Chapter #4

THE ART AND SCIENCE OF ASSESSMENT – CONNECTING CLASSROOM-BASED SPEAKING ASSESSMENTS TO TEACHERS' PRACTICE

Rúben Constantino Correia
CETAPS NOVA FCSH – Centre for English, Translation and Anglo-Portuguese Studies, Portugal

ABSTRACT
Speaking has been increasingly promoted in language syllabuses and curriculums, both in Portugal and internationally, as one of the major aims of foreign language teaching. Naturally, the importance of oral skills has led to increasing research in this area, with the focus largely on the need to measure ability and the best way to do it. Unsurprisingly, considerable attention has been drawn both to assessment and the context in which it operates. However, the unique features of speaking make it the most challenging skill to assess. Bearing this in mind, and my role as both researcher and teacher with a vested interest in speaking, I spent almost a full school year at a Portuguese public school cluster doing classroom observation in an attempt to chart: a) – typical classroom interactions between learners / teachers and learners / learners, and b) – the general nature of most speaking events taking place in the classroom, including that of assessment. Findings seem to evidence that Portuguese EFL teachers appear to be at odds with designing suitable assessment procedures for monitoring students' progress. There is a narrow view of assessment as synonymous with testing, and thus the grading function, which largely contributes to the dominance of summative assessment over formative assessment.

Keywords: English as a foreign language, speaking, assessment, learning, learning-oriented assessment.

1. INTRODUCTION

Learning a foreign language, as a rule, is seen by experts (anthropologists, sociologists and professors/teachers) as a major asset for global understanding and the mobility of people. English is found at the top of the pyramid as the number one language to achieve these goals. Nowadays being able to express oneself proficiently and intelligibly in English is decisive for learner-users who want to thrive both academically and professionally.

The search for more effective ways of teaching English as a Foreign Language (EFL) gave rise to different teaching methods/approaches on both sides of the Atlantic over the past century. From those, the Communicative Language Teaching (CLT) approach emerged as the one adopted by most practitioners, marking "a major paradigm shift within language teaching in the twentieth century, one whose ramifications continue to be felt today" (Rodgers & Richards, 2001, p. 151). CLT argues for genuine communicative exchanges through activities designed to develop the students' ability to use language appropriately and meaningfully. Along with the rise of CLT, the importance of oral skills in language syllabuses and curriculums grew and led to increasing research in this area, with the focus largely on the need to measure ability and the best way to do it. Considerable attention has, then, been drawn both to assessment and the context in which it operates. In view of these premises, this study sought to explore the knock-on effect of Portuguese EFL teachers' practices on classroom-based speaking assessments, which translated into one central research question:

→ Are speaking assessment tasks aligned with the ongoing classroom activities used by the teachers?

The primary goal of classroom-based speaking assessments is to support learners' language learning trajectories. Assessment results must be used to inform instruction, identify areas for improvement, and foster a positive and supportive language learning environment. This is all the more important in EFL contexts like the Portuguese where the learners attending lower-secondary education (5^{th} – 9^{th} grade) should be given ample opportunities to progress through five years of language instruction.

2. BACKGROUND

2.1. Problem Statement

Speaking has unique traits that make it the most distinctive and probably the most difficult skill to assess in classroom-based contexts. Unlike writing, speaking is done spontaneously, greatly restricting the possibility to plan one's discourse before processing and producing it. Thus, the teacher has to judge, in real-time, production and/or interaction related to several aspects of what is being said (range, pronunciation, accuracy, fluency, interaction, coherence). Furthermore, in Portugal the assessment of speaking proficiency seems to face a major challenge – the apparent reluctance of Portuguese state schoolteachers to address it. Most students studying English at the lower levels (5^{th} up to 9^{th} graders) are overloaded with grammar instruction and exercises, usually done via course-books, quizzes or worksheets. Clearly, the emphasis given to linguistic competence outweighs that given to linguistic performance, which in turn hinders the students' speaking proficiency and the assessment process itself.

Right from day one, Portuguese learner-users are faced with the strict grip of this type of assessment, that of diagnostic assessment, which is a common practice usually done via testing related to their past learning. Theoretically, it aims to ascertain the learner's strengths and weaknesses, although it is the latter that is acted upon by teachers. All their efforts seem to be directed at what the students cannot do. This type of assessment neither does what it is meant to do – identify strengths and weaknesses – nor is it designed as a diagnostic tool. Firstly, it hardly ever covers all the major skills, as speaking is usually omitted and secondly, it resembles an achievement test instead of a diagnostic one. As a result, students are not assessed to check what they can or cannot yet do, but instead are assessed on their understanding of language features from previous years with little or no valid feedback available for students or teachers. The effectiveness of diagnostic assessment is undermined and does not contribute as it should to successful learning. Pertaining to the scope of this study, Kazemi and Tavassoli (2020), who investigated the impact of diagnostic assessment on the speaking skills of EFL learners, found that the group of participants who received diagnostic feedback based on their initial assessment showed significant improvement of their speaking proficiency when compared to that of their control group counterparts. Even though the study was conducted with adult learners, it strengthens support for the positive impact of diagnostic assessment on learners' (speaking) achievement. It may even resonate deeper with the cohort addressed here. Identifying strengths and weaknesses is, as Nikolov and Timpe-Laughlin (2021) point out, "particularly important for YLs, as they need substantial encouragement in the form of frequent, immediate, and motivating feedback on where they are in their learning journey so that their vulnerable motivation can be maintained […]" (p. 24).

The root of the problem may lie in the confused nature of diagnostic testing in past and recent literature. Very often diagnostic and placement tests are taken as interchangeable terms serving the same purposes, when in fact they are not. Brown implies they can be indistinguishable, and a placement test can serve the same aim as a diagnostic test (2004, pp. 46, 47). As mentioned above, the latter is supposed to identify strong points and weaknesses, whereas the former is meant to help teachers place their students in a certain proficiency level appropriate to their abilities. Alderson (2007) notes how neglected diagnostic testing is in language testing research: "[…] there is virtually no description, much less discussion, of what the underlying constructs might be that should be operationalized in valid diagnostic tests" (p. 28). In addition to being limited, the information about diagnostic assessment is also rather unclear, leading to multiple interpretations and misconceptions. In the light of such a lack of rationale, Blood (2011) suggests that "in the broadest sense, then, diagnostic second language (L2) assessment refers to any L2 assessment practice, whether in the form of a formal written test or informal teacher questioning, that yields diagnostic feedback" (p. 57). As it happens, one avenue for future research on assessment is that of diagnostic assessment. Several key aspects such as the functions and constructs that should underlie it remain obscure. If we believe that prior knowledge shapes new learning, we need to redefine our conceptualisation of diagnostic assessment.

