

Assess me and I learn:

**Formatively assessing oral communication in the primary English
classroom**

Susana Passos Medina

**Relatório de Estágio de Mestrado em Ensino de Inglês no
1º Ciclo do Ensino Básico**

March 2022

Relatório de Estágio apresentado para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Mestrado em ensino de inglês no 1º ciclo do ensino básico realizado sob a orientação científica da Professora Doutora Carolyn Leslie e do Professor Doutor Rogério Miguel Puga.

Dedicated to my grandfather

ACKNOWLEDGMENTS

I would like to sincerely thank all those involved in the process that led to this action report, especially my students, my cooperative teacher, Carla Cunha and the school. A

warm thank you for always being available and willing to participate.

I would also like to express my appreciation to Professor Carolyn Leslie for guiding me through this research process.

Lastly, a big thank you to everyone who supported me and had the patience to listen to me countless times talk about teaching and assessment.

ASSESS ME AND I LEARN: FORMATIVELY ASSESSING ORAL COMMUNICATION IN THE PRIMARY ENGLISH CLASSROOM

Susana Medina

ABSTRACT

KEYWORDS: assessment, formative assessment, summative assessment, self-assessment, observation grids, oral communication, primary English classroom; foreign language teaching

The present research project focuses on developing assessment tools, namely observation grids, that assessed learners' oral communication skills, both formatively and summatively. This research discusses the observation grids' design and implementation. It also involved the students in their own assessment process through self-assessment and self-assessment criteria creation. It intends to answer the following research questions: (i) how can observation grids be designed and implemented to efficiently assess spoken production and spoken interaction? and (ii) how can I involve my students in the process of formatively assessing themselves regarding their oral communication skills?

The research project was developed in a private school during the first school term and involved a group of twenty third grade learners. Oral tasks in pairs and groups were implemented, and three different observation grids were used to formatively assess the students' oral skills. Lastly, students were asked to orally present a group project, and one observation grid was used to summatively assess it. In total, four observation grids were designed and implemented for this study and filled in by both the researcher and the cooperative teacher. Simultaneously, students participated in creating three self-assessment criteria and were asked to fill in a self-assessment grid after performing the speaking tasks. At the end of the study, a questionnaire was handed out to students to gather their thoughts about the self-assessment process and their perceptions of assessment.

This research project concluded that observation grids are a useful and efficient assessment tool. Regarding design, the more detailed and accurate the grid is, the more guided the teacher's observation becomes, helping the teacher gather important information about the students' oral skills. It has also concluded that observation grids benefit from the addition of a notes and comments section, in which there is space for the teacher to collect more qualitative data to support the checklist section of the grid. In terms of implementation and assessment type, the analysis of the results gathered from the student questionnaire and the self-assessment results show that students seem to feel more nervous and show higher performance anxiety levels when being summatively assessed. On the other hand, when a formative assessment was carried out through informal non-intrusive observation, the results show calmer and less anxious feelings from the students. Lastly, most students showed positive feelings towards self-assessing, and this process seemed to influence their confidence levels and foster positive feelings towards learning the language.

AVALIA-ME E EU APRENDO: AVALIAÇÃO FORMATIVA DA COMUNICAÇÃO ORAL NA SALA DE AULA DO 1º CICLO DO ENSINO BÁSICO

RESUMO

PALAVRAS-CHAVE: avaliação, avaliação formativa, avaliação sumativa, autoavaliação, grelhas de observação, comunicação oral, inglês na sala de aula do ensino básico, ensino de língua estrangeira

O presente projeto de investigação teve como foco o desenvolvimento de ferramentas de avaliação, nomeadamente grelhas de observação, que avaliem formativa e sumativamente as competências orais dos alunos. Este estudo pretendeu discutir o design bem como a implementação das grelhas de observação. Teve também como objetivo envolver os alunos no seu processo de avaliação através da implementação da autoavaliação e da criação de critérios para a mesma. Procurou responder às perguntas de investigação: (i) como é que as grelhas de observação podem ser desenhadas e implementadas para avaliar, eficientemente, a produção e interação oral? e (ii) como posso envolver os meus alunos no processo de se avaliarem formativamente em relação às suas competências orais?

Este projeto foi desenvolvido numa escola privada durante o primeiro período escolar e envolveu um grupo de vinte alunos do terceiro ano. Foram implementadas tarefas orais em pares e em grupo e foram usadas três grelhas de observação para avaliar formativamente as competências orais dos alunos. Foi pedido aos alunos para fazerem uma apresentação oral em grupo, avaliada sumativamente através do uso de uma grelha de observação. Para este estudo foram criadas, no total, quatro grelhas de observação. As grelhas foram preenchidas pela investigadora e pela professora cooperante. Em paralelo, os alunos participaram na criação de três critérios para a autoavaliação, autoavaliando-se após as tarefas orais. Os alunos preencheram um questionário para recolher a sua opinião sobre o processo de autoavaliação bem como as suas perceções sobre avaliação.

Este estudo concluiu que as grelhas de observação são ferramentas de avaliação úteis e eficientes. Relativamente ao design, quanto mais detalhadas e precisas são as grelhas, mais guiada se torna a observação do professor, ajudando-o a recolher informação relevante sobre as competências orais dos alunos. Foi também concluído que uma seção de notas e comentários é uma mais valia à grelha pois permite ao professor recolher dados de teor qualitativo que suportam a seção da grelha a ser preenchida apenas com breves símbolos. Relativamente à implementação e tipo de avaliação, os resultados dos questionários dos alunos e da autoavaliação apontam para que os alunos demonstrem maiores níveis de stress e de ansiedade de performance quando são sumativamente avaliados. Pelo contrário, os resultados mostram que os alunos parecem mais calmos quando avaliados formativamente, através de uma observação informal e não intrusiva. Finalmente, a maioria dos alunos demonstra sentimentos positivos em relação à autoavaliação e este mesmo processo apreça ter influenciado positivamente o seu nível de confiança e sentimentos positivos relativamente à aprendizagem da língua.

TABLE OF CONTENTS

Introduction.....	1
Chapter I: Literature Review.....	2
I.1. Oral communication: production and interaction.....	3
I.2. Assessment: definition and purpose.....	3
I.2.1. Formative and summative assessment.....	4
I.2.2. Observation grids as assessment tools.....	6
I.2.3. Self-assessment.....	6
I.3. Summary.....	7
Chapter II: The Action Research.....	7
II.1. Context.....	7
II.2. Methodology.....	8
II.2.1. Consent letters.....	8
II.2.2. Speaking tasks.....	9
II.2.3. Grid evaluation questionnaires (GEQ).....	10
II.2.4. Teaching journal.....	10
II.2.5. Student questionnaire.....	11
II.3. Results and discussion.....	12
II.3.1. Observation grids and GEQ.....	12
II.3.1.1. Grid criteria and parameters.....	12
II.3.1.2. Observation grid 1.A.....	13
II.3.1.3. Observation grid 2.A.....	15
II.3.1.4. Observation grid 2.B.....	16
II.3.1.5. Observation grid 3.A.....	17
II.3.2. Grid evaluation questionnaire (GEQ).....	19
II.3.3. Self-assessment grid and self-assessment criteria creation.....	21
II.3.4. Student questionnaires.....	24
II.3.5. Summary.....	26
II.4. Conclusion.....	27
II.4.1. Teaching implications.....	30

References.....	31
List of Tables.....	33
Appendix A: Letter of consent to parents.....	34
Appendix B: Letter of consent to the school board.....	35
Appendix C: Letter of consent to students.....	36
Appendix D: Grid 1.A.....	37
Appendix E: Grid 2.A.....	38
Appendix F: Grid 2.B.....	40
Appendix G: Oral interaction task 1 – Where is it?.....	42
Appendix H: Oral interaction task 2 – What’s missing?.....	43
Appendix I: Oral interaction task 3 – Pictionary.....	45
Appendix J: Oral interaction tasks 4 and 6 – Mini cards and Pictionary.....	46
Appendix K: Oral interaction task 5 – Board Game.....	47
Appendix L: Oral interaction task 7 – GuessWho.....	48
Appendix M: Project guidelines.....	49
Appendix N: Grid 3.A.....	50
Appendix O: Grid evaluation questionnaire.....	52
Appendix P: Excerpt from the teaching journal.....	55
Appendix Q: Student questionnaire.....	56
Appendix R: School’s official assessment grid.....	57
Appendix S: Photo of the puppet (Smelly Cat).....	58
Appendix T: Excerpt from classroom conversation about self-assessment.....	59

Introduction

Learning is necessarily and continuously accompanied by assessment, particularly in a non-naturalistic setting such as the foreign language (FL) classroom. As teachers, we are responsible for monitoring the students' progress and, consequently, recording evidence of it (Ioannou-Georgiou & Pavlou, 2003). Therefore, assessing the students' linguistic competences and accomplishments can best guide a reflective teacher in his or her teaching. Apart from this, learners have the right to be systematically, objectively, and fairly assessed on all four skills – listening, speaking, reading, and writing.

I believe that learning a language, especially when it comes to young learners, has quite a clear aim: communication. Our duty as young learner FL teachers is to focus a substantial part of our teaching on developing our students' oral skills, fostering positive attitudes towards speaking, and building the learners' confidence and willingness to put their communicative skills to use. However, how do FL teachers assess and collect evidence of their student's oral communication skills? From what studies show (Moreira et al, 2021) as well as according to my observation, FL teachers seem to continue favouring summative assessment over formative assessment. In this sense, a formal and less frequent assessment technique, often done in a test format, - i.e., a summative assessment - is preferred over an informal, more frequent assessment that provides an ongoing picture of the learners' abilities – i.e., a formative assessment. Furthermore, FL teachers continue to assign most assessment weight on written tests or a few oral presentations - only representative of the students' oral production, to the detriment of oral interaction (Moreira et al., 2021). In order to have complete and well-rounded assessment, teachers need to put both summative as well as formative techniques into practise. Consequently, investing more in formative assessment has become a central issue in my teaching practice.

This research aims to create assessment tools, observation grids, and discuss their design and implementation. Even though observation grids serve both formative and summative assessment, their implementation can be essentially different in terms of the impact they can have on students. This is why, although this research is focused on formatively assessing oral communication, I felt it added interest to also create an observation grid to summatively assess the learners' oral production to compare both assessment types and analyse their differences. It is important to point out that the

beforementioned observation grids aim to assess both spoken production as well as spoken interaction.

Additionally, this research also intends to involve students in formatively assessing themselves. Self-assessment encourages learners to become more responsible and involved in their learning and assessment, allowing them to understand what aims are expected and set learning goals so students achieve their best performance (Brewster & Ellis, 2002).

In this sense, this report will aim to answer the following research questions:

1. How can observation grids be designed and implemented to efficiently assess spoken production and spoken interaction?

2. How can I involve my students in the process of formatively assessing themselves regarding their oral communication skills?

Ultimately, I believe that there is still a gap to fill with regards to formatively assessing oral communication in the FL classroom, hence the need to research, reflect, and experiment with different types of observation grids to perfect and master this assessment technique. At the same time, and because teaching is impossible without students, I find it of the utmost importance to involve my students in this process, helping them become active agents in their learning process. Teaching is ineffective if no learning occurs, and formatively assessing student progress and competences is one way to guide and assign purpose to teaching.

The present report will be divided into two chapters: chapter I is dedicated to reviewing the literature on the topic and providing a theoretical framework, and chapter II will be subdivided into four sections. Sections 1 and 2 discuss the study's context, methodology, and tools used to collect data, and section 3 presents the research results and findings, discussing them to answer the previously mentioned research questions. Lastly, section 4 will draw conclusions and state the present research's teaching implications.

Chapter I: Literature Review

The present review will focus on understanding the main concepts relevant to this research: oral communication, formative and summative assessment definitions, tools and techniques. Furthermore, this review will explore self-assessment as a formative

assessment tool in which learners are the active agent. Lastly, it will be articulated with pertinent research related to the topic.

I.1. Oral communication: production and interaction

Young learners have an intrinsic instinct for interaction and communication, which motivates students to use and experiment with the language (Halliwell, 1992). Foreign language (FL) teachers have to take advantage of this instinct and guide their students to create healthy and positive experiences when orally expressing themselves in a foreign language. The importance of communication as an essential tool in learning a foreign language is recognized by the Portuguese Ministry of Education, and it is a crucial part of the English Curriculum for the primary level.

As the Companion Volume to the CEFR (2020) states, oral production consists of a “long turn” in which the speaker might describe something or give a short presentation. Oral production also entails a certain degree of speech preparation and organization before verbalizing the words. The CEFR adapted descriptor for young learners (2018) expects pre-A1/A1 learners to produce short sentences about themselves, giving basic personal information. The CEFR (2020) states that these types of activities have an important social function and value; this type of ability is not acquired naturally and needs to be practised in order to be learnt.

On the other hand, oral interaction is “the spoken language that takes place between two or more people and (...) it is the type of speaking and listening that occurs in real time (i.e., in the present) in communicative exchanges (i.e., interactions).” (Oliver and Philp, 2014, p.5). The CEFR adapted descriptor for young learners (2018) expects pre-A1 learners to be able to ask and answer questions, using short formulaic expressions and gestures to reinforce meaning. Interaction has been proven to lead to improved learning and, more specifically, peer interaction provides children the opportunity to engage with the language meaningfully, benefiting them socially, academically, and culturally (Leslie, 2021).

In conclusion, amongst the four skills, research and official education documents point to speaking possibly being the most fundamental to develop with young learners and should, for this reason, be more prevalent in the FL classroom.

