporated in the school curriculum and taught as optional courses either by the school teachers or by test developers, but free of charge. The discursive analysis shows how the candidates, despite being briefly instructed how to approach all four parts of the exam, do not manage to capitalize their knowledge regarding the English language.*

*Keywords: discourse analysis, validity, Cambridge exams, oral tests, conversation analysis, B2 First for schools*

### **Literature Review**

Language examinations have long been considered essential historic evidence that provide information about the language classes back in the days. English as a

---

<sup>1</sup> Scientific Coordinator: Teodora Iordăchescu, Professor, PhD “1 Decembrie 1918” University of Alba Iulia

Foreign Language (hereinafter EFL) Cambridge examination, CPE Certificate of Proficiency in English (hereinafter CPE), for instance can be identified as among the first formal language examinations going back to 1913. Even though the exam was organized to last up to 12 hours, it was the first serious attempt to accurately and thoroughly assess the candidates' competences and language skills including translation from English to German/French and vice versa, English essays, English literature and English phonetics. The oral exam consisted of dictation, reading and conversation. The purpose of this exam was for people older than 20 years of age to become teachers of English.

Nowadays, the very thought of such long and exhausting EFL examination might generate a lot of debate and raise countless concerns. Starting back with 1858 in Britain at the University of Cambridge Local Examinations Syndicate (UCLES) the Cambridge Assessment was followed one hundred years later by the USA with the development of TOEFL (Test of English as a Foreign Language) back in 1962. Throughout the years, the optics changed, therefore a full development of major examinations can be observed at all levels of language. Starting from the mid-18<sup>th</sup> century the language assessment industry flourished constantly up until today when such examinations are taken by millions of candidates all over the world. In their pursuit of reliability and validity in test development, the tests developers have struggled to adapt the test format and content to the realistic needs of modern society. This paper is an essential part of a larger research conducted on official Oral Language Examinations (OLE) such as Cambridge (B2 First and CAE), IELTS (International English Language Test) and TOEFL as well as on the Romanian national examination, the Bacalaureate.

This article focuses on the B2 First examination, also known as Cambridge English: First (FCE), which is only one of the Cambridge English Qualifications available worldwide. According to the test developers it is the most popular exam which is accepted by educational institutions and employers around the globe. The Common European Framework of Reference for languages (hereinafter CEFR) which is the international standard for describing language ability provides a useful scale for all levels of language making it easy both for employers as well as for language teachers to interpret the different levels of proficiency. CEFR provides six levels starting from beginner – A1 and up to C2 – proficient user which are applicable to any language. Annex 1 presents the levels of proficiency in detail for a better understanding and reference.

Moreover, the discursive approach to be used in this article has as a starting point the uses of discourse analysis presented by Barbara Johnstone in her book, authored in 2008, entitled *Discourse Analysis* in order to establish “*what happens in a talk*” (2008, p. 6). Moreover, she states that “*language seems to be created by speakers as they interact, noticing, repeating and sometimes making reflective generalizations about what other people do, in the process of evoking and creating a **world***” (2008, p. 268). Last but not least, it is of utmost importance to consider that “*discourse means imagining that ‘language’ could exist prior to being ‘used’, but in fact people learn how to talk by talking*” (2008, p. 268)

The study is also grounded on the theory launched by Norman Fairclough in 1989 with all its subsequent editions within the book entitled *Language and Power*, where he defined language as “a part of society”, or “ a social process” (1989, p. 22) and he also stated that the “idea of ‘power behind discourse’ is that the whole social order of discourse is put together and held together as a hidden effect of power” which is worth investigating, given the fact that the third section of the oral examination under analysis involves a collaborative task where two or three candidates are involved. Furthermore, written in 2003, his work entitled *Analysing Discourse* defines discourse as an important part of social practices and identifies genres and representations of the speakers’ own world also offering more information regarding discursive analysis as well as conversation analysis.

Last but not least, validation of oral examinations is one more issue under investigation within the thesis this article is a constituent part. Therefore, by using the evidence-based approach presented in *Language Testing and Validation* authored by Cyril J. Weir (2005) corroborated with the ample work edited by Alister Cumming and Richard Berwick published in 1996 on *Validation in Language Testing* we have managed to have a clear image on how validity should be achieved in language testing. Furthermore, *Qualitative Approach to The Validation of Oral Language Tests* prepared by Anne Lazaraton in 2002 guided this attempt to raise awareness on the fact that research possibilities on Cambridge speaking tests validation are limitless and are worth pursuing in order to make sure that all the research perspectives are covered and supported by relevant data. Therefore, by conducting this recorded mock Oral Language Examination (hereinafter OLE) this paper has succeeded to demonstrate its initial theory that the personal characteristics of the candidates such as age, gender as well as their knowledge of the exam format are essential elements that must not be overlooked since they might negatively impact their performances. For this particular reason we have initiated our investigation by starting off from the Conceptual framework for performance testing, a model designed by Milanovic and Saville (1996, p. 16) as presented in Anne Lazaraton’s book (2002, p. 118), presented below in Figure 1. The framework designed by Milanovic and Saville has been chosen because it can be considered as one of the most comprehensive and most inclusive models.