Adding to this predicament, summative assessment has become the most dominant type of assessment. Ideally, formative assessment would pick up where diagnostic assessment finishes, through a range of tasks designed to provide further learning opportunities, learners' strong points would be reinforced, and their weaknesses would be improved. Nonetheless, this is not exactly what happens because of the importance given to summative assessment. Teachers tend to undertake assessments which are a replica of Low Stakes testing. Besides, in a similar fashion to diagnostic assessments, summative assessments in Portuguese EFL classrooms usually omit speaking. In fairly analogous EFL contexts (Greece and Cyprus) to the Portuguese, Tsagari (2016) found the same pattern. Assessment was predominantly made of vocabulary and grammar, followed by writing. Listening and speaking were hardly assessed if at all. In Cyprus, speaking was completely absent. Assessment is, then, carried out not to support learning, but by grading tests that normally do not provide useful information feedback. There is "a widespread public expectation of assessment, and while it could be argued that this is insufficiently future-oriented, it would be difficult to mount a case which involved shifting existing well-established perceptions of this purpose" (Boud & Falchikov, 2006, p. 401). The perception of assessment as marking/grading runs deep amongst students as well, whose focus is driven by the desire for higher grades rather than learning. Thus, students' efforts are not volitional, but motivated by the demands of assessment. Input is simply memorized and not transformed into real operational knowledge, prior knowledge is not related to new knowledge, and concepts are simply applied mechanically without reflection. In Portugal, like in many other EFL contexts, summative assessment practices have prompted the devaluation of actual learning. Although we need summative assessment to make decisions, learners' results should contribute to forthcoming learning.

2.2. Speaking's Inherent Character

Speaking is a fundamental aspect of human communication and plays a crucial role in how we interact with others and express our thoughts, ideas, and emotions, and make connections with the world around us. It is a dynamic process that involves both speakers and listeners. However, speaking has not always figured centrally in the field of applied linguistics. Even when it became particularly implicated in language teaching methodology,

speaking was inaccurately seen as similar to writing. Thus, the overall nature of speaking was disregarded in favour of the long writing tradition of teaching and learning a foreign/second language.

Researchers have fairly recently started to dedicate similar attention to spoken language as they did to written language, only to realise that they differ significantly from each other. Unlike writing, where a shared spatio-temporal ground is by definition non-existent, speaking is done in real-time, narrowing greatly the possibility to plan, edit or revise one's discourse before processing and producing it. In addition, the speaker must master and mobilize an array of linguistic knowledge – vocabulary, sound system (segmental features), suprasegmental aspects like stress, intonation and rhythm and language functions – alongside the kinesics usually related to spoken language, to avoid extensive hesitation or communicational breakdowns. Unsurprisingly, speaking seems to be more challenging than writing, or reading for that matter.

Speaking is broadly characterised by the use of incomplete sentences (known as ellipsis) to avoid unnecessary effort, connected or not with conjunctions, what Luoma (2004, p. 12) conceives of as idea units, short turns between interlocutors together with simple interrogative structures, manipulation of strategies to gain time to speak, such as fillers and hesitation markers, repetitions and rephrasings (to correct, alter or improve what has been said by the speaker who is taking the floor or by previous speakers), fixed conventional phrases and use of informal speech (simpler syntax to make improvisation easier) due to its spontaneity and purposes. These devices are employed to both facilitate speaking and compensate for difficulties that (may) arise. Indeed, disfluencies and consequent repairs are quite natural in spoken language. Spoken language is commonly less lexically dense and fragmented, resulting in a high frequency of pro-forms, incomplete clauses, and a low frequency of information-carrying words. The fact that speaking is traditionally an interactional activity contrasts with the detached stance of most writing. While the writer embarks on a solo endeavour, and his/her audience is not present and often is not known, the speaker is directly involved with his/her listener(s), the subject matter, and the context. This involvement is marked by the use of first-person pronouns, vocative forms, and attention signalling. Table 1 (based on Biber, Conrad, & Leech, 2002) identifies and summarises a variety of social and situational features of speaking. The set of features presented are intrinsic to the time-bound nature of speaking's processing conditions. The shape and nature of speaking is intimately connected to its socio-psychological processes, which clearly impact on language use and are responsible for most of the differences between spoken and written language.

Table 1.
Circumstances and Features of Spoken Language.

Social and Situational Traits	Association with Conversational Traits
Typically occurs in a shared context	• High frequency of (personal) pronouns; • Low frequency of nouns; • Use of substitute pro-forms and ellipsis; • Reliance on deictic words; • Use of fragmentary components (frequently inserts).
Avoids elaboration and/or specification of meaning	• High frequency of verbs (especially primary and modal verbs); • Low frequency of elaborated noun phrases; • Use of complement clauses: that and wh-; • Reliance on function words; • Considerable usage of vague language (often hedges).
Is marked by interaction	• Abundance of negative utterances; • Large quantity of question-answer sequences. Responses are often elicited via question-tags; • Profusion of attention-signalling forms; • Common use of vocatives; • Frequent use of discourse markers.
Expresses stance	• Heavy reliance on endearments (e.g., dear), interjections (e.g., wow), exclamations (e.g., good for you!), evaluative predicative adjectives (e.g., nice) and stance adverbials (e.g., fortunately).
Takes place in real time	• Occurrence of disfluencies; • Use of reduced forms (e.g., "gonna"), contractions and elision processes; • Usage of a restricted and repetitive repertoire.
Employs vernacular phraseology	• Style is eminently colloquial; • Occurrence of regional dialect forms.

Considering the cohort to be studied (9th grade students), a final point to be made about the nature of speaking regards age. The language used by young speakers has some distinctive features, when compared to adult speakers, that need to be taken into account by their teachers. For instance, youngsters tend to be more informal than adults, which translates

into using less modality while producing the language. It is uncommon for adolescents to use could or might. They are much more likely to use their present corresponding forms can and may, especially can. Complex modalized sentences do not seem to fit this specific group of speakers. The point I am trying to make is the necessity to rethink if it is fair or even realistic to demand of our adolescent learners styles of speaking that do not conform to their age span and they do not use in their L1, nor that we will ever hear from the mouths of their NS counterparts.

2.3. Key concepts

Measuring and understanding student learning outcomes is a complex but crucial process that involves the assessment of learners' language proficiency, skills and knowledge. Gaining a solid knowledge of assessment has the potential to empower teachers to design and implement effective classroom-based assessments that accurately measure learners' language abilities, encourage language development and support meaningful learning outcomes. Unsurprisingly, assessment has become a popular buzz word but "sometimes misunderstood term in current educational practice" (Brown, 2004, p. 4) and for this reason a distinction between the terms assessment and testing must be made. While these terms are often used interchangeably, they represent distinct methodologies with different purposes and implications. "The debate may seem abstract and theoretical, but it is of considerable importance to classroom teachers of language because it impacts not only how learners are assessed, but how they are taught" (Hamp-Lyons, 2007, p. 488).