I.2. Assessment: definition and purpose

The term “assessment” refers to the process and methods used “...to gather information about the children’s knowledge, ability, understanding, attitudes, and

motivation” (Ioannou-Georgiou & Pavlou, 2003, p. 4). Assessment also involves the analysis of data gathered about the learners’ performance and progress to make valid inferences about learning and teaching (Papp, 2019). It is through this data analysis that the teacher can check if the target objectives are being achieved by everyone, what aspects seem to be the most difficult and which seem to be the easiest and then use these findings to inform teaching.

Cameron (2001) metaphorically connects language learning and its assessment to plant growth. According to the author, the purpose of assessment is to ask how the plant is growing without disrupting its growth which means making use of assessment techniques that indicate learning is happening and, at the same time, generate positive attitudes towards assessment.

I.2.1. Formative and summative assessment

There are two types of assessment: formative and summative. According to Jang (2014), summative assessment provides a snapshot of a particular student’s performance at a given instructional time. Pinter (2006) adds that summative assessment appraises what has been learnt and achieved at the end of a more extended period – at the end of each term or the end of the school year – and it is often associated with a certificate of some kind. Regarding its outcomes, summative assessment techniques provide us with grades, marks, and rankings of students. As Cameron (2001) argues, its outcomes can be used to evaluate the effectiveness of a course or programme or other information such as students and teacher evaluations. Traditional ‘paper and pencil tests’ are often the assessment tool favoured by teachers. Traditional written tests, if well designed, can be efficient tools to assess the learners’ skills (writing, reading, and listening) but speaking is often left unattended. Furthermore, Pinter (2006) argues that because tests are isolated exercises, they might not show what the student knows or can do. Besides, young children are often not proficient at writing which means that traditional tests can be stressful and tiring.

Alternatively, formative assessment is used to “provide feedback at various stages of the learning process” (Jang, 2014, p.12) and tends to occur more often throughout the school year. As Rea-Dickins & Gardner (2000) point out, formative assessment is seen as a means to improve teaching and learning, helping teachers become more aware and responsive to their students’ learning needs. Formative assessment intends to make a difference in teaching through continuous lesson plan adjustments and adaptations, and

in learning by providing students with continuous feedback on their performance. Cameron (2001) states that feedback is helpful to learners, improving their learning process since it offers them the opportunity to understand learning objectives. Studies show that ongoing, or formative, assessment is likely to increase student motivation as it provides feedback and information on the students' performance over a longer period of time, hence providing a rounded view of the child's abilities (Brewster & Ellis, 2002).

A recent study conducted by Moreira et al. (2021) shows that traditional summative 'paper and pencil tests' continue to be the main assessment technique conducted in the FL classroom in Portugal. The study examined the teachers' perceptions, approaches, and practices regarding assessment to "identify possible constraints concerning the national standards and a transformative vision of language education in Portugal" (Moreira et al., p.11). To this end, the authors used a mixed-method approach to the study to collect quantitative and qualitative data: a grid analysing school cluster assessment documents and an online survey and semi-structured interviews to collect the teachers' perceptions and opinions. Regarding the school cluster assessment documents, the study concluded that in most cases, the expectations for English assessment do not match the national standards nor the recommendations in the literature for most assessment school clusters' documents give written tests a weighting of over 40% of the learner's total mark. The most common assessment tool is still the written test, and literacy skills are favoured over oracy. The survey also concluded that even though 89% of the respondents reported using observation grids and that 99% disagreed that assessment should be carried out exclusively using tests, there was some confusion understanding the terms "formative" and "summative" assessment, skill areas and what assessment instruments consist of. In other words, tests remain the primary tool for collecting evidence of learning. Lastly, teacher interviews allowed researchers to conclude that teachers consider speaking assessment difficult due to the size and time constraints of the groups. Also, observation grids are more often used to support summative assessment of speaking (oral production) and reading aloud, rather than being used systematically to support formative assessment of oral skills. Finally, self-assessment is often misused or absent, mainly focusing on behaviour rather than reflecting on learning. These results lead us to believe that there is still a gap regarding formative assessment and assessing a crucial skill in the Portuguese primary FL classroom: speaking.

I.2.2. Observation grids as assessment tools

It is important to stress that all assessment carried out in classroom contexts should provide information about what learners are able or unable to do, as well as informing teachers on how to support the students future learning (Jang, 2014). Checklists or observation grids can scaffold and guide both summative and formative assessment, mainly regarding the assessment of oral skills such as oral presentations or speaking pair and group work tasks. Furthermore, they help the teacher to focus and determine realistic, age-appropriate achievement targets (Nikolov & Timpe-Laughlin, 2020).

Cameron (2001) adds that, despite their convenience and easiness to manage, checklists may limit the information recorded about each student. As an alternative and an addition to checklists, the author proposes the adoption of a loose-leaf record book with a section for each student in which the teacher can note formal records of assessment as well as informal notes that will provide a useful picture of the overall development of the student as a language learner.

I.2.3. Self-assessment

Self-assessment is an integral part of a learner-centred approach and of formative assessment. It entails asking the students to “think about their own performances and achievements on a regular basis” (Pinter, 2006, p.136). According to Cameron (2001), the benefits of self-assessment are immense because learners can (i) understand more about their learning process, (ii) be motivated and involved in their learning, (iii) continue learning outside the classroom, (iv) create a more equal relationship between teachers and learners and for teachers to learn more about their students as individuals. Butler (2016) considers that engaging in self-reflection should enhance the children’s motivation and learning. Most importantly, it helps young learners develop positive attitudes towards language learning and assessment (Jang, 2014). Also, and when used successfully, it promotes autonomy (Nikolov, 2016).

Nevertheless, it is important to point out that young learners might find the process of self-assessment difficult at first, as well as an abstract concept to grasp. This is why it is of the utmost importance to take it slowly and, as Pinter (2006, p.136) advises “self-assessment works best if it is restricted to certain well-known tasks and situations. In the same sort of task, the same sort of criteria can be used and this gives children confidence and a sense of safety”. As Jang (2014) reminds us, self-assessment requires the teacher’s guidance, for learners do not naturally develop the ability to self-assess their performance

hence the need to teach them how and why. The author also adds that the assessment criteria should be shared with the learners prior to the students' self-assessment and even goes on to affirm that involving the learners and inviting them to brainstorm criteria for assessing their work can be a great idea.

I.3. Summary

Assessment should positively support learning and foster positive attitudes. It should motivate learners through the process and outcomes of assessment, help teachers plan lessons, inform the evaluation and improvement of courses and programs and, through assessment activities and feedback from them, support further learning (Cameron, 2001). The more teachers make use of all assessment types, the more well-rounded the assessment becomes, thereby providing fair and objective evidence on the child's abilities. Nevertheless, studies reveal that the Portuguese reality is in dissonance with theory (Moreira et al., 2019) by continuing to favour mainly summative assessment techniques, disregarding oral interaction assessment.

II: The Action research

II. 1. Context

This action research project was conducted in Raiz International Learning School, a private school located in an upper-class neighbourhood in Lisbon, Portugal. The school is composed of three preschool classes, five primary level classes, and two-second cycle classes (level five and six). It is important to mention that the word "international" in the school's name refers not to the adopted curriculum but rather to the pedagogical model in use, known as High Scope (Snyder, 2016). This model is fundamentally based on an active learning approach, working under the premise that children learn best by doing. Learner autonomy is, in this sense, consistently encouraged, as is critical thinking, decision making and expanding on the children's curiosity about the world (Snyder, 2016).

The participants in this study were a group of twenty third-year students, six boys and fourteen girls. There was one student diagnosed with dyslexia. All students were native speakers of Portuguese and were learning English as their only second language. One of them had lived abroad and had a native speaker parent, which affected the student's language proficiency level. The vast majority of these students had been in

contact with the English language since preschool, and two students were currently enrolled in extracurricular English activities.

Students had 60-minute English lessons three times a week. Due to schedule restrictions, I could only solo teach two complete lessons per week and would co-teach the third lesson because I could only be present for the last 30 minutes of the lesson. Together with the workbook adopted in 3rd grade, the coursebook used was *I Wonder 3* (Doodley & Obee, 2019), and lessons were planned following the order of contents in the book.

It is important to mention that all students knew me from previous years because English teachers are also present during other parts of their routine as a second adult. These interactions mainly happened in the afternoons, when the class was having *Estudo do Meio*; this subject happened in a project-based learning format that allowed students to include English in their projects if they wished to.

II. 2. Methodology

Action research was the chosen research methodology for this project, which entailed a qualitative and quantitative approach to data collection. According to Burns (2010), action research extends our teaching skills by allowing the teacher to take a reflective, critical and systematic approach to his or her teaching thus gaining a better understanding of him or herself, the students and the teaching context. The implementation of the action research in this study consisted of the following:

Table 1 - Action Research Plan. Adapted from Anne Burns (2010)

Stage	Description
Planning	Identify the problem (see Introduction); Prepare resources; Send consent letters to all stakeholders
Action & Observation	Implement observation grids during oral communication tasks and self-assessment grid after oral communication tasks; Implement data collection tools, student questionnaire and grid evaluation questionnaire; Collect and analyse data gathered in the observation grids, grid evaluation questionnaire, student questionnaire and teaching journal.; Adjust tools, if necessary
Reflection	Draw conclusions on the effectiveness and usefulness of the observation grids as assessment tools; Draw conclusions regarding the students' feelings and perceptions regarding self-assessment and assessment types

II. 2.1. Consent Letters

In early September, prior to this study, consent letters were prepared, sent to, and gathered from all stakeholders: parents, the school's director, and the students (Appendix A, B, and C). This step was considered essential to make sure the aims and purpose of the

study and data collection tools were clear for everyone involved, ensuring an ethical conduct for the present study.

Regarding the participants, a whole-class conversation was planned and conducted to ensure students understood what was going to happen. The letter was read aloud, and there were pauses to clarify some points and time for students to pose questions. All students agreed to participate.

II. 2.2. Speaking tasks

Students performed two types of speaking activities: (i) oral production and (ii) oral interaction. In order to assess them, four observation grids were used more than once during the speaking tasks to gather evidence from all the students' oral skills. The data gathered from these grids will be examined and discussed in the results section.

The first oral interaction task, integrated into the class's routine, was check-in. Each lesson would start with a brief interaction between teacher and five individual students: the teacher asked how the students were doing and the children would share with the whole class how they were feeling that day and the reason why. Check-in and check-out are common large group activities in the school. It allows the teacher to get a feeling of the group's emotions, provides children with the opportunity to share meaningful information about their personal lives and interests, and keep in touch with their emotions. Grid 1.A (Appendix D) was used to formatively assess this speaking task. I gathered records from individual students' utterances during check-in on seven different occasions, allowing me to have, on average, two records per student of the utterances they were able to produce.

Secondly, students were formatively assessed during other eight oral interaction tasks. The language in use in tasks 1, 2, and 3 related to prepositions of place, house objects, and the verb "to be". Grid 2.A (Appendix E) was used to formatively assess these tasks. The language in use in tasks 4, 5, 6, 7 and 8 related to adjectives (physical and psychological description) and the verbs "to be" and "have got". Grid 2.B (Appendix F) was used to formatively assess these tasks. The table below presents the tasks' summarized description:

Table 2 – Oral interaction tasks

Task	Date	Goal	Procedure	Materials
1. Where is it?	12/10	Form sentences to express where a certain object is in relation to another object	T asks a question about the picture to an individual S at a time. E.g., T.: Where is the CD player? S.: The CD player is on the bed.	Picture of a bedroom Appendix G

2. What's missing?	12/10	Spot the difference	Both Ss have the same picture of a bedroom but with different house objects. Ss talk to each other to spot the differences and draw what is missing. Ss work in pairs.	Worksheet A and B Appendix H
3. Pictionary	19/10	Describe a picture	S takes a picture card and describes it to the group, e.g., "the cat is in the box". The other Ss draw.	Picture card set Appendix I
4. Mini card: Find your opposite	9/11	Find the opposite adjective	There are different pairs of opposite adjectives (tall/short; kind/mean). Each S has one card with an adjective and has to ask other Ss if they have the opposite to the adjective on his card. Ss work as a whole class.	Adjective word set Appendix J
5. Board Game	12/11	Form sentences; name the opposite adjective	Ss take turns rolling the dice. Ss follow the game's instructions. Ss work in groups of 4.	Appendix K
6. Pictionary	16/11	Guess the words.	S takes a word card (fat, tall, etc) from the pile and draws a representative picture. Ss guess the adjective.	Same as Appendix J
7. Guess who	16/11	Describe a person	Ss have a worksheet with the same pictures of people. S chooses one person, describes it and the other Ss have to guess who it is. Ss work in groups of 4 or 5.	Worksheet Appendix L
8. Guess who	30/11	Describe a person	S chooses a colleague and describes him or her. Other Ss listen and guess who it is.	-

Lastly, as an oral production task, students prepared an oral presentation, the final stage of a group project developed in class. The language in use in this project was also related to adjectives and the verbs to be and have got. Students had to create a character, in a poster format, and present that character to the class. The guidelines for this project can be found in appendix M. Grid 3.A (Appendix N) was used to summatively assess the students during their presentations.

Apart from the oral presentation in which students were aware of being graded and assessed, during the other nine speaking activities assessment was informal, and students knew they were being observed, but not assessed.

The grids' design and implementation will be discussed and analysed in greater detail in the results sections.

II.2.3. Grid evaluation questionnaires (GEQ)

Grid evaluation questionnaires (Appendix O) were used to assess the aims, the parameters, the criteria, and the design of the observation grids. My cooperative teacher and I both filled in the same observation grids and, after completing that process we both filled in a GEQ. To each observation grid there is a corresponding GEQ. The aim of the GEQ was to gather our thoughts, opinions and perceptions in an organized and standardized manner. Having my cooperative teacher fill both the observation grids and its correspondent GEQ was a way to ensure a comparative and unbiased opinion. The findings from these GEQs will be analysed both qualitative as well as quantitatively.