### ***Procedure***

This article analyses the texts resulted after transcribing the OLE conducted by one examiner and two candidates while taking an online mock oral examination and the transcript of a video provided by the test developer as sample for the B2 First speaking test. The quantitative analysis is going to be a contrastive one as shown in Table 1 where the results of the two candidates are compared to the results obtained from the text resulted from transcribing a video recording with two candidates which was posted on You Tube by Cambridge Essentials for future candidates’ reference together with the examiner’s comments.

### ***Examination format***

The mock OLE was conducted by using a sample paper available on Cambridge Examinations website for the B2 First level. The examiner, the students' English teacher, who is a non-native speaker of English, is the person administering the test using the free online Zoom platform to conduct the recording. Both parents of the candidates have signed parental consents to allow their children to participate in this research.

### **Candidates**

The candidates a 15-year-old girl (hereinafter C.O.) and a 14-year-old boy (hereinafter B.M.) taking this OLE are both in the 9<sup>th</sup> grade, in Mathematics – Informatics, bilingual English class at the same school. Both candidates have been exposed for the first time to the B2 First OLE format and they have never taken any type of OLEs before. The initial placement tests taken by both of the candidates at the beginning of the school year in September 2019 showed an upper intermediate level.

Before the recording of their performance, they have been briefed about the exam format and what each of the four parts of the oral test require. All their questions regarding the exam format have been answered and only when everything was clear did the examiner start the test.

### **The exam**

According to the official website the exam format is as follows:

*“A face-to-face test taken with one or two other candidates and an examiner. Students have to show how well they can produce spontaneous spoken language, talking with either the examiner, the other candidate, or by themselves.”<sup>2</sup>*

The entire OLE lasts for about 14 minutes and is aimed at assessing the candidates' ability to use general interactional and social language, organise a larger unit of discourse (comparing, describing, expressing opinions), sustain an interaction, exchange ideas, express and justify opinions, etc.

The B2 First OLE consists of four parts as per Annex 2.

#### **Part 1. 2 min**

*“Conversation with the examiner. The examiner asks questions and students may have to give information about themselves, talk about past experiences, present circumstances and future plans.”<sup>3</sup>*

Consisting of three to five individual questions is meant to familiarize the students with the exam atmosphere, the examiners, as well as with the other candidate/s. The exam format allows the organizers to have groups of three candidates provided that the number of candidates is uneven. According to the exam overview, all of the questions asked in this section involve verbal questions where the candidate must be able to provide personal information, talk about present or past experiences and events as well as about future plans.

#### **Part 2. 1 minute per candidate, plus a 30-second response**

<sup>2</sup> <https://www.cambridgeenglish.org/exams-and-tests/first-for-schools/exam-format/>

<sup>3</sup> <https://www.cambridgeenglish.org/exams-and-tests/first-for-schools/exam-format/>

*“The examiner gives the student a pair of photographs to talk about and they have to speak for 1 minute without interruption. The questions about the photographs are written at the top of the page to remind the student what they should talk about. When they have finished speaking, the student’s partner then has to answer a short question from the examiner about their photographs.”<sup>4</sup>*

Each of the candidates receives a set of two pictures and are supposed to speak for about 1 minute giving information, expressing opinions and relating the candidate’s personal experience to the images. The candidates also receive a follow-up question regarding the other candidate’s questions. The answer should be no longer than **30 seconds**.

**Part 3.** *“A 2-minute discussion followed by a 1-minute decision-making task”*

*“Conversation with the other candidate. The examiner gives the students a question and some written prompts. The students discuss these together for two minutes. The examiner will then ask them to make a decision together about the topic they have been discussing.”*

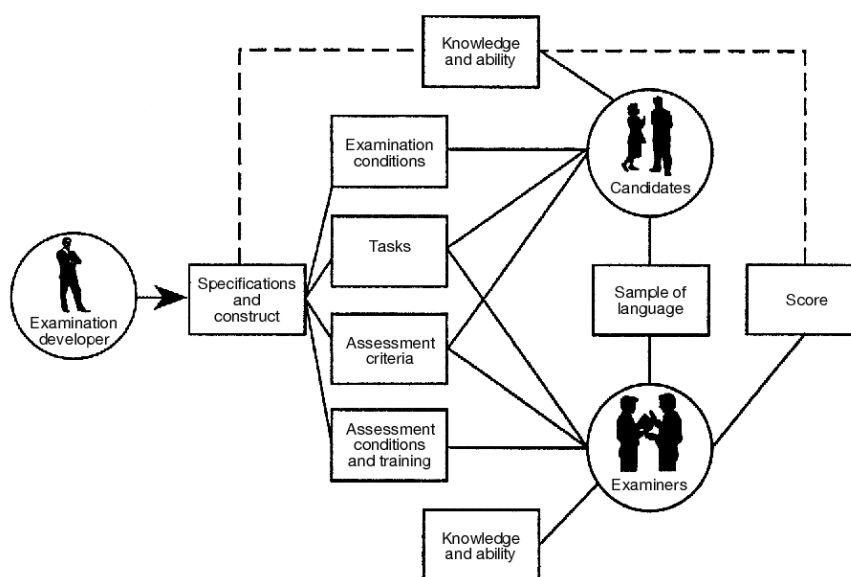
Starting from written prompts the candidates have to solve a task collaboratively. The candidates receive verbal instructions as well as written stimuli and they must exchange information and opinions, express and justify opinions, agree or disagree (partly or fully), make suggestions, and speculate. They have to complete this task in **2 minutes** and then they are allotted another **1 minute** to reach a decision.