Testing is an administrative product-oriented procedure, usually imposed by the teacher, that occurs at specific moments with the purpose of measuring second/foreign language knowledge for scoring and grading. So, tests tend to create immense pressure on learners, leading to heightened stress and anxiety. The fear of failure and the potential consequences on their academic future can impede learners' cognitive functions, negatively impacting their spoken language performance. As advocated elsewhere (Correia, 2015), some speakers' frequent pauses and hesitations, resulting in reduced intelligibility, may positively correlate with speaking anxiety. The words of a Croatian 12-year-old learner cited in Mihaljević Djigunović (2019) tellingly illustrate such view: "Each time our teacher announces a test, I panic. While preparing for the test at home, I feel nervous all the time" (p. 25). Tests are often a norm-referenced instrument – scores are compared amongst students, used to determine individual ability, or demonstrate mastery of a given skill, and offer limited information to identify areas for improvement because they tend to be "one-off" events of speaking proficiency. Again, the issue of feedback must be highlighted. In a similar fashion to diagnostic assessment, little or no valid feedback for learners is obtained from these procedures. Teachers do not critically discuss and review these results with their learners and thus do not provide (non-threatening) information about what they are doing well and what needs to be tackled with more enthusiasm. There is little reflection on what is being tested and why is it being tested. Nikolov's (2017) research in Hungary goes further to show that many teachers failed to see the added value of timely feedback. Ultimately, when a teacher gives a test, s/he is obtaining a narrow sample of the test-taker's performance in a specific domain that does not account for the progress made (or not) based on that performance. For the less proficient learners we are simply telling them they lack ability without pointing to ways for improvement. Perhaps, I would argue, some EFL practitioners do not feel comfortable going beyond what they experienced as learners themselves and now perpetuate this as teachers, while others are simply caught in a predicament between what they believe is best and more helpful for their students (moving away from summative assessment done via testing) and the pressure to achieve the success percentages set by school

boards. Throughout literature we can find examples (Hamp-Lyons, Hood, & MacLennan, 2001; Davison, 2004; Bonner, Torres Rivera, & Chen, 2018) of this conflicting pattern between teachers' beliefs and system pressure.

On the other hand, assessment is an ongoing process-oriented approach that takes many different forms. One of these forms is tests. Thus, testing is a subset of assessment and should be seen as one of the many methods available for assessing students' verbal performance. Teachers may employ a diverse range of assessment methods and tools that align with specific learning objectives and students' unique needs, whilst enabling their differences. In view of the limited nature of tests, alternative assessment procedures such as self-assessment, peer-assessment, portfolios, performance assessment, observation, etc., have been advocated at different times by some experts like Shohamy (1997), Bachman (2002), Hamp-Lyons (2007) and Nikolov and Timpe-Laughlin (2021). I prefer to consider these methods, tests included, as simply assessment, preferably when used in an integrated fashion to help improve learners' speaking skills. Assessment is often a criterion-referenced measurement – students' performance being compared against a set of criteria, used in educational contexts to monitor students' strengths and weaknesses. It is operated in a systematic way for the purpose of helping teachers trace the students' individual learning trajectories throughout the school year. Assessments serve as tools to identify learning gaps and draw inferences that the teachers can rely on about the students' achievements, and to make the necessary adjustments in the teaching-learning environment, i.e., using assessment results to change practices and adapt teaching strategies which in turn assist learners to improve their speaking proficiency by meeting their individual learning needs. In a nutshell, "assessment is the systematic collection, review, and use of information […] undertaken for the purpose of improving student learning and development" (Banta & Palomba, 1999, p. 4), entailing careful planning, implementing and acting upon the results. Unlike testing, which focuses on grading and final outcomes, assessment emphasizes timely continuous feedback that helps learners recognize areas for improvement while the learning experience is still fresh in their minds. Involving students in the assessment process enhances their metacognitive awareness, promoting a deeper understanding of their strengths and weaknesses. Indeed, the iterative feedback loop between students and teachers allows for continuous progress and targeted support. Learners are more likely to demonstrate higher levels of motivation, engagement, and perseverance when they are actively engaged in the assessment process and can make sense of the information given and use it to enhance their spoken proficiency. Assessment goes beyond the question of how much the students have learned; instead, it asks how they learned and what can be done to improve their learning.

Recognizing the differences between assessing and testing is fundamental for teachers to make informed decisions regarding their learners. While both approaches play a role in identifying student learning, assessment's flexible, holistic, and learner-centred nature offers a more comprehensive view of learners' competences. Emphasising assessment whilst deemphasising testing practices can lead to a more comprehensive and meaningful understanding of the learners spoken language proficiency that empowers them and supports their long-term ability and communication success. However, turning this narrow view of assessment as synonymous with testing around seems as yet an optimistic but blurred vision.

2.4. Assessing with a Learning-Oriented Frame of Mind

From the beginning of the twenty-first century, a new framework has steadily gained ground in the field of educational assessment, the learning-oriented assessment approach. This innovative view of educational assessment pedagogy "holds that for all assessments, whether predominantly summative or formative in function, a key aim is for them to promote

productive student learning" (Carless, 2009, p. 80). Hence, whatever form the assessment takes it must be a means of supporting learning and, simultaneously, to acknowledge its centrality. Implementing a learning-oriented assessment approach to speaking "involves the collection and interpretation of evidence about performance so that judgments can be made about further language development" (Purpura, 2004, p. 236) to promote knowledge. Analysing Purpura's words carefully, we conclude that evidence is the core ingredient of learning-oriented assessments. After being collected from multiple sources, evidence helps teachers to monitor students' progress, shows students' acquisition (or otherwise) of what is being taught, and provides meaningful feedback for students and teachers.

Carless (2009) summarizes learning-oriented assessment in three simple principles. Bearing these principles in mind, teachers will be able to engage learners in productive assessment activities. "Principle 1: Assessment tasks should be designed to stimulate productive learning practices amongst students; Principle 2: Assessment should involve students actively in engaging with criteria, quality, their own and/or peers' performance [*sic*]; Principle 3: Feedback should be timely and forward-looking so as to support current and future student learning" (p. 83). To achieve these intended purposes, appropriate tasks should be designed, students have to be involved and feedback has to be significant. First, learning tasks should be conceptualized as assessment tasks and vice-versa, encompassing the anticipated learning goals by promoting interactional authenticity, a reflection of the real-world and collaborative work. Besides mirroring real-life language usage, authentic tasks expose learners to natural language patterns, expressions, and contexts. Engaging with authentic speaking tasks enables learners to navigate various communication situations they might encounter in the target language environment, thereby enhancing their spoken language proficiency. Second, students must be given the opportunity to understand the criteria and standards applied to their work, enabling them to accurately judge whether they meet these criteria and standards or not. "The conceptual rationale for peer assessment and peer feedback is that it enables students to take an active role in the management of their own learning" (Liu & Carless, 2006, p. 280). Indeed, a paradigm shift in focus must occur from merely testing learners' performance to actively involving them in the assessment process. Learners should be encouraged to reflect on their learning progress, set goals, and take ownership of their learning. Yet, traditional beliefs over EFL teaching and learning continue to go counter to these expectations. Another of Nikolov's (2017) findings shed light on the teachers' disagreement with the use of peer and self-assessment. Teachers mistrusted their learners' ability to do so as well as their honesty. Third, feedback must be timely, relevant, and able to be acted upon by the students, i.e., it functions as feedforward. If it does not help students close the gap between their expected learning outcomes and the present state, it does not really qualify as feedback. If we are to succeed as EFL teachers, we have to recognise that timely and constructive feedback is the heart of classroom-based assessment. This goes beyond the traditional feedback practices discussed above, it guides learners towards future improvements (feedforward). Feedback(forward) should be given at regular intervals during the learning process, enabling students to make real-time adjustments and enhance their understanding of the subject matter.