II. 2.4. Teaching journal

As Moon (2006) explains, teaching journals are fundamentally a vehicle for reflection. Journal writing can entail a wide range of purposes (Moon, 2006), and the adoption of this data collection tool for the present study was based mainly on three of those aims. Firstly, by keeping a learning journal, my primary intention was to record experiences and note factual occurrences. Secondly, I intended to support my learning as a teacher through recording personal interpretations about different aspects that happened in the lessons (the students' reactions and opinions to certain activities, quotations of the students' utterances, my opinions in regards to the observation grids) and later reread this information to inform my future teaching. Finally, reflective teachers are better teachers and, according to the literature (Moon, 2006), journal-writing enhances a reflective practice, an attitude I wish to maintain throughout my teaching career.

I adopted a loose-leaf arrangement, and journal entries were always made after each lesson. I used the qualitative data collected in my teaching journal to support and add information to my findings. These results are presented as quotes in the results section. An excerpt of my teaching journal can be found in Appendix P.

II. 2.5. Student questionnaire

The aim of using student questionnaires was to gather the learner's perceptions and feelings regarding self-assessing their oral skills. Pinter and Zandian (2014, p.65) argue, "Children's concerns and agendas can be incorporated into research in different ways at different levels, representing different opportunities on a continuum from more modest involvement to levels where children actually conduct their own research". Since students were active agents in their formative assessment process, it was necessary and interesting to ask for their feedback. Furthermore, the analysis of the questionnaires provided an answer for my second research question, "How can I involve my students in the process of formatively assessing themselves regarding their oral communication skills?".

The same questionnaire (Appendix Q) was handed out to students on two different dates, on the 25th of November and on the 14th of December, to check if the students' perceptions had changed after a second whole-class conversation/discussion about self-assessment prior to the second questionnaire. The questionnaires consisted of eight questions, six of which required students to colour one of three faces to provide their answer - a happy face for "yes", a not so happy face for "more or less", and a sad face for

“no”. Apart from this, there were four “why” open questions, providing more qualitative data. The two last questions were multiple-choice with two answer options. The results from the questionnaires were analysed both quantitatively and qualitatively.

II. 3. Results and discussion

II. 3.1. Observation grids and GEQ

The following two subsections will discuss and analyse the design and implementation of the different observation grids to evaluate their efficiency and usefulness as oral communication assessment tools. This analysis also takes into account the data collected from the grid evaluation questionnaires (GEQ). The results and discussion in section 3.1.1 to section 3.2 intend to answer the first research question, “How can observation grids be designed and implemented to efficiently assess spoken production and spoken interaction?”. Sections 3.3, 3.4 and 4 will provide answers for the second research question, “How can I involve my students in the process of formatively assessing themselves regarding their oral communication skills?”.

II.3.1.1. Grid criteria and parameters

The four grids designed for this research share similar, yet not exactly the same, criteria, since they were designed for different types of speaking activities. In this sense, it is important to clarify the four most common parameters. Whenever there is a mention of telegraphic language, it refers to the use of very short unarticulated sentences. As Paradis (2007, p.388) states, “L2 children’s first utterances in English tend to be either formulaic or telegraphic, meaning that children rely heavily on memorized or unanalyzed phrases and use few grammatical morphemes.”. The pronunciation parameter relates solely to individual words correct or incorrect pronunciation. The accuracy parameter focused on accurate use of language, (e.g., if students could form correct sentences using the correct word order of the vocabulary and sentences in use). Lastly, the fluency parameter observes the learners’ pauses and hesitations while speaking and whether, even if inaccurately, the students could convey a message.

The coursebook I followed presented the language divided into vocabulary and structures. Additionally, the CEFR adapted descriptors for young learners (2018) is divided into vocabulary range and grammatical accuracy. This seemed to be the simplest way to schematize the language in use which is why grids 2.A, 2.B, and 3.A were designed following the same parameters. Although grid 1.A did not have a vocabulary

section, it featured a “one-word utterance” subsection that played the same role. The structures section was subdivided into standardized subheadings: “mixes L1 with L2”, “uses telegraphic language” and, “uses complete sentences”. The reason behind this subdivision was that having standardized criteria made the grid’s design applicable to all grammar contents and all types of structures being assessed. In this sense, the grid’s overall design would not change every time new content was taught. In grid 3.A, the structures parameter was subdivided into the specific structures being assessed for it facilitated filling it in while students were doing their oral presentation.

All four grids included a checklist section, to be filled in with a check mark or the symbols (+) for achieved, (+/-) for partially achieved, and (-) for not achieved. There was also at least one notes/comments section to record students’ utterances to ensure objective and qualitative evidence to support the information in the parameters.

The grids’ criteria were not shared with the students, since I believe they were too complex for such young learners. However, due to the collective creation of the self-assessment criteria, students were aware of their learning goals.

In the end, all evidence of the students’ oral skills collected through these grids informed the students’ final term mark. My cooperative teacher uses an official assessment grid (Appendix R) approved by the school board and the evidence we both collected through the observation grids affected the grading of the parameters “classroom participation”, “check-in/speaking” and “projects”.

II.3.1.2. Observation grid 1.A

Observation grid 1.A (see Appendix O) was designed to formatively assess the students’ oral performance during check-in. The *Aprendizagens Essenciais em articulação com o Perfil dos Alunos*’ document (DGE, 2018) foresees that 3rd grade students should be able to express themselves using short and simple sentences. The CEFR adapted descriptors (2018) adds that pre-A1/A1 learners should also be able to say how they are feeling using simple words such as “happy” or “tired”, accompanied by body language. Observation grid 1.A allowed me to observe the following:

Table 3 – Findings from Grid 1.A

One-word utterances	Uses more L1 than L2	Uses telegraphic language	Uses full sentences
1 (0%)	2 (10%)	4 (20%)	13 (65%)

As table 3 shows, the students’ language level is adequate, since only one student is unable to use more than one-word utterances, and more than half can use full sentences.

On seven different occasions I wrote down exactly what students said. These quotes from students added important and more qualitative information to the four parameters in the table above. After examining those utterances, I observed that 95% of students would consistently use the language chunk “I’m happy because” (e.g., “I’m happy because in six weeks of November is my birthday”); that 35% of students would consistently use the language chunk “how do you say...?” (e.g., “I’m happy because I this Saturday how do you say ‘tarde’ [afternoon]?”); and that 80% of students would justify their feelings by asking their peers, or the teacher, how to say a particular word or sentence, willingly repeating the word or expression (e.g., “I’m happy because...como se diz segunda feira [how do you say Monday]? (teacher says Monday) I’m happy because Monday we competition first place”). In fact, only 25% of students did not wish to share or ask how to say a particular word or sentence, which corresponds to the number of students still producing either one-word utterances or using simple telegraphic language. It was observed that students would only resort to Portuguese whenever they lacked the word or expression in English (e.g., “I’m happy because tomorrow is feriado [public holiday]”, “I’m happy because em Dezembro [in december] I’m going to circo [the circus]”).

These findings were helpful in mapping out the class’s oral skills as well as giving learners the opportunity to be creative with the language. As Leslie (2021) defends, students need to be challenged to communicate their own thoughts and meanings, which may oblige them to engage with both form and meaning.

In terms of teaching implications, I systematically observed that learners would only use the feeling “happy”. Not knowing if this was due to lack of vocabulary knowledge or just because they felt happy every day, I constructed a feelings poster to scaffold their speaking during this activity. The results were not clear, although there was one student that said “I’m nervous”, probably due to this feedback strategy.

The feedback I provided students with during this activity was recasting their utterances whenever there was an error that needed correction. I would also encourage students to seek and provide help amongst themselves to foster a collaborative environment. The uptake of this feedback mostly translated into students spontaneously rephrasing or repeating a sentence.

Grid 1.A also helped make note of a very shy and nonparticipative student’s progress, helping me realize that this particular student evolved from one-word utterances such as “happy” (19/10) to producing short telegraphic sentences, “I’m happy” (23/11) two months after we started implementing the check-in into our lesson’s routine.

The data gathered from the grid evaluation questionnaires did not lead to any changes in the observation grid because my cooperative teacher and I considered this grids' design to be efficient, useful and adequate to assess this speaking task.

II.3.1.3. Observation grid 2.A

Observation grid 2.A (see Appendix P) was designed to assess the students' oral interactions during the speaking activities 1, 2 and 3 (see table 2). The language in use was related to prepositions of place, house objects and the use of the structure “the (house object) is/are (preposition of place) the (house object)”, e.g., “the guitar is under the bed”. Observation grid 2.A allowed me to observe the following:

Table 4 – Findings from Grid 2.A

	Vocabulary		Structures			Pron. ¹
	House Objects	Prepositions of place	Mixes L1 with L2	Uses telegraphic sentences	Uses complete sentences	
Achieved	15 (75%)	15 (75%)	2 (10%)	5 (25%)	13 (65%)	17(85%)
Partially achieved	5 (25%)	4 (20%)				3 (15%)
Not achieved	0	1 (5%)				0

Overall, table 4 shows that the vocabulary was successfully learnt. Furthermore, overall pronunciation was clear and the vast majority of students were able to use complete sentences. The “recurrent problems” and “comments/notes” sections featured in observation grid 2.A were designed to collect and add qualitative data to the quantitative data shown in table 4. After close analysis, the qualitative data showed that 25% of students are producing unarticulated sentences, omitting the verb “to be”, e.g., “CD player in wardrobe” and “Cell phone on the bed”. Also, even though 65% of students were able to use complete sentences, there were two main accuracy issues. First, there was a lack, or misuse, of the articles (the/a/an), e.g. “A guitar is in front of a bed” and “Roller skates in front of table”. Secondly, there was a word order problem with the verb “to be”, e.g., “Is computer on the table”. This data highlighted the importance of including sections dedicated to recording students' utterances or teacher's comments in the observation grids, since they provide relevant and more detailed information which a simple checklist is unable to provide.

A limited number of lessons were planned to cover this topic, which means I was unable to plan a feedback intervention session to address these specific inaccuracies. However, in terms of teaching implications, this information influenced the planning of

¹ Pron.: Pronunciation

the content that followed which also involved an accurate use of the verb “to be”. Noticing these issues helped me focus my observation while teaching the following content which led to a corrective feedback intervention, discussed in the next subsection.

III.3.1.4. Observation grid 2.B

Observation grid 2.B (Appendix Q) was designed as an alternative to grid 2.A. It took into account the data and comments gathered in the grid evaluation questionnaire that corresponded to grid 2.A, altering its design in order to compare both and ascertain whether its effectiveness was improved (see table 7). This grid was designed to assess the students’ oral interactions during the speaking activities 4, 5, 6, 7 and 8 (see table 2). The language in use was related to adjectives and physical description, and the use of the verbs “to be” and “have got” in affirmative and interrogative sentences, e.g. “Is she tall?”, “Have you got dark hair?”. Observation grid 2.B allowed me to observe the following:

Table 5 – Findings from Grid 2.B

	Vocabulary		Structures			Pron. ²	Acc.	Flu.
	Adjectives	Physical description	Mixes L1 with L2	Uses telegraphic sentences	Uses complete sentences			
Achieved	16 (80%)	14 (70%)	1 (5%)	2 (10%)	17 (85%)	14(70%)	7 (35%)	15(75%)
Partially achieved	4 (20%)	6 (30%)				6 (30%)	12(60%)	3 (15%)
Not achieved	0	0				0	1 (5%)	2 (10%)

Overall, table 5 shows that the majority of students learnt the vocabulary, four students were still consolidating the adjectives and six were still consolidating the physical description vocabulary. Fluency was, in general, adequate, as was pronunciation. Nevertheless, if compared with grid 2.A, pronunciation had lower levels of achievement because words such as “thin”, “kind”, “straight”, “beard” and “moustache” were noticeably harder to pronounce.

In terms of structure use, more students were using complete sentences when compared to grid 2.A. There were, however, three core accuracy problems: (i) use of the verb “to be” instead of “have got” (e.g., “She is a brown eyes”, “She is the moustache”); (ii) issues formulating interrogative sentences, i.e. canonical word ordering with rising intonation (e.g., “You are tall?”, “You are fat?”); (iii) using both verbs, one immediately after the other (e.g., “She is has brown hair”, “She is has got dark hair”). These findings

² Pron.: Pronunciation
 Acc.: Accuracy
 Flu.: Fluency

show that, when interacting with each other, students were using the verb “to be” more accurately which was good progress when compared to the data gathered in grid 2.A. Furthermore, students were also progressively making more complex and longer sentences when compared to what was previously observed.

I opted for corrective feedback intervention in which I used a puppet (Appendix S), asking students to correct the puppet’s utterances. I noted on my teaching journal that students could easily spot two of the issues: (i) using the verb “to be” instead of “have got” and (ii) using both verbs, one immediately after the other. Nevertheless, students were unable to fully understand the interrogative inversion issue. After the feedback intervention, I observed that students improved on their accuracy with regards to the two first issues, leading me to believe the data observation grid 2.B allowed me to gather was extremely useful since it led to identifying language problems as well as addressing them with a positive outcome.

III.3.1.5. Observation grid 3.A

Grid 3.A (Appendix S) was designed to gather data to summatively assess the learners’ oral production skills in an oral presentation resulting from a group project. Unlike the other three observation grids, the implementation of 3.A happened formally, and all students knew there were being assessed.