**Part 4.** *“The discussion should last 4 minutes”*

*“Further discussion with the other candidate, guided by questions from the examiner, about the same topic as the task in Part 3.”*

Continuing with the same topic as in Part 3 the candidates have to answer general opinion questions starting from verbal prompts and they have to exchange information and opinions, to justify their opinions and to disagree, agree or partly agree.

### Data analysis



<sup>4</sup> <https://www.cambridgeenglish.org/exams-and-tests/first-for-schools/exam-format/>

**Figure 1<sup>5</sup>.** A conceptual framework for performance testing in OLE (Milanovic and Saville, 1996, p. 16)

To start with, the framework presented by Milanovic and Savile (1996, p. 16) offers myriads of research possibilities. However, given the exam under analysis, we shall apply the abovementioned model on this particular test alone – the oral part – and on the pair of candidates. As provided by the model from Figure 1, the variables to be included in processing the data relevant to this OLE are the candidates, the examiner, the assessment principles and the tasks, together with the interrelations existing between these variables. These variables have been considered while designing this test in order to assess the performance of candidate without any threats connected to the validity or reliability of the test. As it results from the diagram above there are three stages:

- The first stage where the examination designer develops the speaking test;
- The second stage, test administration, the candidates provide samples of language corresponding to the tasks prompted by the examiner according to the descriptions from the Exam format section of this paper;
- The third, and the last stage is marking the candidates' performance as per their knowledge and ability.

Due to the fact that there have been numerous studies to discuss the examiners' role within the oral exam as well as their impact on the candidates' performance during this type of oral examination, this study shall concentrate on other variables such as: the candidates' personal characteristics and, as presented in Figure 2, inserting Exam format training for the candidates. In order to justify the existence of these alterations the following remarks are to be made:

- ❖ Data on candidates' personal characteristics should be collected to help test developers design more reliable tests. Therefore, pre-testing the candidates' language level should be considered in order to increase the test's validity and reliability. Tests are not always able to predict the future performance of the candidate in the actual test situation. The washback effect or the impact of the exams on the candidates appears (Cheng and Curtis, 2004, p. 3-17) and its dimension can influence the candidate's learning but it can also influence the teaching. Charles Alderson from the Department of Linguistics and Modern English Language, Lancaster University in a paper presented in a Conference in Barcelona in July, 2001 wrote:

*"Ten years ago, washback was a common concept, and the existence and nature of washback was simply accepted without argument. Tests affect teaching. Bad tests have negative effects on teaching; more modern, good tests will have positive effects; therefore, change the test and you will change teaching." (Ed. Milanovic and Weir, 2001, p. 2)*

<sup>5</sup> Permission request sent for the use and amendment of figure 1 as presented in "A Qualitative Approach to the Validation of Oral Language Tests" published by Cambridge University Press in 2002

If 20 years ago the recommendation was to change the tests and the teaching would follow suit, why hasn't it materialised? The change did not take place. The vast majority of EFL teachers in Romanian schools teach from coursebooks that were published at the same date with the quote above. Thus, the question is still there, necessary to be reiterated over and over again.

Therefore, the purpose of this article is to raise awareness on the fact that official language examinations such as the one in discussion here - B2 First Cambridge examination - are firstly meant to augment the sales of books and training materials for publishers all around the globe rather than caring about their clients and their needs when the actual need is to help potential future candidates have access to materials and teaching methods that could improve their exam performance. It is of course a matter for discussion and debate for those schools where the parents cannot afford to purchase coursebooks that have adapted their curriculum for this type of exam. The major problem is that the Ministry of Education does not see that the national curriculum should really be adapted to the needs of Romanian students. It is mandatory to clarify that it is not only the case with the Cambridge examinations. IELTS and TOEFL are also included here but the discussion is to be addressed in a different paper. Hence, considering all the aforementioned information corroborated with the data analysed in this specific case study and with the increasing numbers of students willing to take such examinations it is of utmost importance to compel the relevant authorities to implement this format into the national curriculum or even adjust and amend it in such a manner to allow students who are willing to take such official language examinations to be able to take them without this spectacular financial effort from the part of their parents in hiring tutors and purchasing materials.