Learning-oriented approaches to speaking should not be concerned only with measuring ability, but also with the actual learning of pronunciation (segmental and suprasegmental aspects), vocabulary, language functions, register, turn-taking and breakdowns compensation. Thus, teachers must make sure that learning/assessment tasks represent spontaneous, real-life spoken interaction and target the speaking aspects the learner-users are supposed to use, i.e., how assessment relates to and can help promote speaking acquisition.

3. METHODOLOGY

The study followed a qualitative driven approach, drawing on two complementary sources of data – classroom observations and field notes. The rationale for conducting qualitative research in this study was threefold: first, due to the exploratory nature of the study; second, because it allowed a longitudinal detailed examination of the unit of analysis (regular weekly intervals were made over the course of a school year); and third, for its potential representativeness and contribution to understanding the connection between typical classroom interactions between learners / teachers and learners / learners and the general nature of most speaking assessment events occurring in the classroom.

3.1. Participants, Instruments and Procedure

The EFL teachers taking part in this study were recruited based on nonprobability convenience sampling from Portuguese state school contacts known to the researcher. Out of the available pool of 9th grade English teachers, four gave me their written informed consent to observe their lessons. All the teachers have English teaching experience, ranging from 14 to 22 years of teaching. They hold undergraduate teaching degrees in Languages (either English and Portuguese or English and German) from Portuguese state universities. As it happens, the limited number of teachers in the sample may well be considered one of the study's limitations. Although research results are never so context-dependent that they have no implications for other settings, nor are they so generalisable that may apply to every single setting, the results offered here are, then, suggestive only. Therefore, given the qualitative character of this study, another avenue for future research would be to delve into this topic with larger cohorts of teachers, either across the country or abroad, namely through large-scale questionnaires. It would yield a different source of information to be matched against classroom observations, thus providing more reliable information to answer the research question.

The source of data for this study were classroom observations. Altogether I observed 41 lessons. Apart from the odd exception (e.g., national holidays) observations were evenly distributed. My own degree of participation in the setting moved back and forth between complete observer and observer as participant, starting with the latter and then moving back as much as possible to the former. Students and teachers knew my identity and the teachers the purpose of my stay, but I did not interfere with the natural development of the lesson. To capture the phenomenon under study first-hand, I followed a combination of structured and unstructured classroom observation. The former hinged on an adapted version of the Communication Orientation of Language Teaching (COLT) (Spada & Fröhlich, 1995) observation scheme while the latter hinged on field notes. The combination of both served the purpose of on the one hand coding events as they occurred to have a clearer gist of the teachers' speaking practices over time and on the other hand to make detailed descriptions, which progressed into a running narrative, of important nuances about the nature of speaking assessment events as they unfolded.

The procedure was carefully negotiated because it involved a weekly presence in the classroom for almost a full school year. Complying with the teachers' instructions, I sat at the rear right-hand side of the classrooms to disturb as little as possible.

4. DATA RESULTS AND INTERPRETATION

Most of the lessons observed (92%) were teacher-led, either teacher to learner or teacher to class, which translated in a lion's share of learner's individual work performing the same activity. Only 7% of the lessons were fully learner-led, either learner to learner or learner to class. Yet, it must be stressed that in these occasions, learners were engaged in speaking assessment activities. All of them were asked to do the same activity, being organised once in groups and five times in pairs. Teacher-centred instruction clearly outweighs learner to learner interaction, either in pairs or groups. Indeed, this type of freer activity that allows language acquisition by experimentation was seldom encouraged thus restricting the learners' possible use of the language as well as opportunities to engage in sustained speech outside the strictures of the textbook. Learners rarely had the opportunity to explore the language collaboratively and engage in extensive speaking with their peers on meaningful topics. Such finding parallels that of Koizumi (2022) in Japan. Koizumi highlights the lack of formats for speaking assessment which prompt spoken interaction amongst learners and extemporaneous talk. However, one remark must be made. Some learners, either by anxiety or lack of proficiency, even if given the opportunity refuse to speak. As stated elsewhere (Correia, 2015), the fear of making mistakes in front of peers and being subjected to potential general mockery or laughter, plus the frustration caused by the inability to express oneself clearly in the same way as using one's mother tongue, hinders learners' willingness to communicate and participate in class activities.

As for language itself, a strong emphasis continues to be attributed to grammar. Three out of the four teachers observed spent entire lessons dwelling on grammar, either from the workbook, worksheets, quizzes or all of the above. This pattern is not new in many EFL contexts. For instance, Al Hosni (2014) found it too in Oman. The observation data from her study revealed a marked preference for grammar, whilst speaking was repeatedly not included. More recently (2021), although referring to tertiary education, so did Nguyen, Hung, Duong, & Le (2021), whose study of the Vietnamese EFL context showed an emphasis on linguistic form rather than on communication skills. The issue of grammar instruction raised in the problem statement section is confirmed here. The focus of many classes was on accuracy, grammar rules and the printed word, which translated in being heavily dependent on the textbook and its additional resources. Language functions were coded in all lessons, but their importance in accurately conveying and/or interpreting meaning was never discussed nor was there a connection between them and the co-construction of meaning in intercultural encounters, which would help to combat erroneous stereotypes and prejudiced views of the other. In addition, learners spent most of the time restricted to topics, again usually determined by the textbook, with a narrow range of reference that apply to the classroom domain and/or their first-hand experiences instead of being prompted more regularly to engage with topics that go beyond their immediate environment; and I would venture to say, far more appealing and in line with their interests. Similarly, Adem and Berkessa's (2022) observation of Ethiopian teachers' classroom practices shows that in 15 out of a possible 20 lessons, the nature of the speaking topics stemmed from the textbook. Perhaps, learners' willingness to initiate sustained discourse and interact in the TL would be boosted by taking advantage of their integrative oriented motivation. For now, form, particularly grammar, clearly outweighs meaning. Admittedly, the pervasive testing culture of Portuguese EFL classrooms is mirrored by this teaching behaviour.