It is important to point out that this was the first English project and oral presentation of the school year. Students had to create one or more characters, design a poster and present it to the class. The language in use was the same language assessed in grid 2.B. There were only nineteen students present the day of the oral presentation. Observation grid 3.A allowed me to observe the following:

Table 6 – Findings from grid 3.A

	Vocabulary		Structures		Pron. ³	Flu.	Acc.
	Physical Description	Adjectives	Verb to be	Verb have got			
Achieved	15 (79%)	17 (90%)	19 (100 %)	19 (100%)	14 (74%)	14 (74%)	14 (74%)
Partially achieved	4 (21%)	2 (11%)	0%	0%	3 (16%)	4 (21%)	4 (21%)
Not achieved	0%	0%	0%	0%	2 (11%)	1 (5%)	1 (5%)

Overall, the oral presentation results showed that both vocabulary and structures use were achieved. Seven students were marked “very good”, eight “good” out of which

³ Pron.: Pronunciation
 Flu.: Fluency
 Acc.: Accuracy

two were “good plus”, and the remaining four were “sufficient”. All marks took into account the speech coherence, voice level and adequate posture, and whether the student respected all formal parts required for the presentation – introducing the group, describing the character and ending the presentation. The “very good” mark meant that the student used five or more adjectives and physical traits, always used both verbs correctly, had little to no hesitations while speaking and could pronounce individual words correctly. The “good plus” mark meant that the student used four adjectives and physical traits, used both verbs correctly most of the times, had little to no hesitations while speaking and could pronounce individual words correctly. The “good” mark meant the student used three adjectives and physical traits, used both verbs with some inaccuracies, had some hesitations while speaking and could pronounce most individual words correctly. The “sufficient” mark meant the student only use two or less adjectives and physical traits, presented several structure inaccuracies, often hesitated while speaking or resorted to reading and had pronunciation issues.

When cross-referencing the findings from observation grid 2.B (see table 5) with 3.A (see table 6), it is possible to conclude that grid 3.A confirmed students could (i) use the vocabulary appropriately and (ii) form correct affirmative sentences using the verbs “to be” and “have got”. This data, together with formative assessment data, allowed me to confirm that learning took place and that core speaking goals were met. However, it is important to emphasise that this was an oral production task, meaning students were able to practise the language multiple times before being assessed, which, in turn, led to some students showing higher levels of accuracy than they would in an oral interaction task using the exact same language as shown on grid 2.B. Accuracy levels were immensely improved, since not a single student misused either of the verbs – a recurrent issue during oral interaction tasks (see table 5). This oral presentation did not require students to use the interrogative sentences, which meant I could not record evidence of this language aspect as I did during observation of oral interaction tasks. Regarding fluency and pronunciation, the findings gathered from grid 3.A match the findings gathered from grid 2.B. However, and compared to an oral interaction scenario, during an oral presentation it is easier for the teacher to notice all mispronunciations and speech hesitations, becoming more detailed and demanding when assessing pronunciation and fluency. So, in reality, students showcased better pronunciation and fluency during the oral presentation when compared to the results gathered from grid 2.B.

The data gathered from the grid evaluation questionnaires is not pertinent enough to mention, since it did not lead to any changes in the grid’s design. Due to time constraints, it was not possible to reuse this grid to assess a second oral presentation and see if the overall design continued to be useful.

In conclusion, oral production tasks such as oral presentations serve the purpose of allowing students to practice the language multiple times, forcing them to think about both form and meaning while constructing their presentations. Nevertheless, summatively assessing the students’ oral skills alone was not enough to inform an objective and fair final grade. Assessment needs to be carried out on multiple occasions to inform teaching on an ongoing basis. If summative and formative assessment are both put into practise, their collected data cross-referenced and both are fed into the student’s final term grade, the closest of painting a full picture of the student’s competences a teacher can get.

II.3.2 Grid evaluation questionnaires (GEQ)

Each observation grid had a correspondent grid evaluation questionnaire (GEQ), filled by me and by the cooperative teacher. Nevertheless, results gathered from GEQs corresponding to the observation grids 1.A and 3.A did not lead to any subsequent changes in the grids’ design hence irrelevant to mention. The table below presents the results from GEQs 2.A and 2.B, related to evaluations of the observation grids 2.A and 2.B, respectively. This data was considered relevant since it led to important changes in the grid’s design.

Table 7 – Findings from GEQ 2.A and 2.B

	GEQ 2.A			GAQ 2.B		
1. Style and structure						
The grid was visually simple to understand	100%	-	-	100%	-	-
The grid was easy to interpret	100%	-	-	100%	-	-
The grid was easy/simple to fill in	-	100%	-	100%	-	-
The grid has enough space for me to write down some students’ utterances that I feel are important	100%	-	-	100%	-	-
2. Assessment parameters/criteria						
The parameters in the grid were adequate for this activity (vocabulary, structures in use, creative use of the language)	100%	-	-	100%	-	-
I feel that there were parameters missing	-	50%	50%	-	100%	-
The wording of the parameters was clear	100%	-	-	100%	-	-
I need a lot of lessons to fill in the grid because there are too many parameters	-	100%	-	-	50%	50%
There were too many parameters and I felt overwhelmed	-	-	100%	-	-	100%
There were a lot of parameters but I feel that, with time, I can get acquainted with all of them and it will be easier	-	-	100%	-	-	100%

3. Aims							
Using the grid made me more aware of what to look for when observing students speak	100%	-	-	100%	-	-	
Using the grid provided a clearer focus for my observation and assessment	100%	-	-	100%	-	-	
Using the grid helped me understand in which parameters my students are struggling	100%	-	-	100%	-	-	
Using the grid helped me understand in which parameters my students are excelling	100%	-	-	100%	-	-	
Using the grid made me feel more confident when planning future lesson	100%	-	-	100%	-	-	
Using the grid made me feel more confident if I need to justify the grade of a particular student to the stakeholders (parents, school).	100%	-	-	100%	-	-	
Using the grid was useful for me.	100%	-	-	100%	-	-	

	Agree
	Partially agree
	Disagree

When examining the GEQ 2.A, that both my cooperative teacher and I filled in after using observation grid 2.A, our opinion was unanimous. While qualitatively analysing our comments to different statements on the GEQ, I observed that both of us noted that having two “recurrent problems” sections and a “comments/notes” section made the grid longer and more challenging to fill in. We also felt that the two lessons in which the speaking tasks occurred were insufficient. In addition, my cooperative teacher noted that there was no column to check for fluency and accuracy. Ultimately, we both considered the grid useful mostly because it helped us identify mistakes made most frequently. In conclusion, observation grid 2.A helped guide our observation, since we both felt it accomplished the aims stated on the GEQ; however, its design was not yet as effective as it could have been.

When analysing the GEQ 2.B, corresponding to observation grid 2.B, both my cooperative teacher and I found this design to be more efficient, clear, and simple to fill in during observation. When compared to grid 2.A, the main structure of grid 2.B remained the same: two main parameters, “vocabulary” and “structures” subdivided into more specific criteria. The most significant differences in design were that grid 2.B had separate columns to be filled in with a check sign regarding pronunciation, fluency, and accuracy, which was an important aspect GEQ 1 allowed me to see – missing parameters. Also, the “recurrent problems” sections were condensed into only one section “comments/notes”.

In conclusion, observation grid 2.B’s design turned out to be an improved version of grid 2.A, as its design was more effective and easier to fill in during observation. As

our findings show, this grid guided our observation in a more detailed way, which led to a successful feedback intervention that reflected an overall improvement in the students' oral skills.

II.3.3. Self-assessment grid and self-assessment criteria creation

Section 3 and 4 intend to answer the second research question, “How can I involve my students in the process of formatively assessing themselves regarding their oral communication skills?”.

A learner centred classrooms actively involves the learners in their learning process. A study conducted by Ozdemir (2015) revealed that students learn better when they are actively involved. The same study shows that students believe that learning would be more effective if conducted in a more inductive and “fun” way. Taking this information into account, to answer my second research question, I decided to use a puppet as a strategy to involve learners in the creation of the self-assessment criteria I would then ask them to use as a formative self-assessment tool.

On the 19th of October, I conducted a whole class discussion in which I introduced the puppet, “Smelly cat”, to the class, as previously mentioned in section 3.1.3 (see Appendix R). I told students that the puppet was going to talk to us and that they had to assume the teacher's role and pay close attention to what the puppet was saying, in case there was a need to correct it. During this activity I allowed myself and the students to use as much L1 as we felt necessary for two main reasons, (i) students lack the language skills to express themselves in this circumstance and (ii) I wanted to ensure everyone understood what was being said.

My goal was to guide the class, in an inductive way, to establish the following self-assessment criteria: (i) I spoke English, (ii) I know the vocabulary and (iii) I used complete sentences. I established these three criteria because I considered them to be the core, most essential aspects when speaking – using the language as often as possible and practising new vocabulary integrated into full sentences. Also, and because they are general standardized criteria, they are easy to adapt to different contents. This means that there is no need to constantly create new criteria or parameters whenever the content taught changed. Furthermore, by maintaining the same set of criteria, students could easily get used to them thus facilitating the self-assessment process. Lastly, establishing criteria made learners aware of what they were learning, ensuring learning goals were clear and focusing the learners' attention on achieving the learning outcomes.

To this end, I used a set of flashcards (house objects), the same ones I had been using with the class, so they were familiar with the vocabulary; I then asked the puppet to name the pictures. When the puppet only used Portuguese to say the names of the objects in the cards, students reacted and spontaneously started to name the objects in English. When I asked what was wrong one of the students commented “Ele não está a dizer em inglês” [He is not saying it in English] (19/10/21).

Secondly, and continuing to use the same flashcards, I asked the puppet to name the pictures but the puppet would either give a wrong answer or say “I don’t know”. Again, some students would spontaneously correct the puppet; when I asked what was wrong, students concluded that the puppet did not know the vocabulary and it was important that it did.

Thirdly, I recorded the following dialogue in my teaching journal (19/10/21):

Teacher: How old are you, Smelly cat?

Puppet: Nine.

Students (spontaneously): I am nine years old!!

Teacher: What’s your favourite animal, Smelly cat?










Puppet: Mouse.

Students (spontaneously): My favourite animal is [the] mouse!

I was quite impressed with how easily the class spotted what was missing and that when we speak, we are expected to use complete sentences. On my teaching journal entry after that lesson, apart from the excerpt above, I wrote “Puppet was very successful. At the end of the class, during reflection time, I asked students to tell me what we did with Smelly cat and they were able to do so” (19/10/21). I also noted that the class was extremely excited with the puppet and I felt that, even though this activity was a bit long, the puppet helped them to maintain focus throughout the whole process. A study conducted by Toledo and Hoit (2016) concludes that exposing children to puppets not only generates positive feelings but it also keeps the students engaged and motivated in the activity: learning is less stressful, it flows smoothly and children are more willing to participate which was exactly what I observed during this activity.

After this intervention, I designed the self-assessment grid presented below. Students were asked to fill in the self-assessment grid after having performed speaking tasks 4, 6, 7 and 8 (see Table 2) and after their oral presentation.













Table 8 - Self-assessment grid

1. I spoke English (<i>Falei em Inglês</i>)	  
2. I know the vocabulary (<i>Eu sei o vocabulário</i>)	  
3. I used complete sentences (<i>Usei frases completas</i>)	  

Even though I felt this intervention generated very positive effects on students, I noted the following on my teaching journal: “We talked about self-assessment in general but I still need to reinforce this conversation” (19/10/21). As Pinter (2006, p. 136) states, “it is important to emphasize that careful and gradual training is needed and the ability of children to assess themselves cannot be taken for granted”. To this end, during reflection time on the 26th of October, I asked the students’ opinion regarding the self-assessment grid and recorded the following on my teaching journal: “I asked for the students’ opinion and they told me it was easy. When I asked them what the criteria was, they were able to tell me. Self-assessment happened with no setbacks, there were no questions about it.” (26/10/21). Finally, on the 9th of November, I planned another whole-class conversation directing the students’ attention to the self-assessment grid to ensure everyone understood it and to get some feedback from the students. On my teaching journal, I recorded an excerpt of the conversation (Appendix T) in which students explained the three self-assessment criteria and provided language examples. From this conversation excerpt, it is safe to assume three main and very important aspects: (i) students seemed to understand what the standard criteria meant, since they were able to explain it, (ii) they were able to apply those standard criteria to the language content being taught and (iii) it seemed that the speaking goals were clear, since they could verbalize the type of language they were expected to use, both isolated words as well as two types of sentences, affirmative and interrogative.

The table below presents the results from the self-assessment grid students were asked to fill in after the speaking tasks:

Table 9 - Self-assessment grid results

	Task 1 (17 participants)			Task 2 (19 participants)			Task 3 (16 participants)			Task 4 (19 participants)		
												
I spoke English	88%	12%	0	95%	5%	0	100%	0	0	84%	16%	0
I know the vocabulary	76%	24%	0	84%	16%	0	94%	6%	0	79%	21%	0
I used complete sentences	71%	29%	0	63%	37%	0	81%	19%	0	74%	26%	0

It is interesting to note that students have a more or less accurate perception of their oral performance that matches the results gathered from the observation grids (1.A,

2.A, 2.B and 3.A). As previously discussed, the observation grids allowed me to conclude that the aims for vocabulary and structures were overall achieved and students seem to have a similar perception which showed that, if guided and scaffolded, self-assessment might have quite positive outcomes regarding raising language and performance awareness. It is also of note that the self-assessment results in task 4 (oral presentation) are the ones furthest from the truth because, for example, even though all students spoke English the entire presentation, three of them felt they could have spoken more. Also, if we look at table 7, we see that accuracy levels during the oral presentation were very high but the students' perception is somewhat different. This might be due to the different implementations of the observation grids since the oral presentation was a summative formal assessment moment and students were aware of it. On the other hand, in tasks 1, 2, and 3, students knew they were being observed but were unaware they were being assessed. In this sense, summative assessment appeared to have negatively influenced the students' perceptions and raised their performance anxiety levels. In tasks 1, 2 and 3, which were formatively assessed, students seem to show much more confidence in their oral performance.