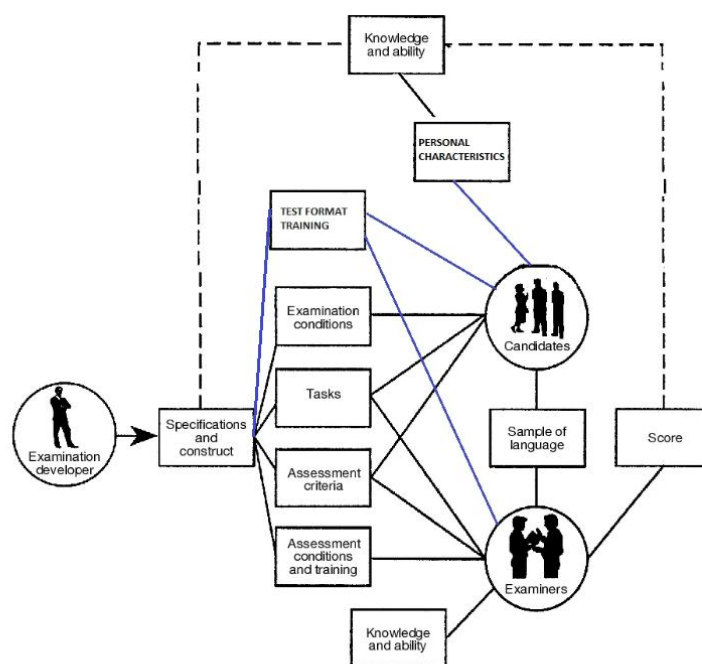
The most efficient method would be to encourage exam developers to distribute their materials for free in schools and even to train the teachers to prepare students for their exams as a separate subject or only for those students who want to take such examinations at the end of their academic cycle.

❖ The second amendment involves the test format training

It is of utmost importance to underline the fact that all other types of standardized tests such as an ordinary driving license or a certificate to become Cambridge, IELTS or TOEFL examiner involve training sessions on how to successfully take the final test. In the case of the language tests, students are on their own. Either rely on their own ability to master all the skills necessary to take the test and learn from different available sources or they rely on their classroom teacher who may or may not have access to updated materials for the test or who does not know the test format at all.

It is this last situation that is described in this study. A pair of candidates who took their Cambridge OLE for the very first time without any prior preparation. Supposedly, their level of English should be enough to pass the test. But as we all know, this is not enough and every candidate wants to obtain a good score to be able to use the certificate in order to enter a university, get a job or a better job or even to apply for a

grant, an internship. For all these reasons we have made the adjustments as presented in Figure 2 below. All the changes are marked with blue lines.



**Figure 2.** Amended conceptual framework for performance testing (Milanovic and Saville, 1996, p. 16) in OLE including test format training for the candidates as well as analysis of candidate personal characteristics.

Starting from the first stage of the model created by Milanovic and Saville, that is the designing of the test, the test format training for the future potential candidates should be considered in order to anticipate who are the potential clients, and how to better design the tests to assess their skills and abilities. Moreover, the personal characteristics have to be considered also, in order to form well balanced pairs in the oral test and to make sure that the candidate has chosen the right exam to test his/her language skills. These two amendments to the initial model can be implemented by organizing training sessions for the future candidates as well as organizing free mock examinations to help the students choose their appropriate test level.

### **Data analysis**

The data under analysis here consists in a text resulted after the transcription of an online mock oral examination conducted by one examiner (the candidates' English teacher) and two candidates C.O. and B.M. students in the same class and having the same level of English. This test was recorded in November 2020. The length of the recording is 11 minutes and 20 seconds showing the candidates while solving the four tasks associated to B2 First OLE. What needs to be kept in mind is the fact that both students have never taken any type of OLEs before and their level of English was measured before this speaking session by using a placement test where C.O. scored 86 out of a total of 100 and B.M. scored 83 therefore, after the interpretation of results they

both master the English language at B2 level (upper-intermediary level according to the Common European Framework of Reference for Languages (hereinafter CEFR) (according to the placement test the maximum score is 100 representing C2 level-proficient)).

Before the administration of the oral test the examiner presented the number of parts involved in this type of OLE and what each part required. Only after both candidates understood what they had to do, did the test begin.

From the very beginning it can be noticed the fact that the time necessary to cover all the 4 parts is only 11 minutes of actual speech while the official format usually covers about 14 or even 15 minutes. The reason why this happened was because the two candidates did not practice this type of long turn speaking and were not aware neither of how much time they actually did spend speaking nor the remaining time thus they did not know if they had to add more information or not to their answers. Out of the four parts, Part 2 and Part 3 are the ones our investigation is focused on because they assess Discourse Management and Interactive Communication.

***Candidate - C.O., girl, 15.***

She shows a good degree of control of a range of simple and even more complex grammatical forms using a good range of appropriate vocabulary; in Part 2 she covers the topic pretty well speaking with no interruption or hesitation for about 20 seconds with accurate stretches of language. She stops for about 8 seconds and only resumes her discourse after the examiner's deviation (hand gesture) signalling the fact that she has to keep on speaking, consequently she speaks for another 30 seconds ending with a question directed to the examiner asking for the translation of a word and saying that she does not know the English word for "autocar" when she ultimately abandons the topic altogether. However, this part (the one produced by the candidate after the examiner's interference) has not been taken into consideration in the overall assessment of the C.O.'s performance across the whole investigation. Even though C.O. managed to speak with little or no hesitation it was obvious that her discourse suffered precisely because she was not familiarized with the format and the whole setting as well as the type of task were new to her despite the fact that the exam format was clearly presented at the very beginning. The consequence of her shorter response in Part 2 is that the amount of language produced is less than it could have actually been thus there is a smaller amount of language to be evaluated and fewer chances to impress the examiner with rich vocabulary, fluency and the quality of the discourse.