Of special interest for the scope of the study was the category "student modality". As it happens, speaking is the least practiced skill, whereas writing tops the rank. Tellingly, not only is speaking the least coded skill in isolation but also the skill that systematically has

a subordinate role when in combination with the rest of the skills. The most popular combination rests with writing-reading, and usually the primary focus is on writing. Only once was speaking given the spotlight in instruction. The data shows that speaking practice lags far behind writing, but also reading and even listening practice. This state of affairs translates in an exceedingly small number of self-initiated turns by the learners and a sparing use of the TL. In line with the rationale offered for the original COLT, "it's important to note that self-allocations, such as calling out an answer, are not considered to be *Discourse initiations*" (Spada & Fröhlich, 1995, p. 87). Although learners are sensitive to turn-taking, they are either left in response mode for most of the time or simply use their first language, mostly for clarification requests. Observation in other settings (Al Hosni, 2014; Adem & Berkessa, 2022) juxtapose this excessive use of the mother tongue. Authentic interaction was scarce, occurring almost exclusively between the teachers and the native-speaker (NS) learners of each class. When they do use English, most learners moved back and forth between ultraminimal (one or to two words) and minimal (three or more words, long phrases and/or one or two main clauses) speech. Sometimes the difference in coding was truly small, minimal speech could easily become ultraminimal (e.g., "Ok teacher" vs. "I need help"). Many learners did not go beyond five word stretches of spoken language. Sustained speech (at least three main clauses) was coded in as little as 11 lessons, of which 5 matched up with speaking assessments whilst another 5 with NS learners. This may be the combined result of low proficiency, language-skill-specific anxiety (negative self-confidence and self-efficacy), and the teacher-centred nature of the class. As alluded to above, the use of non-threatening, freer activities could easily encourage learners to initiate discourse and use the TL more often and for longer stretches, thus positively impacting on both their overall spoken language proficiency and intelligibility. Furthermore, "[…] talk-based instructional practices provide for feedback, self, and peer assessment without the need for formal assessment products" (Shepard, Diaz-Bilello, Penuel, & Marion, 2020).

How exactly, then, do these typical daily lessons influence the general nature of most speaking assessment events taking place in the classroom. First and foremost, teachers seem to be letting themselves be negatively guided by the impact of washback and not by learning. Indeed, most activities carried out reflect summative assessment demands instead of catering to the learners' needs. This teaching-to-the-test effect leads to a narrow focus on test-related content, thus creating a gap between instruction and what should be the intended learning outcomes, i.e., speaking the language proficiently and intelligibly. Teachers tended to prioritize test-specific knowledge, sacrificing a broader and more comprehensive approach that would foster problem-solving abilities through spoken language germane to their lives in and outside the classroom. The disconnection of learning from real-world applications hinders the learners' ability to transfer knowledge and competences to authentic contexts. Meaningful speaking learning tasks were systematically overshadowed by decontextualised test-preparing tasks that hardly, if at all, contributed to the development of the learners' speaking skills and internalisation of new knowledge. Further evidence of the mismatch between traditional forms of assessment and learning, this time from the Czech Republic (Seden & Svaricek, 2018), substantiate the claim for a need to align learning outcomes, classroom activities, and assessments. Seden and Svaricek's results "indicated that the majority of the teachers used assessment for managing behaviour and for certification rather than to improve teaching and learning" (p. 119). Classroom-based testing tasks that concentrate on recalling isolated details or facts are ineffective and promptly forgotten by the learners. When the teachers' primary focus is on preparing for tests, often learners resort to rote memorisation and superficial learning strategies to simply reproduce information. In point of fact, the learners observed seemed to perceive their test-like speaking assessments

as a means to an end rather than a tool for deeper learning. Yet, it must be said that teachers should not carry all the blame. Even though the available legal framework (Ministry of Education and Science, 2016) for assessment in Portugal (an EFL specific framework does not exist) states that formative assessment should be privileged and thus connecting formative assessment to the eligible teaching approaches in (EFL) teaching; as a teacher myself, I am no stranger to the pressure of summative assessment. It results in pressure to achieve success percentages set by school boards, who in turn are under the veiled pressure of the Ministry of Education and national school rankings whose hierarchy is based on the gradings of national exams.

Unsurprisingly, teachers set up oral presentations, role-plays, and description tasks with a grading frame of mind instead of a formative one. Learners ended up restricted by the teacher, textbook or activity to produce pre-set language expected from them in near future test-like speaking assessments. In this manner, the relationship between task and assessment task is one and one alone, that of grading. As a result, learners paid attention solely to their grades, not taking agency over their present state of mastery of the language and showed even less interest in potential feedback sought to guide proficiency improvement. Tailoring speaking tasks to ensure that learners perform well on speaking assessments does not get us beyond surface-level learning whilst limiting the understanding of important speaking skills and intelligibility patterns. In formative contexts such as classrooms, "assessment will mainly go on in the classroom with more formal periodic assessments designed to blend into the learning continuum and recognize what the student has learned and what progress s/he has made" (Hamp-Lyons, 2007, p. 479). Adding to the challenge, many of the tasks conducted often took after the printed word. Learners thought and/or discussed amongst themselves, if it involved pairs, in Portuguese and wrote down their sentences/text in English. This uncharacteristic planning in advance for speaking was followed by plenty of memorisation and rehearsal. As could be expected, learners struggled with their speaking or even came to a halt when they forgot their lines and had to restart their script all over again. Such evidence correlates positively with Koizumi's (2022) findings in Japan where prepared and scripted talk for assessment was the rule. Although resorting to speaking, this behavioural pattern does not match the characteristics of spoken language but the printed word instead. In this vein, learners' speech sounds unnatural, bookish, and too formal.

5. IMPLICATIONS FOR PRACTICE

Out of the data results, the following question can naturally be raised – how are students supposed to provide extensive chunks of spoken language for assessment purposes, or otherwise, if oral practice is not part of normal lessons? The starting point must, then, revolve around effective oral practice as part of normal lessons and from there to a properly functioning assessment system (learning-oriented assessment), which in turn implies a constructive alignment of the curriculum and its intended learning outcomes (intended because we do not necessarily always get the outcomes that we have planned), teaching methods and assessment tasks (for a deeper understanding of this approach see Biggs & Tang, 2011). For present purposes, a constructive alignment in the context of speaking skills with an English as a Lingua Franca (ELF) frame of mind would make the connection between (intended) learning outcomes, teaching methods and assessment in the following manner. First, we want learners to be able to develop their ability to speak proficiently and pronounce the language intelligibly, from a NNS standpoint as future ELF users in situational speaking communities. It is important to share the (intended) learning outcomes with our learners to make them part of the process and thus enhance the learning outcome and for the sake of

transparency. Second, a move from audiolingualistic pedagogy (memorisation and scripted talk) to CLT principles (meaningful communicative tasks that foster intelligibility-like pronunciation) must occur. Classroom activities and procedures must be appropriate to achieving the desired learning outcomes (e.g. collaborative group work activities). Third, assessment is not used with a grading function only but primarily to monitor learners' progress and language acquisition, whilst providing timely feedback by judging the qualities of their performance against what was intended in the learning outcomes. Assessment tasks should match classroom activities and mirror the intended learning outcomes. Unlike many assessment events observed that are separate from classroom activities, there must be a logical thread and consistency to optimise the outcome otherwise we are tricking the learners. Biggs and Tang (2011) go further to assert that when assessment is aligned to what students should be learning, washback can even work positively:

> From the teacher's perspective, summative assessment is at the end of the teaching–learning sequence of events, but from the student's perspective the assessment is at the beginning. However, if the intended outcomes are embedded in the assessment, the teaching activities of the teacher and the learning activities of the student are both aligned towards achieving the same goal. In preparing for the assessments, students will be learning the intended outcomes (p. 198).