In conclusion, it would be plausible to assume that self-assessing helped learners become more aware of the language being learnt, since they had to reflect on their performance during a given speaking activity. Furthermore, being involved in the criteria creation process, as well as using the grid to self-assess often, probably made students feel responsible for their learning and more aware of their speaking achievements and shortcomings. This last topic regarding the students' feelings and perceptions is going to be explored in the next section, through the analysis of the student questionnaires.

III.3.4. Student questionnaires

Student questionnaires were handed out on two different dates, namely the 25th of November and the 14th of December. On both occasions, 18 participants filled in the questionnaire. The table below compares the quantitative results gathered on both questionnaires:

Table 10 – Student questionnaires’ results

	Questionnaire 1 - 25/11			Questionnaire 2 – 14/12		
	Yes	More or less	No	Yes	More or less	No
Q.1. I like to speak English	15 (83%)	1 (6%)	2 (11%)	16 (89%)	0 (0%)	2 (11%)
Q.2. I like to self-assess my oral skills	14 (78%)	3 (17%)	1 (6%)	15 (83%)	3 (17%)	0
Q.3. Self-assessing my oral skills makes me feel more confident when I speak English	13 (72%)	2 (11%)	3 (17%)	14 (78%)	1 (6%)	3 (7%)
Q.4. Self-assessing my oral skills helps me learn English	12 (67%)	1 (6%)	5 (28%)	13 (72%)	2 (11%)	3 (7%)
Q.5. Self-assessing oral skills is important	16 (89%)	1 (6%)	1 (6%)	17 (94%)	1 (6%)	0
	Calm		Nervous	Calm		Nervous
Q.6. When Susana assesses my oral skills, I feel...	11 (61%)		7 (38%)	15 (78%)		3 (15%)
	Responsible		Confused	Responsible		Confused
Q.7. When I self-assess my oral skills, I feel...	11 (61%)		7 (39%)	17 (89%)		1 (5%)

The first five questions aimed to ascertain the students’ perceptions about the process of self-assessment their oral skills. It is important to mention that this class had never had a self-assessment experience regarding their oral skills. It is possible to observe that the students’ perceptions and opinions did not vary greatly between questionnaires. Nevertheless, the second questionnaire shows slight differences possibly due to the whole-class conversation about the importance of self-assessment I reinforced before handing out the second questionnaire. Overall, the students’ feelings towards learning the language are positive. Students also demonstrate positive feelings towards self-assessing their oral skills.

Most students feel that self-assessing is important and, when analysing the open-question “why” in both questionnaires, the participants’ answers tend to be “because I can correct myself”, “because I can check if I spoke English” and “because it helps me to learn English”. However, these answers seem to contradict the answers to question 4, mostly the “why” question in questionnaire 1. In their answers we can read “no, because I don’t do anything in English”, “because I don’t learn English”, “because I don’t write nor speak” or “because it is only colouring smiley faces”. When analysing these answers, it is plausible to assume that students took the question too literally and were focusing not on the self-assessing process but rather on the logistical process of filling in a grid. Also, this question might be cognitively difficult since it implies the understanding of the importance of learning to learn strategies.

Lastly, and despite effect sizes being small, a study conducted by Butler and Lee (2010) found self-assessment improves learners’ confidence in learning English. In fact,

both questionnaires indicate that a large part of the group feels it contributes to their confidence. When analysing the students' answers to the "why" question in both questionnaires, most participants could not explain themselves. The two main strands of answers were either "I don't know", or "because it makes me feel confident". The participants that could develop their answers a bit further, linked their confidence to enjoying learning English, which indicates a correlation between positive feelings towards the subject and confidence level. However, one might argue that self-assessment could have influenced the students' positive feelings towards the subject, since, as the students themselves state, it allowed them to improve, correct themselves and learn.

In regards to question 6, it is interesting to note the difference between the first and second questionnaires: out of the seven participants who said they felt nervous when assessed, four changed their answer in questionnaire 2. I believe that the reason for this change might be linked to the fact that learners were summatively assessed on the 23rd of November, two days prior to the first questionnaire thus influencing their feelings in regards to being assessed. This data might corroborate the idea that summative assessment does indeed negatively affect the learners' feelings towards assessment because it induces stress (Pinter, 2006). Therefore, 'non-invasive' formative assessment techniques such as the use of observation grids seem to foster a calmer and anxiety free environment.

Lastly, the results gathered in question 7 were also quite different. This fact might corroborate the idea that the more students self-assess, the more proficient they become hence the less confused they feel. The teacher needs to guide and scaffold the students' self-assessment process (Jang 2014) and the use of the same criteria seems to also be helpful because, over time, learners get better acquainted with it (Pinter, 2006) hence feeling more in charge of their own learning process which, in these questionnaires, translates into the students' feeling of responsibility.

II. 3.5. Summary

The findings gathered from the observation grids as well as the results of the grid evaluation questionnaires answer the first research question, "How can observation grids be designed and implemented to efficiently assess spoken production and spoken interaction?" by providing data on the usefulness and efficiency of the beforementioned grids. This research project has concluded that observation grids are a useful and efficient assessment tool. Regarding design, the more detailed and accurate the grid is, the more

guided the teacher's observation becomes, helping the teacher gather important information about the students' oral skills. It also concluded that observation grids gain from the addition of a notes and comments section, in which there is space for the teacher to collect more qualitative data to support the checklist section of the grid.

In addition, the discussion and findings from the self-assessment criteria creation and the self-assessment grid results answer the second research question, "How can I involve my students in the process of formatively assessing themselves regarding their oral communication skills?" by showing that involving the learners in their learning and assessment process was an efficient strategy. Most students showed positive feelings towards self-assessing, and this process seemed to influence their confidence level, fostering positive feelings towards learning the language.

Furthermore, the results from the students' questionnaires provide additional information on the learners' opinions and perceptions regarding assessment techniques. Results show that students seem to feel more nervous and showcase higher performance anxiety levels when being summatively assessed. On the other hand, when a formative assessment was carried out through informal non-intrusive observation, the same results point to more calm and less threatening feelings from the students.

II.4. Conclusion

Despite the importance that literature gives formative assessment, research shows that most FL teachers in Portugal tend to mainly summatively assess their students' oral production skills (Moreira et al., 2021). This approach translates into two important limitations that have to be taken into consideration: (i) formative assessment tools and techniques are, overall, disregarded when it comes to oral skills, and (ii) oral interaction is left unassessed.

This research concludes that observation grids are fundamental assessment tools that can serve, guide, and scaffold the teacher's observation both in summative and formative assessment moments. An observation grid is efficient and valuable when (i) it is clear and straightforward to fill in during observation, (ii) it entails all the necessary parameters that help the teacher focus the observation of the language in use, (iii) it allows space to record examples of the students' utterances, (iv) it provides essential information about the students' use of the language – recurrent problems – that give the teacher clear focus to plan a feedback intervention and address the problems, (v) it focuses and shows individual progress and accomplishments, (vi) it provides a map of the class's overall

language level which, consequently, will be fed into future lesson planning (increasing or decreasing the challenge level), and (vii) it provides objective evidence of each student's language abilities in order to ensure a fair and objective assessment that will be translated into a grade at the end of each term.

It is also important to stress that designing efficient and useful observation grids is a process that needs constant adjusting and readapting. The grid evaluation questionnaires were extremely helpful in this process of pinpointing positive and negative aspects in the observation grids that consequently led to their polishing and perfecting.

Even though the assessment tool to collect data about the students' oral skills is the same – namely, observation grids –, its implementation can be quite different and, therefore, have different outcomes. Oral presentations serve the vital purpose of making students engage with the language and getting progressively familiar with it by practising it multiple times. Furthermore, young students have a long academic path ahead in which they will face formal assessment moments and public speaking situations, so, in this sense, it is good practice. The findings from grid that assessed the oral presentation, 3.A, were useful mainly to confirm that learning had happened and for students to consolidate that knowledge. Nevertheless, by its nature, summative assessment moments are few. Moreover, summative assessment is an assessment of learning which means its aim is to measure to what extent students have mastered what they were taught (Nikolov & Timpe-Laughlin, 2020). Therefore, it leaves little space for learning to happen based on feedback and for teachers to adjust their teaching regarding a particular content. Moreover, and realistically speaking, it is complicated to summatively assess the students' oral skills in all taught contents throughout the term because teachers need to allocate several classes for the students to carry out their projects and oral presentations take a lot of lesson time.

Alternatively, and guided by observation grids, formative assessment can be carried out frequently throughout the term, allowing the teacher to realize how students are learning a particular content, adjusting teaching as she or he goes. Also, the literature says that formative assessment entails feedback which, in its turn, aims at supporting future learning (Cameron, 2010) which this research was able to confirm. The teacher can plan feedback interventions that, hopefully, lead to improved learning. In addition, apart from only assessing oral production, oral interaction can now be addressed. The overall use of the observation grids – 1.A, 2.A, and 2.B – and their analysis led to the pinpointing of language issues which subsequently led to changes in teaching and lesson planning.

Moreover, observation grid 2.B led to a corrective feedback intervention with positive outcomes.

By formatively assessing the students' oral interaction, observation grid 2.B shed light on language issues that the summative assessment grid 3.A failed to. This aspect reinforces and confirms that entirely basing assessment on only one assessment type is insufficient to paint a complete picture of the students' abilities, issues, and, more importantly, language learning needs. Therefore, there needs to be a balance between assessment types, and formative assessment results can be fed and influence the students' final grades. Due to the schools' policy, my cooperative teacher was responsible for grading students at the end of the first term. However, fortunately, she decided to use the results from the observation grids we both filled in when formatively assessing students and incorporated them into their final grades.

Lastly, the students' questionnaires and the self-assessment results show that students apparently feel more nervous and showcase higher performance anxiety levels when being summatively assessed. On the other hand, when a formative assessment was carried out through informal non-intrusive observation, the results point to more calm and less threatening feelings from the students.

Because a classroom is composed of both teachers and students, I felt it was of the utmost importance to involve my students in their assessment process, challenging them to be active agents in their learning and assessing process. To this end, using a puppet as a classroom strategy to guide learners in creating self-assessment criteria was successful and seemed to foster positive feelings towards self-assessment. Additionally, the results gathered from the students' questionnaires show that learners manifested positive feelings towards learning English. Furthermore, students considered that self-assessing their oral skills made them feel more confident, which leads us to assume there might be a correlation between enjoying learning the language, self-assessing, and feeling confident. It is also interesting to note that the student questionnaires' results indicate and reinforce a correlation between summative assessment fostering nervous feelings and a more stressful environment, and formative assessment fostering feeling calmer and anxiety-free environment. Lastly, the student questionnaire findings also corroborate what literature has been saying regarding guiding and teaching students to self-assess, which will progressively lead to learners becoming more proficient performer appraisers.

The limitations found for this research were mostly related to time constraints. It would have been very interesting to continue to use observation grid 1.A and 2.B

throughout the school year to fully confirm their usefulness and effectiveness in assessing oral skills. Besides, having the opportunity to use grid 2.B when assessing other types of contents would help confirm the efficiency of its design in assessing oral interaction and ensure it did not need more alterations to become a progressively more sophisticated tool. The same time constraints apply to the results gathered from the students' questionnaires. The more time students had to learn how to self-assess, explore this learning to learn and assessment strategy, and, hopefully, become more proficient in it, might affect the results of the questionnaires. Furthermore, it would be interesting to have other teachers use the grids and collect their perceptions and opinions through the grid evaluation questionnaire, significantly contributing to the grids' refinement.

II.4.1. Teaching implications

This research has enriched me as a teacher, and its findings have turned out to be valuable to inform my future teaching practice. As a teacher, my final goal is to teach efficiently and ensure that, ultimately, it all leads to learning. If assessment is structured, frequent, and results are fed into the tailoring of lessons, it becomes one of the most efficient tools a teacher can rely upon.

I have continuously felt the lack of oral formative assessment tools during my teaching career. The present research led to the creation of valuable palpable materials – observations grids – that can continue to be put into practise. These grids have proven to be efficient, functional, and provided structure and focus to my observation. In this sense, this research had a very relevant impact on my present and future teaching: it facilitated the assessment process I will continue to carry out throughout my teaching career and enhanced my confidence regarding all stakeholders and, most importantly, when planning lessons tailored to my students' needs. It has also promoted a reflective approach to my teaching, which is an attitude that I truly wish to maintain. I feel it has helped me progress professionally and as an individual.