Yet again, the follow-up question received an 11-second forthright answer, without any examples to support C.O.'s ideas and to allow her to demonstrate her fluency and lexical richness. A longer answer might have offered the candidate the possibility to demonstrate her ability to use the language while expressing personal ideas and opinions on a familiar topic thus scoring even higher.

*[C.O.] I visit museums often, but not as much galleries, and because I find them very interesting and stimulating for the brain. [11sec]*

***Candidate - B.M., boy, 14***

In Part 2 he managed to speak for a full minute despite some hesitation, several disruptions and certain grammatical and vocabulary inadvertencies showing a good degree of control of simple grammatical forms appropriate use of vocabulary to express his views on a range of familiar topics.

*[B.M.] ...the woman is (pause) probably looking for some clothes, what she likes or wants to buy for someone else (pause) and (pause) in the second picture (pause) she (pause) looks very (pause) happy and calm (pause) looking at something that maybe she likes and in the second (pause) [aaa] first picture, they seem very serious, (pause) maybe they are, they are (pause) there are rules, so they have to keep the(pause) silence in a museum. [1 min]*

The answer given to the follow-up question by B.M. lasted for 14 seconds, thus a longer unit of discourse would have been appropriate in order to accurately address the task and score high. Even though the answer was grammatically and functionally accurate, it lacked details and examples that might have provided a full and rich answer that could have secured a maximum score.

*[B.M.] I never travel by plane (pause) anywhere, when I go on holidays with my family, usually we go with our own car. [14s]*

After being awarded 15 seconds to look at the task, the candidates have to talk to each other and collaboratively discuss about the items presented in Part 3 and answer a question. From the very beginning C.O. proves to be a strong speaker, initiating and discussing with a lot of ease about all the items given in the task and complementing nicely the answers provided by her partner. Both candidates have quite short contributions and the turn taking pace was quite energetic. However, they jointly decided to stop after 1 minute and 30 seconds even if the discussion seemed quite vivid and flowing. They both could have extended some more in order to offer longer stretches of fluent speech for the examiner to assess. Being engaged into the collaborative task they both brought relevant contributions with little repetition, using appropriate linking devices and clear and intelligible pronunciation and appropriate intonation.

With respect to the interactive communication, they both engage into the communication appropriately, bringing their own ideas into the verbal exchange. Moreover, the stress is accurately placed and individual sounds are articulated clearly.

*[C.O.] May I start? Ok. I think sleeping eight hours every night is very important because not sleeping can actually reduce the number of neurons you have in your brain. What about you?*

*[B.M.] I think that eating at regular times is one of the most important things since you let your body accommodate to precise hours, when you can prepare to eat.*

*[C.O.] Also, visiting the doctor has to do with eating, since you can also visit a nutritionist, maybe, to help you eat healthier [ahem](pause) and visiting the doctor, like a normal doctor, even a dentist, also help you with possible health problems.*

*[B.M.] Yeah, and people really would recommend you to spend time outdoors each day, because it can be refreshing and can calm you.*

*[C.O.] And you also can get a lot of vitamin D from the sun and strengthen your bones or go to the gym and strengthen your muscles and, you know, just cardio. Do cardio and stuff.*

*[B.M.] Yeah (longer pause)*

*[Examiner] Ready? Ok. Now, you have about a minute to decide which two are most important for keeping fit in the long term. You have a minute. You have to talk to each other and decide.*

*[C.O.] [aham] I think that eating at regular times and going to the gym, because these two go hand in hand to keep you fit and eating at regular time can also mean eating meals that do not necessarily have a lot of fat and carbs because you couldn't eat that many, so I think these are the most important*

*[B.M.] Yeah, I think the same, because visiting the doctor you don't really need to, because you can search on internet and by going to the gym you can actually spend time each day outside and eating at regular times even helps you to maintain a good sleep pattern.*

The two candidates have produced a clear discourse with little hesitation, good control of simple as well as more complex grammatical forms and good and appropriate range of vocabulary while exchanging views on this familiar topic regarding things people do to keep fit and healthy.

Even though B.M. displayed good vocabulary and prompt answers, his third intervention was only a one-word answer where he could have extended his answers more. Moreover, he did not invite C.O.'s response at all and he only provided answers as responses to C.O.'s contributions thus making the conversation relevant but one-sided.

*[C.O.] And you also can get a lot of vitamin D from the sun and strengthen your bones or go to the gym and strengthen your muscles and, you now, just cardio. Do cardio and stuff.*

*[B.M.] Yeah.*

However, in the last section of Part 3 the answers of both candidates are well-balanced bringing arguments and longer stretches of language to support their ideas. The same can be said about Part 4 where while answering the examiner's questions they managed to expand more on the topic in discussion.

### **Discourse analysis**

The candidates' discourses were consistent with the tasks, providing pertinent, straightforward answers to the examiner's questions. Throughout their speeches they have both used linking devices, appropriate discursive strategies as well as some discourse markers thus employing rather argumentative styles. However, given the format of the exam, and the fact that they have never taken such oral exams before, we can infer that the amount of produced discourse would have been significantly larger and fluency would have been proven even better through longer stretches of speech if they had been trained for it accordingly.