By explicitly aligning learning outcomes, classroom activities, and assessments, teachers will promote a supportive learning environment that deepens conceptual understanding and thus improved spoken language proficiency and intelligibility, whilst potentially minimising the negative effects of washback. The vital point when discussing (speaking) assessment is making sure it reflects instruction (frequent opportunities to engage extensively with the language), supports learning, and is meaningful for learners.

Perhaps, some input both for pre- and in-service teacher training which takes into consideration the rationale offered should be adopted for enhanced learner outcomes, both in Portugal and abroad, in view of the similarities between the evaluation procedures used in other countries for assessing speaking in EFL teaching. Despite the cultural differences between countries – Portugal, Czech Republic, Oman, Ethiopia, Japan – they do not seem to significantly impact EFL speaking assessment practices and outcomes. Indeed, teacher training has a role to play in filling the gap of formative speaking assessment. If, on the one hand, most, if not all, Portuguese universities offering teaching degrees for prospective teachers include plenty of curricular units related to assessment, this is not the case with continuous professional development (CPD) courses for in-service teachers. This is paramount if we think of the Portuguese teachers' profile. Novice teachers are few and far between in the classroom. School staff is made up of older teachers who did not have up-to-date assessment pre-service training. The teachers' profiles provided by the Portuguese Directorate-General for Statistics of Education and Science (2021) shows that more than half in-service teachers are over 50 years old. It is, then, the responsibility of local policy makers to act upon the gap identified here, by providing hands-on teacher training on classroom-based formative speaking assessment. In addition, thinking about the findings of this study, some refinement of the current guidelines offered by the Ministry of Education and Science for EFL teaching and learning (2018), encouraging speaking and pronunciation (focus on intelligibility) practice would be advisable. In a similar vein to policy makers, materials writers must also be held accountable for their potential role in a positive knock-on effect of Portuguese EFL teachers' practices on classroom-based speaking assessments. Drawing on the tendency of over-reliance on the textbook by the teachers observed, if it ever incorporates the development of oral proficiency and intelligibility as one of its main goals it may duly contribute to steadily closing the gap alluded to above.

Thinking of local policy makers, one interesting initiative, referred to by Koizumi (2022), that could eventually be mimicked by the Portuguese government, is being developed in Japan by the National Institute for Educational Policy Research (research institute integrated in the Japanese Ministry of Education, Culture, Sports, Science and Technology). Teachers are being provided with "with handbooks that include good assessment practice samples along with explanations; encouraging performance-based speaking; and conducting learning-oriented and criterion-referenced teacher, self-, and peer assessment" (Koizumi, 2022, p. 145). Yet another commendable and freely actionable initiative developed by teachers for teachers with the support of the Education and Training department of the Victoria State Government can be found at https://teal.global2.vic.edu.au/. The Tools to Enhance Assessment Literacy (TEAL) project is an online resource for teachers of primary and secondary level pupils who are learning English as an Additional Language (EAL) in Australia. This online open resource underscores the assessment for learning theoretical principles, providing an assessment tool bank, guiding principles and advice, and exemplars of assessment tasks with formative feedback, to name but a few. Although stemming from a rather distinct context, TEAL may prove useful to gather practical insight on how to implement classroom-based speaking assessments with a learning-oriented frame of mind.

Another possible way to tip the scales in favour of learning and thus of formative speaking assessment values would be process speaking. This suggestion is not entirely new, it draws from the widely recognized pedagogical method process writing. "Process writing refers to a set of beliefs and strategies that enable teachers to work with student writers while they are writing, rather than waiting until a piece of writing is finished and then marking or critiquing it" (Hamp-Lyons, 2007, p. 477). Process writing has been advocated for its pedagogical advantages over product writing (Onozawa, 2010; Rao & Durga, 2018), including its benefits for reduced anxiety (Bayat, 2014). Naturally different from writing, the underlying rationale for speaking is the same. Rather than simply focusing on grading learners' product, i.e., their speaking assessment performance, teachers would closely support learners' language learning trajectory, i.e., their process. Instead of putting the focus solely on the final spoken product, the process speaking approach would place equal emphasis on the various stages prior to the final assessment task, such as pre-speaking preparation (brainstorming ideas, organizing thoughts), practicing on a regular basis (collaboratively with their peers), receiving ongoing constructive feedback (ideally both from the teachers and peers, guiding learners towards improving their proficiency and intelligibility), and refining speaking performance based on the feedback received. In this fashion, learners are pivotal in the learning process, being encouraged to discuss topics relevant to their lives and interests and consequently more engaged, motivated, and confident to use the language in its spoken form. Throughout the process, alternative teaching methods that go beyond traditional approaches whilst taking advantage of technology can also be implemented, e.g. flipped classroom. Along these lines meaning is paramount, whilst form is deemphasised, grammar is a means not an end, the textbook plays a subsidiary role, being used only and if it contributes to genuine learning, the link between teaching, learning and assessment is easily traced (constructive alignment). The process speaking approach offered above is envisioned to follow a gradual increase in complexity throughout the school year, where learners initially engage in simpler speaking tasks and progressively advance to more challenging ones gradually scaffolding their spoken language proficiency and intelligibility, whatever their starting point may be. Yet, if truth be told, going against the dominance of the testing culture in ELT is a tough row to hoe. It would require a joint effort of the teaching community, educational policy makers, and publishing materials writers.

6. CONCLUDING THOUGHTS

There is considerable evidence throughout the literature (Swain, 2000; Oliver, 2009; Leong & Ahmadi, 2017; Correia, 2021) to demonstrate the significance of spoken production. It is Swain who claims that "output may sometimes be, from the learner's perspective, a 'trial run' reflecting their hypothesis of how to say their intent" (2005, p. 476) whilst El Majidi, Graaff and Janssen say it "can assist language learning through prompting learners to notice their language gaps, testing out hypotheses and reflecting consciously on forms" (2021, p. 3) and Loewen and Sato argue that "it facilitates production practice and, thus, the development of fluency and automaticity" (2018, p. 292). Despite its importance, in Portugal, as in other EFL contexts (see Thornbury, 2005 or Goh, 2017, for instance), there seems to be a mismatch between the perceived value of speaking and the ways it is put into practice.