Reference List

- Brewster, J., & Ellis, G. (2002). *The primary English teacher's guide*. Penguin Longman.
- Burns, A. (2010). *Doing action research in English language teaching: A guide for practitioners*. Routledge.
- Butler, Y. G. (2016). Self-assessment of and for young learners' foreign language learning. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 291–315). Springer.
- Butler, Y.G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, 27(1), 5-31.
- Cameron, L. (2001). *Teaching languages to young learners*. Oxford University Press.
- Council of Europe. (2018). *Collated representative samples of descriptors of language competences developed for young learners: Resource for educators. Volume 1: Ages 7-10*. <https://rm.coe.int/collated-representative-samples-descriptors-young-learners-volume-1-ag/16808b1688>
- Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing.
- DGE, Direção-Geral da Educação. (2018). *Aprendizagens essenciais em articulação com o perfil dos alunos*. http://www.dge.mec.pt/sites/default/files/Curriculo/Aprendizagens_Essenciais/1_ciclo/ingles_1c_3a_ff.pdf
- Dooley, J., & Obee B. (2019). *I wonder 3: Activity book*. Express Publishing.
- Dooley, J., & Obee B. (2019). *I wonder 3: Pupil's book*. Express Publishing.
- Halliwell, S. (1992). *Teaching English in the primary classroom*. Longman.
- Ioannou-Georgiou, S. & Pavlou, P. (2003). *Assessing young learners*. Oxford University Press.
- Jang, E.E. (2014). *Focus on assessment*. Oxford University Press.
- Leslie, C. (2021). Interaction patterns and support for learning in the primary foreign language classroom. *Porta Linguarum*, 36, 65-82. <https://doi.org/10.30827/portalin.v0i36.15712>

- Moreira, M.A., Mourão, S., Leslie, C., & Monteiro, E. (2021). *Avaliação das aprendizagens de Inglês no 1ºCEB: Um estudo em Portugal continental*. Lisboa: CETAPS, NOVA FCSH.
- Moon, J.A. (2006). *Learning journals: A handbook for reflective practice and professional development*. Routledge.
- Nikolov, M. (2016). Trends, issues, and challenges in assessing young language learners. In M. Nikolov (Ed.), *Assessing young learners of English: Global and local perspectives* (pp. 1–18). Springer.
- Nikolov, M. & Timpe-Laughlin, V. (2020). Assessing young learners' foreign language abilities. *Language Teaching*, 54(1), 1-37. doi:10.1017/S0261444820000294
- Oliver, R. & Philp, J. (2014). *Focus on oral interaction*. Oxford University Press.
- Ozdemir, K. (2015). Enhancing a learning-centered classroom rather than a teacher-centered one. In K. Dikilitas, R. Smith & W. Trotman (Eds.), *Teacher-researchers in action* (pp. 207-224). IATEFL.
- Paradis, J. (2007). Second language acquisition in childhood. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 387-405). Malden, MA: Blackwell.
- Papp, S. (2019). Assessment of young English language learners. In S. Garton & F. Copland (Eds.), *The routledge handbook of teaching English to young learners* (pp. 389-408). Routledge.
- Pinter, A. (2006). *Teaching young language learners*. Oxford University Press.
- Pinter, A., & Zandian, S. (2014). 'I don't ever want to leave this room': Benefits of researching 'with' children. *ELT Journal*, 68(1), 64-74.
- Rea-Dickins, P. & Gardner, S. (2000). Snares and silver bullets: disentangling the construct of formative assessment. *Language Testing*, 17(2), 215-243.
- Snyder, C.M. (2016). *Lesson plans for a strong start: The first 30 days for infants*. HighScope.
- Toledo, F., & Hoit, S. (2016). Developing speaking skills in the first grade: The impact of puppets on young learners' spoken interactions and motivation. Cambridge University Press Teacher Research Programme.

LIST OF TABLES

Table 1 - Action Research Plan. Adapted from Anne Burns (2010).....	8
Table 2 – Oral interaction tasks.....	9
Table 3 – Findings from Grid 1.A.....	13
Table 4 – Findings from Grid 2.A.....	15
Table 5 – Findings from Grid 2.B.....	16
Table 6 – Findings from grid 3.A.....	17
Table 7 – Findings from GEQ 2.A and 2.B.....	19
Table 8 – Self-assessment grid.....	23
Table 9 - Self-assessment grid results.....	23
Table 10 – Student questionnaires’ results.....	25

Appendix A: Letter of consent to parents

Pedido de Autorização

Caros Pais e Encarregados de Educação,

O meu nome é Susana Medina e irei, ao longo do 1º período do corrente ano letivo, realizar a Prática de Ensino Supervisionada (PES II) no âmbito do Mestrado em Ensino de Inglês no 1º Ciclo do Ensino Básico pela Faculdade de Ciências Sociais e Humanas (FCSH), Universidade Nova de Lisboa. Afim de completar a PES II, ou estágio final, terei de desenvolver um projeto de investigação. Para tal, e em colaboração com a professora Carla Cunha, acompanharei o seu educando nas aulas de inglês entre os meses de setembro a dezembro.

O meu projeto de investigação intitula-se *Assess me and I learn: Formatively assessing oral communication in the FL classroom* (Avalia-me e eu aprendo: Avaliação formativa da comunicação oral na sala de aula de língua estrangeira). Este projeto visa a reflexão e criação de ferramentas de avaliação que possam ser utilizadas de forma útil e eficiente de maneira a melhor conseguir pintar um quadro objetivo das competências orais dos alunos. Este projeto tem como objetivo envolver os alunos na criação dos parâmetros para esta avaliação oral de forma a promover a consciencialização dos mesmos. Se somos parte da criação dos critérios então estaremos todos conscientes daquilo que é esperado de nós. Para além disso, haverá momentos de autoavaliação e de avaliação de pares de maneira a reforçar essa autoconsciencialização sobre as nossas competências da oralidade. O seu educando terá de preencher um questionário no fim da minha intervenção e participar normalmente nas aulas.

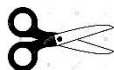
Assim, venho solicitar a vossa autorização para que o seu educando possa participar neste projeto. O seu educando foi também informado e já expressou a sua vontade. É importante frisar que a participação é voluntária e o seu educando pode, a qualquer momento, decidir não participar mais. A instituição e os participantes serão sempre anónimos e não serão recolhidas imagens/fotografias de nenhum tipo em nenhum momento.

Peço-vos que, em caso de dúvida ou se sentirem necessidade de mais esclarecimentos, não hesitem em contactar-me através da Coordenadora de 1º Ciclo.

Agradeço, desde já, a vossa atenção e fico a aguardar a vossa resposta mediante o preenchimento do destacado (a entregar, se possível, até dia 8 de outubro).

Lisboa, 23 de setembro de 2021
Susana Passos Medina

Prof.^a Dra. Carolyn Leslie
Orientadora de Estágio
FCSH, Universidade Nova de Lisboa



Eu, _____

Encarregado de Educação de _____, declaro

que fui informado(a) dos objectivos do projeto intitulado *Assess me and I learn: Formatively assessing oral communication in the FL classroom* (Avalia-me e eu aprendo: Avaliação formativa da comunicação oral na sala de aula de língua estrangeira), e autorizo o meu educando a participar no estudo. Data: _____

Assinatura: _____

Appendix B: Letter of consent to the school board

Pedido de Autorização

A/C Direção,

O meu nome é Susana Medina e irei, ao longo do 1º período do corrente ano letivo, realizar a Prática de Ensino Supervisionada (PES II) no âmbito do Mestrado em Ensino de Inglês no 1º Ciclo do Ensino Básico pela Faculdade de Ciências Sociais e Humanas (FCSH), Universidade Nova de Lisboa. Afim de completar a PES II, ou estágio final, terei de desenvolver um projeto de investigação que levarei a cabo na vossa instituição com a colaboração da professora Carla Cunha.

O meu projeto de investigação intitula-se *Assess me and I learn: Formatively assessing oral communication in the FL classroom* (Avalia-me e eu aprendo: Avaliação formativa da comunicação oral na sala de aula de língua estrangeira). Este projeto visa a reflexão e criação de ferramentas de avaliação que possam ser utilizadas de forma útil e eficiente de maneira a melhor conseguir pintar um quadro objetivo das competências orais dos alunos. Também, este projeto tem como objetivo envolver os alunos na criação dos parâmetros para esta avaliação oral de forma a promover a consciencialização dos mesmos. Se somos parte da criação dos critérios então estaremos todos conscientes daquilo que é esperado de nós. Para além disso, haverá momentos de autoavaliação e de avaliação de pares de maneira a reforçar essa auto-consciencialização sobre as nossas competências da oralidade. Os alunos terão de preencher um questionário no fim da minha intervenção e participar normalmente nas aulas.

Venho por este meio solicitar a vossa autorização para desenvolver o supracitado projeto com a turma do 3º ano, de setembro a dezembro de 2021. Igualmente, solicitarei por escrito a autorização aos alunos e aos respetivos encarregados de educação. A recolha de informação será feita apenas através dos questionários preenchidos pelos alunos e da análise das ferramentas de avaliação criadas por mim. Toda a informação recolhida ao longo do projeto constituirá e será mencionada no meu relatório final de estágio e, eventualmente, em publicações no âmbito académico.

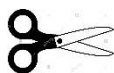
É importante frisar que a participação neste projeto é voluntária e que, como tal, qualquer aluno é livre de não participar. Sublinho, também, que tanto a instituição bem como os participantes serão anónimos e que nenhum tipo de imagens ou fotografias serão obtidas ao longo do estudo.

Em caso de dúvida ou necessidade de mais esclarecimentos, peço-vos que se sintam livres de me contactar através do email susanamedina@hotmail.com.

Agradeço, desde já, a vossa disponibilidade, atenção e apoio prestados até à data e fico a aguardar a vossa autorização.

Lisboa, 23 de setembro de 2021
Susana Passos Medina

Prof.ª Dra. Carolyn Leslie
Orientadora de Estágio
FCSH, Universidade Nova de Lisboa



Eu, _____ Diretor(a) de _____
_____ declaro que fui informado(a) dos objetivos do projeto intitulado *Assess me and I learn: Formatively assessing oral communication in the FL classroom* (Avalia-me e eu aprendo: Avaliação formativa da comunicação oral na sala de aula de língua estrangeira) e autorizo os alunos da turma do 3º ano a participar no estudo.

Data: _____

Assinatura: _____

Hello.

Sei que já me conheces, mas volto a apresentar-me. Sou a Susana e, como sabes, estou a estudar para ser uma super especialista a dar aulas de Inglês ao 3º e 4º ano. Mas, para poder acabar o meu curso, preciso da tua ajuda! Vou desenvolver um pequeno projeto com a tua turma e gostava muito de poder contar com a tua participação e saber a tua opinião!

Um projeto? Isso deve ser difícil!  Mas o que tenho de fazer?

Apenas 4 coisas:

1. Preencher um questionário (mas só no mês de Dezembro)
2. Todos juntos vamos pensar em como se deve avaliar a oralidade
3. Autoavaliar-te e avaliar os colegas depois de realizarmos algumas tarefas em sala.
4. Participar normalmente nas aulas, ou seja, falar muiiiito inglês!!

É simples, não achas?!



Outros pontos importantes que precisas de saber:

- a) Não vou usar o teu nome.
- b) Não vou tirar nenhuma fotografia tua.
- c) Vou enviar uma carta aos teus pais a pedir autorização para que possas participar e a explicar-lhes tudo sobre este projeto.
- d) Não és obrigado/a a participar e podes desistir a qualquer momento! No problem!
- e) Se tiveres dúvidas, por favooooor vem falar comigo e pergunta-me o que quiseres!

THANK YOU

Susana Medina



com um (X)

Assinala

Quero participar no projeto da Susana.

Sim

Não

Appendix D: Grid 1.A

Grid 2.A
Teacher's name: Susana Medina
Activity: Check in

N°	Name	Date	1 word utterances	Uses more L1 than L2	Uses telegraphic language	Uses full sentences	Examples of utterances	Obs. (asks how to say something, helps others, etc)
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Extra obs.:

Teacher's name:

Date(s):

Class: 3rd grade

Activity/activities:

Vocabulary in use:

Structures in use:

Other relevant aspects:

Assessment scale (symbols):

(+) Achieved/ consolidated

(+/-) Partially achieved/ Consolidating

(-) Not achieved/ not consolidated

(v) Check - indicates what the learner is able to do

Abbreviations:

Acc.: Accuracy

Flu.: Fluency

Pron.: Pronunciation

R. Prob.: Recurrent problems (mistakes, words the learners do not know, words the learner never includes in the structure, etc)

L1: Portuguese

L2: English

Nº	Name	Vocabulary			Structures				Pron.	Comments/notes
		Objects	Prep. place	R. Prob.	Mixes L1 with L2	Uses telegraphic sentences	Uses complete sentences	R. Prob.		
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

Teacher's name:

Date(s):

Class: 3rd grade

Activity/activities:

Vocabulary in use:

Adjectives:

Structures in use:

Verb to be:

Verb have got:

Other relevant aspects:

Assessment scale (symbols):

(+) Achieved/ consolidated

(+/-) Partially achieved/ Consolidating

(-) Not achieved/ not consolidated

(v) Check - indicates what the learner is able to do

Abbreviations:

Acc.: Accuracy

Flu.: Fluency

Pron.: Pronunciation

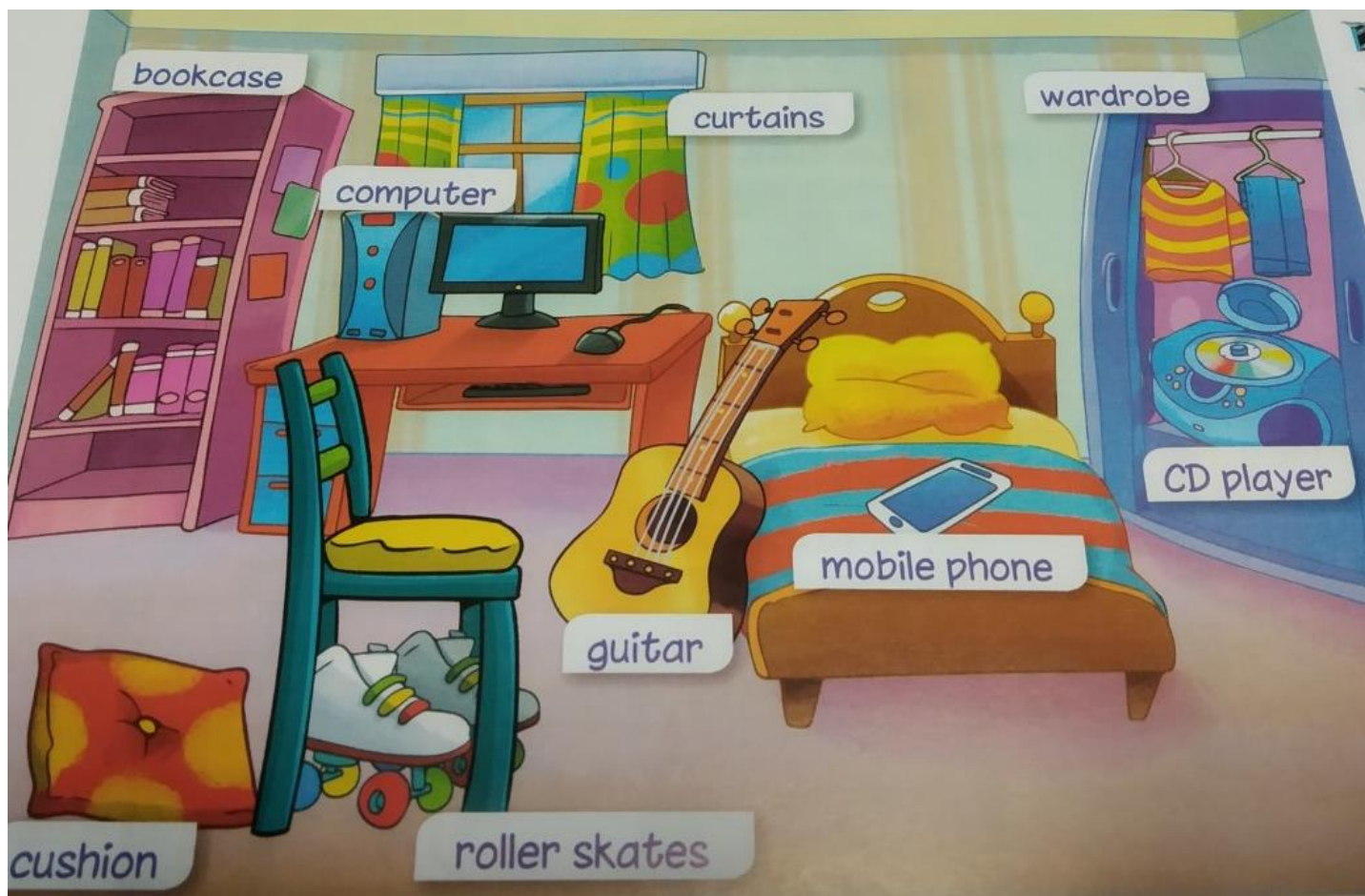
R. Prob.: Recurrent problems (mistakes, words the learners do not know, words the learner never includes in the structure, etc)

L1: Portuguese

L2: English

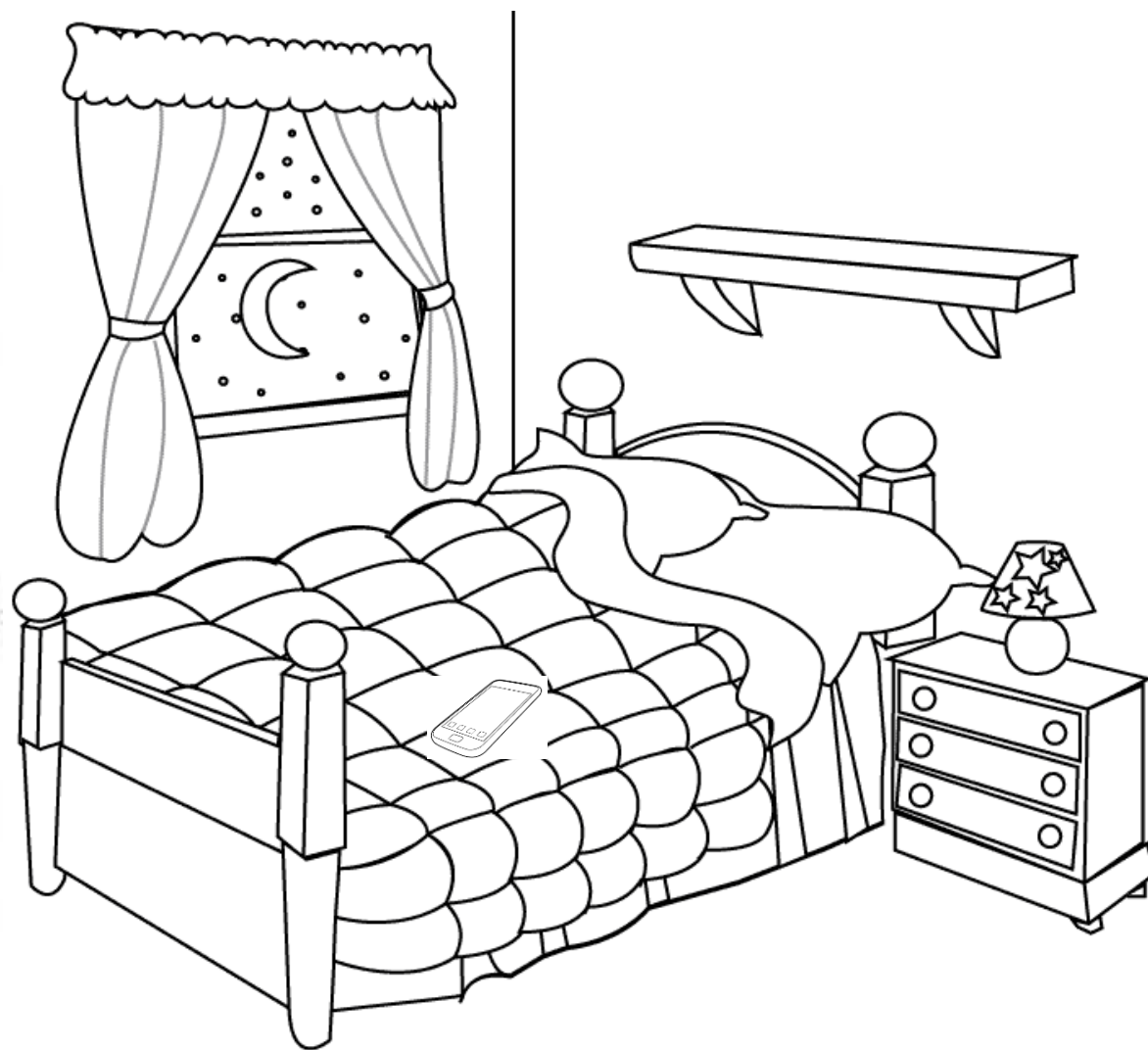
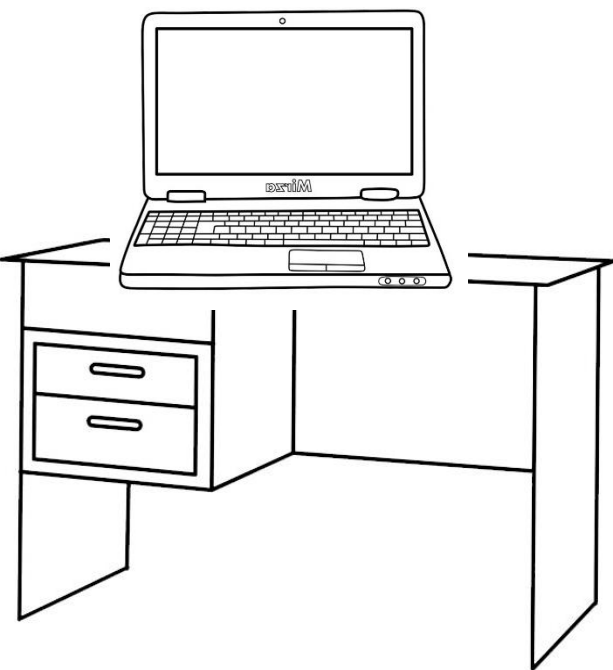
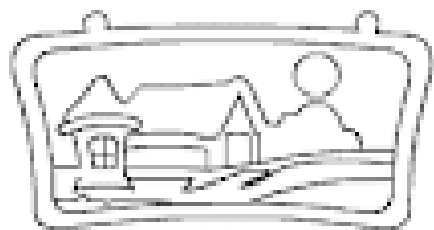
Nº	Name	Vocabulary		Structures			Pron.	Acc.	Flu.	Comments/notes
		Adjectives	Physical description	Mixes L1 with L2	Uses telegraphic sentences	Uses complete sentences				
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										

Appendix G: Oral interaction task 1 – Where is it?

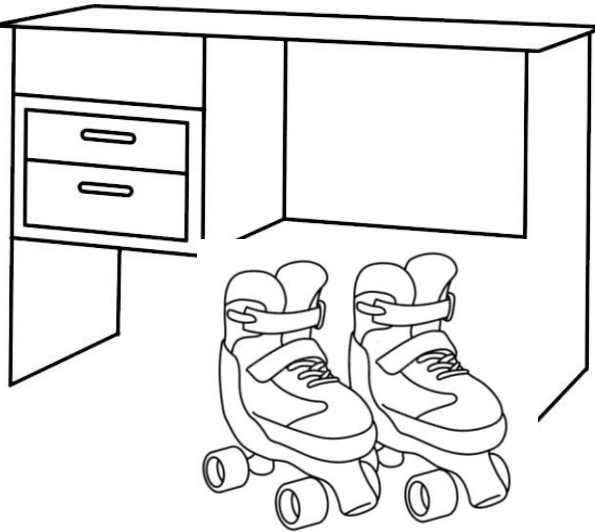
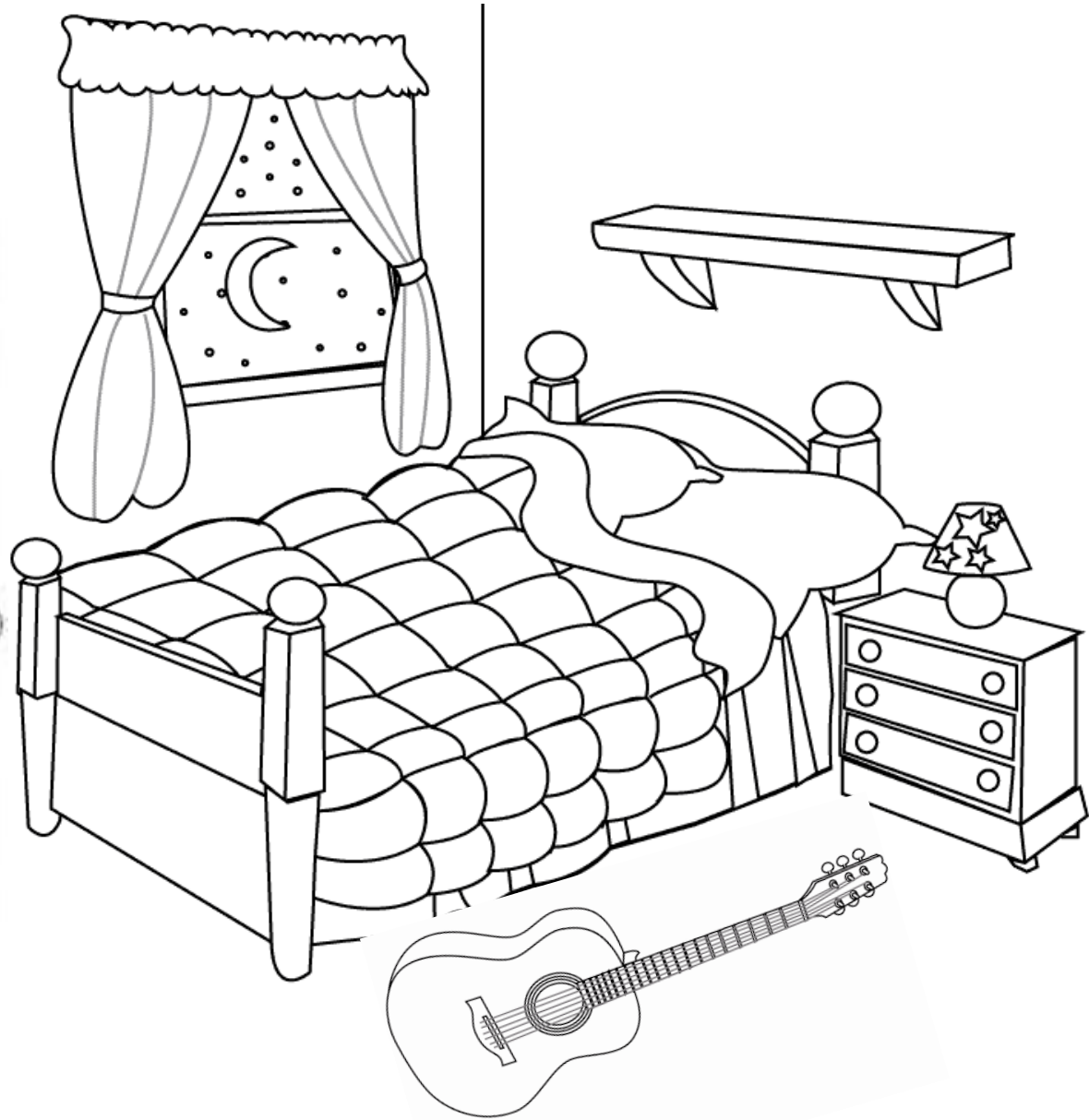
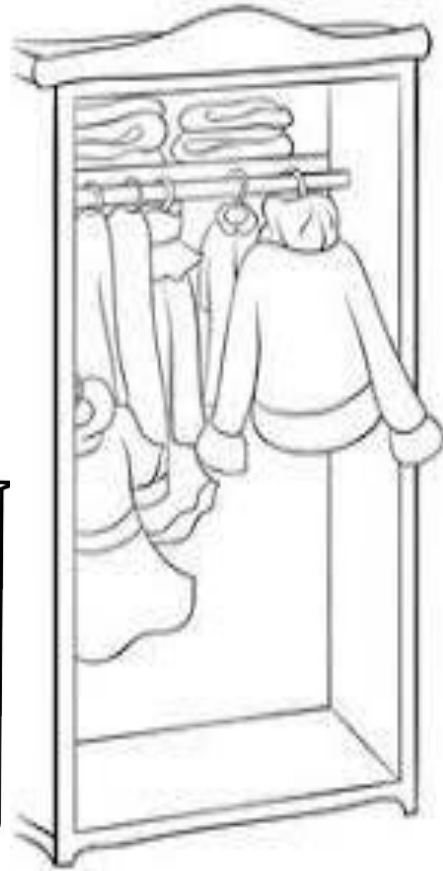
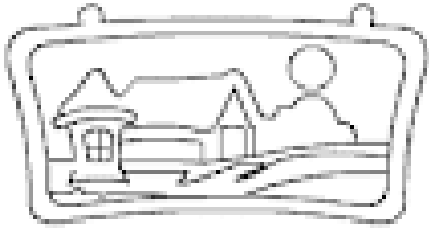


Appendix H: Oral interaction task 2 – What's missing?