According to Halliday and Hasan (1976) the linguistic instruments necessary to organise text into meaningful units of language are the following: reference, ellipsis,

conjunction, as well as lexical cohesion are. For this reason and for the purpose of this paper we have identified the cohesive devices used by the speakers, the verbs, the pronouns in order to conduct an accurate discourse analysis and we have managed to do that by using the free software Tropes.

### **Connectors**

**C.O.** used connectors such as *so* (2 instances) and *because* (5 instances) to express cause, addition was expressed by using *and* (17 instances), disjunction by using *or* in 4 instances, opposition was expressed by using *but* (3 instances) and *otherwise* once. C.O. Used *like* once to make a comparison and once C.O. used *like* as a discursive marker. Time was expressed twice by using *since* and *while*.

**B.M.** also used a fair number of connectors to express cause with *because* (6 instances) and *so* (1 instance), addition was also presented with 12 instances of *and* while disjunction by using *or* (2 instances) and time by using *when* (3 instances) and *since* (1 instance)

### **Verbs**

The verbs used by **C.O.** were 35.1% factive (26 occurrences among which *play, draw, make, strengthen, and do*), 36.5% stative (27 occurrences such as *seem, have, can, keep, and be*) and 28.4% reflexive (28 occurrences such as *think, might, would, may, decide, and want*)

The verbs used by **B.M.** were 39.7% factive (29 occurrences among which *do, go, prepare, eat, accommodate, spend, help and spend*), 31.5% stative with 23 occurrences such as *are, can, get, have, seem, and keep*, 8% reflexive verbs with 6 occurrences such as *find, like, let, think, and recommend*, and 2.7% performative with 2 occurrences *agree* and *maintain*.

### **Pronouns**

The personal pronoun *I* was used by C.O. in 17 instances (48.6%), *you* in 12 instances (34.3%) and *they* in 2 instances (5.7%) denoting a strong speech that involves and captures the audience.

The personal pronoun *I* was used by B.M. in 9 instances (23.1%), *you* in 3 instances (7.7%), *we* was used one time (2.6%), *they* in 10 instances (25.6%) and *somebody* once (2.6%).

The use of the first personal pronoun *I* in both of the cases indicates that the speaker makes reference to his/her personal experiences in order to make himself/herself better understood and to express his/her opinions, emotions and beliefs of the world. The use of the second personal pronoun *you* shows the fact that the speakers have the intention to engage the listener or the interlocutor into their speech generating an inclusive thus engaging discourse.

### **Discussion**

In order to prove our case further, we have also analysed the transcript of a Cambridge English: First Speaking (from 2015)<sup>6</sup> Sample test, where two candidates are recorded while sitting the oral exam. Their performance assessment is not included in

<sup>6</sup> <https://youtu.be/EdeZp0n0JHw>

this paper because it can be found together with the examiner's comments and the recommendations online<sup>7</sup>. Hence, for this particular investigation we have extracted the quantitative data to demonstrate our hypothesis.

Consequently, by using the software [textinspector.com](http://textinspector.com), we have managed to quantify the data resulted from the transcribed texts into data that demonstrate that the two candidates recorded and presented within the video displayed on YouTube by Cambridge English with the associated examiner's comments have significantly higher numbers of sentences and words during their overall performance within the oral test. As it is presented as a sample video it can be inferred that the candidates were pretty familiar with the exam format and this is why they managed to speak within the time limits or even attempted to speak more. Even though the two pairs have different linguistics skills, the numbers presented in Table 1 below show that the candidates who were familiarized with the exam format have produced longer and more extensive stretches of language, within the time limits imposed by the exam format as compared to the candidates who have taken such an oral examination for the very first time.

**Table 1 Data collected by using the software *textinspector.com***

Item No.	B.M.	C.O.	Edward	Victoria
Sentence count	9	10	38	33
Token count (excluding numbers)	317	388	678	716
Type count (unique tokens, excluding numbers)	152	182	211	221
Average sentence length	35.22	38.80	17.84	22.34
Pauses	11	5	7	9
Words with more than 2 syllables	28	39	30	40
Lexical diversity VOCD <sup>8</sup>	85.55	83.37	74.68	58.51

English Vocabulary Profile <sup>9</sup>	tokens	tokens	tokens	tokens
A1	236	284	556	585
A2	42	47	64	75
B1	17	21	26	17
B2	5	16	4	2
C1	1	0	0	1
C2	0	2	0	0
Unlisted <sup>10</sup>	17	20	28	36
Lexical profile scorecard - percentage	49% B2	56% B2+	38% B1+	39% B1+

### **Test validity**

According to Alister Cumming (1996) the concept of validity and validation in language testing is a salient process involving “many types of evidence, analyses, and

<sup>7</sup> [https://www.cambridgeenglish.org/images/173977-cambridge-english-first-speaking-test-examiner-comments.pdf?utm\\_medium=social%20media&utm\\_source=youtube&utm\\_campaign=youtube%20organic&utm\\_content=b2\\_first\\_speaking\\_test](https://www.cambridgeenglish.org/images/173977-cambridge-english-first-speaking-test-examiner-comments.pdf?utm_medium=social%20media&utm_source=youtube&utm_campaign=youtube%20organic&utm_content=b2_first_speaking_test)

<sup>8</sup> (Duran, Malvern, Richards, Chipere 2004, p. 238)