Underlying speaking practice is a written-based orthodoxy reminiscent of a long writing tradition of teaching and learning a foreign/second language which continues to fall into the trap of considering spoken writing as speech. Unwittingly or not, teachers display a tendency to resort excessively to test-like assessment preparation instead of catering to the learners' real needs, hypothetically compromising their spoken language proficiency and intelligibility improvement. Consequently, extensive speaking occurs mainly as the spin-off of assessment events. Often, scripted dialogues are used, which differ significantly from ordinary spoken language – functions and structures typically occur with unnatural frequency; utterances tend to be very short and overly well-formed; backchannel responses, discourse markers and colloquial expressions are seldom used; and a shared knowledge of context is not assumed. Everyday speech rarely generates continuous correct complete sentences, clearly articulated words, and a lack of stance by the interlocutors. Complications arise from the preference of accuracy over fluency, form over meaning, and grammar rules over language in use. It seems, then, that speaking assessment tasks are not aligned as they should with the ongoing classroom activities used by the teachers. Assessment results are used haphazardly to inform instruction and fail to identify areas for improvement, which makes learners' language learning trajectories hardly traceable. Portuguese EFL teachers appear to be at odds with designing suitable assessment procedures for monitoring students' progress. There is a narrow view of assessment as synonymous with testing, and thus the grading function, which largely contributes to the dominance of summative assessment over formative assessment.

This chapter is an attempt to highlight the significance of the connection between classroom tasks, (speaking) assessment, and language development, adding to the existing body of knowledge whose interest has mainly dealt with the general nature of assessment practices, speaking constructs, rater effects, and factors that affect speaking performance (Fan & Yan, 2020). Despite the wealth of research on speaking assessment, it addresses marginalized topics and fills in some of the gaps recently identified by Fan and Yan (2020). First and foremost, the study contributes to increase the limited research committed to classroom-based speaking assessment; second, it taps on two underrepresented topics: learning-oriented speaking assessment and speaking assessment for young speakers (9^{th} grade students); third, it explores the *utilization* inference "which concerns the communication and use of scores" (Fan & Yan, 2020, p. 6), i.e. how do teachers provide feedback, if any, and how does assessment impact teaching practices (washback). Hopefully, a wide readership from the teaching EFL community, but also educational policy makers and publishing materials writers will find food for thought in this chapter.

REFERENCES

Adem, H., & Berkessa, M. (2022). A case study of EFL teachers' practice of teaching speaking skills vis-à-vis the principles of Communicative Language Teaching (CLT). *Cogent Education, 9*(1). doi:https://doi.org/10.1080/2331186X.2022.2087458

Al Hosni, S. (2014). Speaking Difficulties Encountered by Young EFL Learners. *International Journal on Studies in English Language and Literature, 2*(6), 22-30. Retrieved from https://www.arcjournals.org/international-journal-on-studies-in-english-language-and-literature/volume-2-issue-6

Alderson, C. (2007). The Challenge of (Diagnostic) Testing: Do we Know what we are Measuring? In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe, *Language Testing Reconsidered* (pp. 21-39). Ottawa: University of Ottawa Press.

Bachman, L. (2002). Alternative Interpretations of Alternative Assessments: Some Validity Issues in Educational Performance Assessments. *Educational Measurement: Issues and Practice, 21*(3), 5-18. doi: https://doi.org/10.1111/j.1745-3992.2002.tb00095.x

Banta, T., & Palomba, C. (1999). *Assessment Essentials: Planning, Implementing,and Improving Assessment in Higher Education.* San Francisco: Jossey-Bass Publishers.

Bayat, N. (2014). The Effect of the Process Writing Approach on Writing Success and Anxiety. *Educational Sciences: Theory & Practice, 14*(3), 1133-1141. doi:10.12738/estp.2014.3.1720

Biber, D., Conrad, S., & Leech, G. (2002). *Longman Student Grammar of Spoken and Written English.* Harlow, Essex: Pearson education.

Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University* (4th ed.). OUP/Society for Research into Higher Education.

Blood, I. (2011). Diagnostic Second Language Assessment in the Classroom. *TESOL & Applied Linguistics, 11.1*, 57-58. Retrieved from https://academiccommons.columbia.edu/doi/10.7916/D8DJ5SFH/download

Bonner, S., Torres Rivera, C., & Chen, P. (2018). Standards and assessment: coherence from the teacher's perspective. *Educational Assessment and Accountability,, 30*(1), 71-92. doi:https://doi.org/10.1007/s11092-017-9272-2

Boud, D., & Falchikov, N. (2006). Aligning Assessment with Long-Term Learning. *Assessment & Evaluation in Higher Education, 31*(4), 399-413. doi:https://doi.org/10.1080/02602930600679050

Brown, H. D. (2004). *Language Assessment Principles and Practices.* New York: Longman.

Carless, D. (2009). Learning-Oriented Assessment: Principles, Practice and a Project. In L. Meyer, H. Davidson, R. Anderson, P. Fletcher, & M. Rees, *Tertiary Assessment & Higher Education Student Outcomes: Policy, Practice & Research* (pp. 79-90). Wellington, New Zealand: Ako Aotearoa.

Correia, R. C. (2015, August). Classroom Anxiety and Young Language Learners: Broadening the Scope of the Research Agenda. *Humanising Language Teaching, 17*(4). Retrieved from http://www.hltmag.co.uk

Correia, R. C. (2021). *Can We Speak?: Approaching Oral Proficiency in the EFL Classroom* (Unpublished doctoral dissertation). NOVA FCSH - Faculty of Social Sciences and Humanities, Department of Languages, Cultures and Modern Literatures , Lisbon. Retrieved from http://hdl.handle.net/10362/125772

Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing, 21*(3), 305-334. doi:10.1191/0265532204lt286oa

Directorate-General for Statistics of Education and Science. (2021, September). *Perfil do Docente 2019/2020* [Teacher Profile 2019/2020]. Retrieved from https://www.dgeec.medu.pt/api/ficheiros/64f6ea3f885805b7cfbf0738

El Majidi, A., de Graaff, R., & Janssen, D. (2021). Debate as a pedagogical tool for developing speaking skills in second language education. *Language Teaching Research*, 1-22. doi:10.1177/13621688211050619

Fan, J., & Yan, X. (2020). Assessing Speaking Proficiency: A Narrative Review of Speaking Assessment Research Within the Argument-Based Validation Framework. *Frontiers in Psychology, 11*, 1-14. doi:https://doi.org/10.3389/fpsyg.2020.00330

Goh, C. (2017). Research into practice: Scaffolding learning processes to improve speaking. *Language Teaching, 50*(2), 247-260. doi:10.1017/S0261444816000483

Hamp-Lyons, L. (2007). The Impact of Testing Practices on Teaching: Ideologies and alternatives. In J. Cummins, & C. Davison, *International Handbook of English Language Teaching* (pp. 487-504). Boston, MA: Springer.