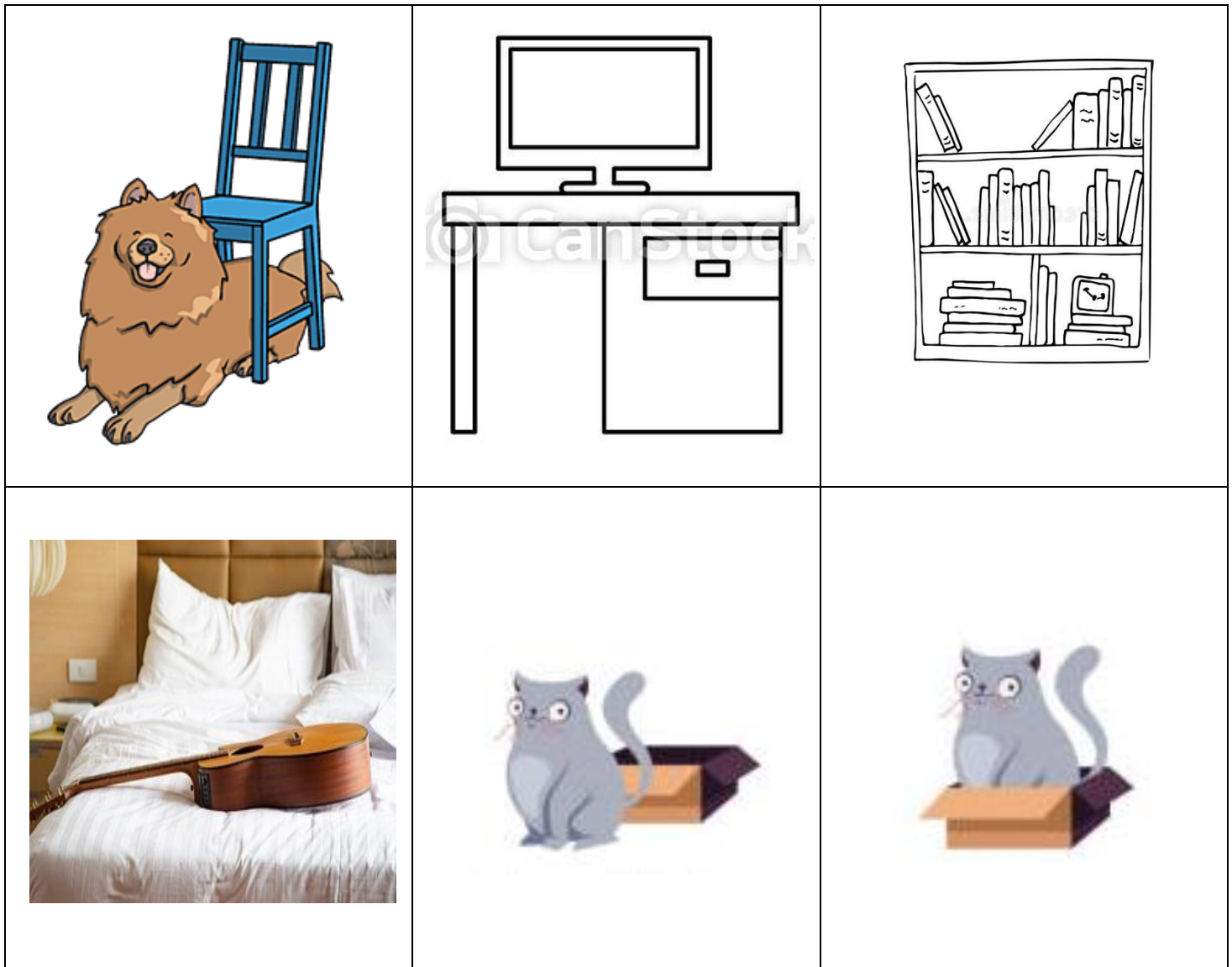
A



B




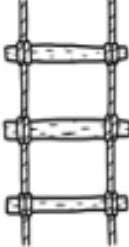

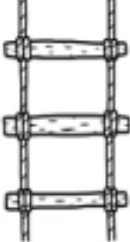


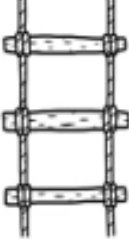

Appendix I: Oral interaction task 3 – Pictionary



Appendix J: Oral interaction tasks 4 and 6 – Mini cards and Pictionary

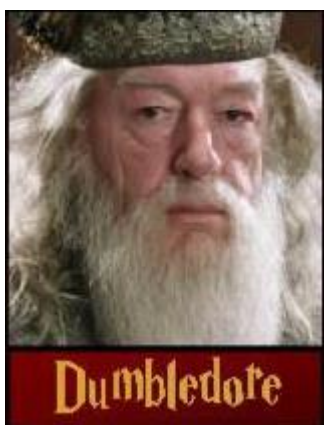
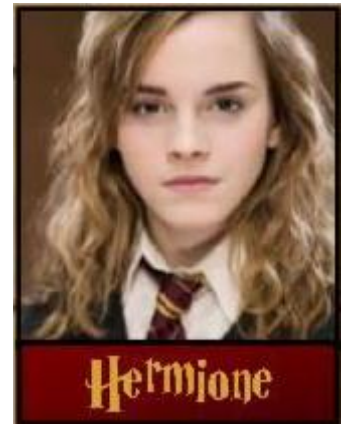
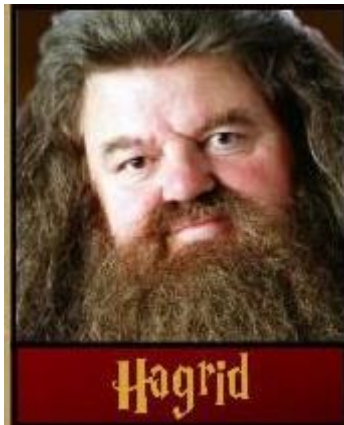
Tall	Short	Ugly
Beautiful	Weak	Strong
Fat	Thin	Old
Young	Kind	Mean

Appendix K: Oral interaction task 5 – Board Game

START	What's the opposite of tall?	What's the opposite of fat?	Go down the stairs	Take one card. Say 3 things about it.	Ups! You are in prison! Wait one round to play	How do you say <i>corajoso</i> in English?
 <h1 data-bbox="465 491 981 596">Adjectives</h1>					How do you say <i>feio</i> ou <i>feia</i> in English?	
How do you say <i>gordo</i> or <i>gorda</i> in English?	Take one card. Say 3 things about it.	What's the opposite of ugly?	How do you say <i>forte</i> in English?	Take one card. Say 3 things about it.	Go down the stairs	Ups! You are in prison! Wait one round to play
What's the opposite of weak?						
How do you say <i>pobre</i> in English?	Go up the stairs	What's the opposite of thin?	Take one card. Say 3 things about it.	Ups! You are in prison! Wait one round to play	What's the opposite of strong?	FINISH!

Appendix L: Oral interaction task 7 – Guess Who

Guess Who





My name is _____

PROJECT – This is...

Mandatory information:

<ol style="list-style-type: none"> 1. Name 2. Age 3. Nationality 4. Physical description (eyes, hair, nose, mouth) 5. Adjectives (tall, fat, young, kind, etc)

You can also talk about...

1. Family
2. Likes and dislikes
3. Favourite things (pets, food, colour, etc)

Read the examples to help you

Beginning	Middle	End
<ul style="list-style-type: none"> ❖ Hello, my name is... ❖ This is our group ❖ We are going to talk about... 	<ul style="list-style-type: none"> ❖ This is... ❖ She is five years old ❖ He is tall and fat ❖ She has got glasses ❖ He has got curly brown hair ❖ Mary likes pizza but Mary doesn't like broccoli ❖ Lucas has got a sister and a brother 	<ul style="list-style-type: none"> ❖ We hope we liked our presentation ❖ The end ❖ Thank you for listening

Appendix N: Grid 3.A

Observation Grid: 3.A (Oral presentation)

Teacher's name:

Date:

Activity:

Language in use:

Structures:

Names	Content - vocabulary		Structures		Pron.	Flu.	Acc.	Comments and observations
	Physical description	Adjectives	Verb to be	Verb Have got				

Appendix O: Grid evaluation questionnaire

Scale: 1 – Agree; 2 – Partially agree; 3 – Disagree

Name:

Grid:

Date:

	Agree	Partially agree	Disagree	Comments
1. Style and structure				
The grid was visually simple to understand	1	2	3	
The grid was easy to interpret	1	2	3	
The grid was easy/simple to fill in	1	2	3	
The grid has enough space for me to write down some students' utterances that I feel are important	1	2	3	

	Agree	Partially agree	Disagree	Comments
--	-------	-----------------	----------	----------

2. Assessment parameters/criteria				
The parameters in the grid were adequate for this activity (vocabulary, structures in use, creative use of the language)	1	2	3	
I feel that there were parameters missing	1	2	3	Which ones?
The wording of the parameters was clear	1	2	3	
I need a lot of lessons to fill in the grid because there are too many parameters	1	2	3	
There were too many parameters and I felt overwhelmed	1	2	3	
There were a lot of parameters but I feel that, with time, I can get acquainted with all of them and it will be easier	1	2	3	

	Agree	Partially agree	Disagree	Comments
--	-------	-----------------	----------	----------

3. Aims				
Using the grid made me more aware of what to look for when observing students speak	1	2	3	
Using the grid provided a clearer focus for my observation and assessment	1	2	3	
Using the grid helped me understand in which parameters my students are struggling	1	2	3	
Using the grid helped me understand in which parameters my students are excelling	1	2	3	
Using the grid made me feel more confident when planning future lesson	1	2	3	
Using the grid made me feel more confident if I need to justify the grade of a particular student to the stakeholders (parents, school).	1	2	3	
Using the grid was useful for me.	1	2	3	Why?

Appendix P: Excerpt from the teaching journal

Week 6 (9/10) (3P)

- fiz um follow-up a seguir à mini card activity - pq senti q havia erros nas perguntas ent fui perguntar: full no a question you used in the game.
Dizem: Ade you... e a preverte para corrigir q toda a turma.
- O trocaram estações / vão trocar na próxima aula.
- Davonei ⊕ tempo do que queria e as estações q mereciam ⊕ tempo do que planeado

Pergunto: quando diz aqui I know the vocabulary? o que é que isso significa?

1. O vocabulário é as palavras novas e estas a aprender

M: como por exemplo?















M.R: beautiful, tall / fat "

Lúlia: "Vocabulary of adjectives"

2. OK. E quando diz I used complete sentences? que tipos de frases será com I beautiful?

Questionário

Pinta a tua resposta.

	Sim	Mais ou menos	Não
1. Gosto de falar inglês			
2. Gosto de fazer a minha autoavaliação da oralidade Porquê? _____ _____			
3. Fazer autoavaliação da oralidade faz-me sentir mais confiante a falar inglês Porquê? _____ _____			
4. Fazer autoavaliação da oralidade ajuda-me a aprender a falar inglês Porquê? _____ _____			
5. Fazer autoavaliação da oralidade é importante Porquê? _____ _____			

6. Quando a Susana avalia a minha oralidade sinto-me...

Tranquilo/a

Nervoso/a

7. Quando faço a minha autoavaliação da oralidade sinto-me...

Responsável

Confuso/a

Appendix R: School’s official assessment grid

Nome	Competências específicas										Nota Final			
	Participação nas atividades	Testes/Projetos		Check in/ Speaking	Reading corner/ Writing	Reading	Listening	Empenho/ Registos dossier	Pontualidade/ Assiduidade	Material/ Organização	Cumprimento de regras	Total	Nota	Auto-avaliação
6	4		2	2	2	1	1	0,5	0,5	1	20	(5)		
														2 (Insuficiente) - A partir de 8,
														3 (Suficiente) - A partir de 11,
														4 (Bom) - A partir de 14,60
														5 (Muito Bom) - A partir de 17,

Appendix S: Photo of the puppet (Smelly Cat)



Appendix T: Excerpt from classroom conversation about self-assessment

Teacher: Quando diz aqui ‘I know the vocabulary’, o que é que isso significa? [When it says here ‘I know the vocabulary’, what does it mean?]

Student 1: O vocabulário é as palavras novas que estamos a aprender. [The vocabulary is the new words we are learning]

Teacher: Como por exemplo? [For example?]

Student 1: Beautiful, tall, fat...

Student 2: Vocabulary of adjectives.

Teacher: Okay. E quando diz ‘I used complete sentences’? Que tipos de frases? Será ‘Can I beautiful’? [Okay. And when it says here ‘I used complete sentences’? What types of sentences? Could it be ‘can I beautiful?’]

Student 2: No! Are you beautiful? Are you tall?

Student 3: She is tall.

Teacher: Boa. E ‘ele é forte’? [Good. And ‘he is strong?’]

Student 3: He is strong.