<sup>9</sup> <http://englishprofile.org/index.php/wordlists/text-inspector>

<sup>10</sup> The unlisted tokens are words such as: names, locations and sounds that have been made during the recording but have nothing to do with the meaning or the message such as [ahm], [mhm]...the examiner's words have not been computed

interpretation” (Cumming, 1996, p. 2). William H. Angoff’s study on test validity interprets the concept after a thorough process of grasping the specific literature in this field, and his view on test validity is as follows:

*“Although different writers chose to define validity in slightly different language, the language, and the work dictated by that language in studying the validity of tests at that time, was characteristically pragmatic and empirical, even atheoretical, and validity data were generally developed just to justify a claim that a test was useful for some particular purpose” (Angoff, 1988, p.19)*

Moreover, by discussing other views on this topic as those presented in Lindquist’s book in 1942 reliability is also defined by concluding that:

*“the square root of the test reliability provides the upper limit for validity” (Angoff, 1988, p. 20).*

Thus, by analysing the aforementioned ideas we hereby conclude that this OLE can be more reliable and its validity improved if the test developers take into consideration the personal characteristics of the candidates as well as if in designing their test, they do not just simply take for granted that the candidates know their exam format but also prepare them by making sure that all future candidates are accustomed to all of the parts of the exam and that the exam format is not one of the challenges that might interfere with the final scores. Test developers should prepare clusters of trained EFL Teachers who can prepare their students within the school hours as an important element of the school curriculum that is meant to prepare students for their future and responding precisely to their needs. Of course, there are some more questions awaiting answers such as:

- ❖ Can candidates who are not familiarized with the exam format score the same as those who are familiar with the format?
- ❖ How can a candidate not knowing the format enrol for any of the exams? The answer is that students in their search for university programs find out the level of English required by the university they want to apply to, hire a tutor and intensely prepare for that specific level. The level achieved in class with the coursebooks offered for free by the Ministry of Education in Romania are not preparing students for their future careers nor do they train the students about the Cambridge exams formats.
- ❖ Who establishes what exam should be taken and what level should be attempted? The answer. The real answer, as demonstrated above: the tutor or the classroom teacher or the university program where the student intends to apply.

### **Conclusion**

By and large, even though many studies have been conducted on investigating the role of the examiners within the examinations why not change the focus and investigate the candidate element in Cambridge EFL Speaking Tests and why not bring some new information and perspective connected to the candidates’ profiles as essential components of the test.

It is also worth investigating the hypothesis where if examiners are trained to administer the test and examine the candidates why not pre-test, prepare and train the

candidates to take the test too. This might be possible by integrating the test formats into the national curriculum and coursebooks to include practice tests, videos and activities that are meant to prepare the students for such official language examinations.

Moreover, in a different study we shall demonstrate how reliable this test is by testing the same pair of students but this time after a thorough explanation and sufficient practice and training. In addition, the candidates will have also attended a Public Speaking Program by the time the next recording is going to be conducted.

According to the definition of test reliability, a test is reliable if it is the same candidate is tested again and with similar results. Therefore, the candidates are going to take the test again, in the same conditions, and their results are going to be investigated to see if they are similar or not.

How the candidate understands what is said and his or her personal interpretation together with the answer provided stand at the core of the OLE under investigation. This particular issue stands at the core of any type of oral language examination for it is this particular ability that enables the candidate to generate adequate and accurate answers in order to obtain the targeted score associated to the OLE. Given the fact that the candidate formulates his or her answers with this goal in mind, it is only easy to draw the conclusion that this type of speech acts are neither conversations, nor natural occurrences of language as advertised by the OLE developers and organizers, but they belong to the category of institutionalized discourse. Thus, from the very beginning we are dealing with a formal type of setting, that influences, at a certain level the oral production of the candidates, be it due to the fact that the candidates are trying to impress and perform at their best capabilities, be it due to the fact that the candidates' performance is influenced in a negative manner exactly because of the pressure exercised by their peers, who in most cases are complete strangers and by the examiners thus causing distress with a noticeable and a quantifiable impact even if Moder and Halleck (1998) concluded that the oral language examination setting does not discourage the communicative outcome of the speech act.

Another valid reason which has been under investigation here is the fact that the candidates might not be familiarized with the OLE format thus, the entire experience seems to be forced, not natural, and the candidates do not succeed to display their full potential in using the language just because they did not know how much time to speak or they did not practice collaborative speaking in school or they are just not that talkative. Therefore, less talkative, and introverted candidates might be at a disadvantage together with the candidates who did not have access throughout the high school years to practice tests, practice materials due to various reasons be it financial or just because the coursebooks provided by the school did not include such activities as those envisioned by the exams in question.