Hamp-Lyons, L., Hood, S., & MacLennan, C. (2001). Promoting quality teaching in the tertiary context. *Higher Education Review, 34*(1), 60-76.

Kazemi, N., & Tavassoli, K. (2020). The Comparative Effect of Dynamic vs. Diagnostic Assessment on EFL Learners' Speaking ability. *Research in English Language Pedagogy, 8*(2), 223-241. doi:https://doi.org/10.30486/relp.2019.1878561.1155

Koizumi, R. (2022). L2 Speaking Assessment in Secondary School Classrooms in Japan. *Language Assessment Quarterly, 19*(2), 142-161. doi:https://doi.org/10.1080/15434303.2021.2023542

Leong, L.-M., & Ahmadi, S. M. (2017). An Analysis of Factors Influencing Learners' English Speaking Skill. *International Journal of Research in English Education, 2*(1), 34-41. doi:10.18869/acadpub.ijree.2.1.34

Liu, N.-F., & Carless, D. (2006). Peer Feedback: the Learning Element of Peer Assessment. *Teaching in Higher Education, 11*(3), 279-290. doi:https://doi.org/10.1080/13562510600680582

Loewen, S., & Sato, M. (2018). Interaction and instructed second language acquisition. *Language Teaching, 51*(3), 285–329. doi:10.1017/S0261444818000125

Luoma, S. (2004). *Assessing Speaking.* Cambridge: Cambridge University Press.

Mihaljević Djigunović, J. (2019). Affect and assessment in teaching L2 to young learners. In D. Prošić-Santovac, & S. Rixon, *Integrating assessment into early language learning and teaching practice* (pp. 19-33). Bristol: Multilingual Matters. doi:10.21832/9781788924825-005

Ministry of Education and Science. (2016, April 5). Ordinance n.º 1-F/2016. Lisbon, Portugal. Retrieved from https://diariodarepublica.pt/dr/detalhe/despacho-normativo/1-f-2016-74059570

Ministry of Education and Science. (2018, July). *Aprendizagens Essenciais - Inglês 9º Ano* [Essential Learning - English 9th Grade). Lisbon.

Nguyen, L. T., Hung, B. P., Duong, U. T., & Le, T. T. (2021). Teachers' and learners' beliefs about pronunciation instruction in Tertiary English as a foreign language education. *Frontiers in Psychology, 12*. doi: https://doi.org/10.3389/fpsyg.2021.739842

Nikolov, M. (2017). Students' and teachers' feedback on diagnostic tests for young EFL learners: Implications for classrooms. In M. d. García Mayo, *Learning foreign languages in primary school: Research insights* (pp. 249–266). Bristol: Multilingual Matters. doi:https://doi.org/10.21832/9781783098118-014

Nikolov, M., & Timpe-Laughlin, V. (2021). Assessing young learners' foreign language abilities. *Language Teaching, 54*(1), 1-37. doi: https://doi.org/10.1017/S0261444820000294

Oliver, R. (2009). How Young is too Young? Investigating Negotiation of Meaning and Feedback in Children Aged Five to Seven Years. In A. Mackey, & C. Polio, *Multiple Perspectives on Interaction: Second Language Research in honor of Susan M. Gass* (pp. 135-156). London: Routledge.

Onozawa, C. (2010). A Study of the Process Writing Approach: A Suggestion for an Eclectic Writing Approach. *Research Note, 10*, 153-163. Retrieved from https://core.ac.uk/works/42002358

Purpura, J. (2004). *Assessing Grammar*. Cambridge: Cambridge UP.

Rao, C., & Durga, S. (2018). Developing Students' Writing Skills in English - A Process Approach. *Journal for Research Scholars and Professionals of English Language Teaching, 2*(6), 1-5. Retrieved from https://www.researchgate.net/publication/325489625_Developing_Students%27_Writing_Skills_in_English-A_Process_Approach

Rodgers, T., & Richards, J. (2001). *Approaches and Methods in Language Teaching.* Cambridge: Cambridge University Press.

Seden, K., & Svaricek, R. (2018). Teacher subjectivity regarding assessment: Exploring English as a foreign language teachers' conceptions of assessment theories that influence student learning. *CEPS Journal, 8*(3), 119-139. doi:https://doi.org/10.26529/cepsj.500

Shepard, L., Diaz-Bilello, E., Penuel, W., & Marion, S. (2020). *Classroom assessment principles to support teaching and learning.* University of Colorado Boulder: Center for Assessment, Design, Research and Evaluation. Retrieved from https://www.colorado.edu/cadre/2020/02/11/classroom-assessment-principles-support-teaching-and-learning

Shohamy, E. (1997). Second Language Assessment. In R. Tucker, & D. Corson, *Encyclopedia of Language and Education* (Vol. 4: Second Language Education, pp. 141-149). Dordrecht: Kluwer Academic Publishers.

Spada, N., & Fröhlich, M. (1995). *Communicative Orientation of Language Teaching Observation Scheme: Coding Conventions and Applications.* Sydney: National Centre for English Language Teaching and Research.

Swain, M. (2000). The Output Hypothesis and Beyond: Mediating Acquisition through Collaborative Dialogue. In J. Lantolf, *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 471-483). Routledge. doi:https://doi.org/10.4324/9781410612700

Thornbury, S. (2005). *How to teach speaking.* Harlow: Longman.

Tsagari, D. (2016). Assessment orientations of primary state school EFL teachers in two Mediterranean countries. *CEPS Journal, 6*(1), 9-30. doi:https://doi.org/10.26529/cepsj.102

ACKNOWLEDGEMENTS

I wish to express my gratitude to the director of CETAPS, Professor Carlos Ceia, and the administrative officer, Mariana Gonçalves, for their support.
Funded by national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the project UIDB/04097/2020, https://doi.org/10.54499/UIDB/04097/2020.

AUTHOR INFORMATION

Full name: Rúben Constantino Correia
Institutional affiliation: CETAPS NOVA FCSH – Centre for English, Translation and Anglo-Portuguese Studies, Portugal
Institutional address: Campus de Campolide CAN - Colégio Almada Negreiros, 1099-085 Lisbon
Short biographical sketch: Rúben Constantino Correia is a Portuguese and English Lower Secondary teacher with almost twenty years of experience. As of September 2022, he started teaching English for Specific Purposes at the Algarve University. At the same time, he is also a CETAPS full member (TEALS team researcher) and an individual member of EALTA – European Association for Language Testing and Assessment. His main research interest is Oral Proficiency and Speaking Skills, supported by other research areas which intersect and contribute to the core that forms speaking – Teacher Education; Interculturality; Assessment; The Common European Framework of Reference; and Affective Variables. He holds a teaching degree from the Algarve University, a Master's in English and North American Studies and a PhD in Foreign Language Methodology from the Faculty of Social and Human Sciences of Nova University of Lisbon.