For this reason, this paper has studied the points that must be considered when a student prepares for such OLE. First and foremost, candidates must practice using spoken English effectively together with the appropriate and adequate way to participate in a discussion. Students must learn about the fact that if in a conversation the turn taking is aleatory, in an OLE the turn taking system belongs to the institutional

talk where each speaker must wait his or her turn to speak. Students must also learn how to master the technique of generating longer and yet natural stretches of speech in order to provide answers to specific questions (between 30 seconds and up to 2 minutes). Last but not least the candidates have to be able to speak loudly and accurately at all times and to request for clarifications if the case may be. All these elements are presented at the beginning of each practice test book with tips and tricks to successfully pass the test. Students have to pay the full price to have access to such information. Of course, there always is the online source of information but the information is presented in such an intricate way, that the students as well as parents always come back to their teachers to try and find out specific information regarding the exam. Teachers also have to spend a lot of money and time in order to stay updated with the test formats and to attend seminars and webinars to learn how to offer extra explanations to their students who plan on taking such exams. It is a common practice nowadays for test developers to organize free webinars for teachers to improve their teaching but the truth behind every free webinar is that it is only a marketing manoeuvre to sell another new product which, let's face it, is not available in schools. Everything connected to these very important language examinations happens outside the classroom, in the students' / teachers' free time during long sessions organized by teachers for students outside the school hours to familiarize with the future candidates with the format.

These techniques or skills are not among the techniques or skills to be taught in schools where the most talkative of students always have the "floor", nor can they be taught at home with parents. A teacher or a trainer is needed to moderate and manage the learning process - the speaking activities in such a way that all the students get the chance to learn how to solve this particular type of exam task - for example two-way collaborative task. Nonetheless, even though the national curriculum encourages learning new and foreign languages it does not stipulate among the key competences the ones dealing with discourse management or peer interaction. The national curriculum for each level of study encourages interaction and communication but the official exam, the Baccalaureate does not test peer interaction and does not provide a setting where two or more students are supposed to collaborate and generate stretches of language in order to prove their level of English, therefore both the teachers as well as the students are inclined to concentrate their efforts towards improving the skills necessary for the Baccalaureate exam or just practice their English during the EFL classes and hire a tutor if an official exam such as Cambridge, IELTS or TOEFL are needed for their academic path.

All the books available for future candidates encourage potential clients to prepare for the exam, presenting all the stages and parts of the examination. However, without purchasing these types of materials, which are constantly updating and changing format the high school students willing to take such exams do not stand a fair chance if he or she is not financially prepared for it. It is a world-wide prolific industry that makes it possible only to those who can afford such examinations to have access to their materials and exams. Of course, there are many online tutors worldwide who are

willing to share their knowledge for free but their sole interest is to score high in their daily number of views, followers and likes and not the excellence of their work. Moreover, many teachers do specialize in teaching courses to prepare students for these official examinations but outside the classroom. Subsequently, students have to pay more time and money, by attending additional classes and spending more and more of their time learning how to successfully pass such examinations which should be the object of the school curriculum. As long as this need is acknowledged worldwide, why not include it into the national curriculums around the globe because schools meant to prepare students for their future careers and the school curriculums are meant to meet the students' needs.

### References

- Angoff, W. (1988) *Validity: An evolving concept*. In H. Wainer and H. Braun (eds) *Test Validity* (pp. 19-32). Hillsdale, NJ: Erlbaum.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke: Palgrave Macmillan.
- Cumming, A. (1996). The concept of validation in language testing. In A. Cumming & R. Berwick, (Eds.), *Validation in language testing* (pp. 1–14). Clevedon, UK: Multilingual Matters LTD.
- Duran, P, D. Malvern, B. Richards, N. Chipere (2004) "Developmental Trends in Lexical Diversity" *Applied Linguistics OUP* 25/2: 220-242
- Fairclough, N. (1989) & (2001). *Language and power* (1st ed. & 2<sup>nd</sup> ed.). New York, NY: Longman.
- Fairclough, N. (2003). *Analysing Discourse: Textual Analysis for Social Research* London: Routledge.
- Halliday, M. & Hasan, R. (1976). *Cohesion in English*, London: Longman.
- Halliday, M.A.K. (2014). *Halliday's Introduction to Functional Grammar* (4th Ed.) Revised by Christian M.I.M. Matthiessen.
- Hughes, A., (2003). *Testing for Language Teachers*, Cambridge: Cambridge University Press.
- Johnstone B. (2008). *(Discourse Analysis, (2nd ed.)* Blackwell Publishing
- Lazaraton A. (2002). *A Qualitative Approach to the Validation of Oral Language Tests*. Cambridge University Press
- O'Sullivan, B. & Lu, Y. (2006). The impact on candidate language of examiner deviation from a set interlocutor frame in the IELTS Speaking Test (Research Report Vol. 6). Retrieved from IELTS Research Reports website: [http://www.ielts.org/PDF/Vol6\\_Report4.pdf](http://www.ielts.org/PDF/Vol6_Report4.pdf)
- Tannen, D., Heidi E., Hamilton, & Deborah Schiffrin (Eds.) (2015). *The Handbook of Discourse Analysis Second Edition* (vol. 1 & vol. 2) (2nd ed.) Blackwell Publishers Ltd (1e, 2001)
- Widdowson, H. G. (2007). *Discourse Analysis*, Oxford University Press.
- Young, R. and He, W. A. (1998). *Talking and Testing: Discourse approaches to the assessment of oral proficiency*. John Benjamins Publishing Company
- Young, Richard. (2006). *Conversational Styles in Language Proficiency Interviews*. *Language Learning*. 45. 3 - 42. 10.1111/j.1467-1770.1995.tb00961.x.
- <https://youtu.be/EdeZp0n0JHw>