

ASSESSMENT FROM WITHIN – UNDERSTANDING THE KNOCK-ON EFFECT OF TEACHERS’ PRACTICES ON CLASSROOM SPEAKING ASSESSMENTS

Rúben Constantino Correia

Centre for English, Translation and Anglo-Portuguese Studies – Languages, curriculum, and teacher education research strand (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 seem 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. As a result of such procedures, functions and structures regularly arise with atypical frequency, utterances are exceedingly short and exaggeratedly well-formed; backchannel responses, discourse markers and colloquial expressions are seldom used, and a shared knowledge of context is not assumed. To turn the tables on this state of affairs teachers need to assess with a learning-oriented frame of mind, i.e., to be the link between instruction and what is learned and to promote effective student learning. 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.

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. Naturally, 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.

1.1. 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/assessor 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 faces a major challenge – the reluctance of Portuguese state schoolteachers to address it. Most students studying English at the lower levels (5th up to 9th 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.

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 transposable 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 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).

2. Speaking's inherent character

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 with 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. 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.

3. Key concepts

Assessment has become a popular but “sometimes misunderstood term in current educational practice” (Brown H. D., 2004, p. 4) and for this reason a distinction between the terms assessment and testing, which are repeatedly used interchangeably, must be made.

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. 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. When a teacher gives a test, he/she 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. On the other hand, assessment is an ongoing process-oriented approach that takes many different forms. One of these forms are tests. Thus, testing is a subset of assessment and should be seen as one of the many methods available for assessing students’ verbal performance. 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 by some experts like Shohamy (1997) and Bachman (2002). 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 find out what students are learning in the classroom and how well they are learning it” (Angelo & Cross, 1993, p. 4). Assessments serve as tools to 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 which in turn assist students to improve their speaking proficiency. 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. Assessment goes beyond the question how much the students have learned; instead, it asks how they learned and what can be done to improve their learning.

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 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).

Learning-oriented approaches to speaking should not be concerned only with measuring ability, but also with 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.

5. Methodology

The study followed a qualitative approach. It involved observing four different 9th grade classes (once a week) for almost a full school year in a combination of an observation scheme supplemented by descriptive linguistics field notes. The observation scheme was adapted from Spada and Fröhlich (1995)

original COLT – Part A and Part B, therefore named COLT PT – Part A and COLT PT – Part B. Besides the scheme, I always took blank sheets of paper to each lesson, allowing abundant space to make various entries about the events taking place inside the classroom.

6. Data analysis and discussion

Most of the lessons observed (87%) were teacher-led, either teacher to learner or teacher to class, which also translated in learner's individual work performing the same activity. Only 9% 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, allowing for few opportunities to engage in sustained speech and thus restricting the learners' possible use of the language. However, I reinforce the term possible because some learners, either by anxiety or lack of proficiency, even if given the opportunity refuse to speak. As for language itself, a strong emphasis continues to be attributed to grammar. Language functions were coded in all lessons but their importance in accurately conveying and/or interpreting meaning was never discussed. In addition, learners spent most of the time restricted to topics, usually determined by the textbook, which 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 nearest environment (e.g., international events). It would seem that form outweighs meaning and within it grammar is the front runner.

Of special interest for the scope of the study was the category student modality. As it happens, speaking is the least practiced skill. 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. Only once was speaking given the spotlight in instruction. This state of affairs translates in an exceedingly small number of self-initiated turns by the learners and a sparing use of the target language. 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. When they do use English, learners move back and forth between ultraminimal (one or two words) and minimal (three or more words, long phrases and/or one or two main clauses) speech. Sometimes the difference in coding is truly small, minimal speech could easily become ultraminimal (e.g., "Yes teacher" vs. "I don't know"). Many learners do 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. I would say 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.

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. Yet, it must be said that teachers should not carry all the blame. As a teacher myself, I am no stranger to the pressure of summative assessment, which results in pressure to achieve success percentages projected by school boards. Consequently, set up oral presentations, role-plays, and description tasks with a grading frame of mind instead of a formative one. Adding to the challenge, these often take after the printed word. Learners think and/or discuss amongst themselves, if it involves pairs, in Portuguese and write down their sentences/text in English. This uncharacteristic planning in advance for speaking is followed by plenty of memorisation and rehearsal. As could be expected, learners struggle with their speaking or even come to a halt when they forgot their lines and have to restart their script all over again. 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.

7. Concluding thoughts

There is considerable evidence throughout the literature (Swain, 2000) (Oliver, 2009) (Correia, 2021) to demonstrate the significance of spoken production, yet in Portugal extensive speaking occurs mainly as the spin-off of assessment events. More often than not, 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. Speaking seems to fall through the cracks of the Portuguese EFL

classroom. Complications arise from the preference of accuracy over fluency, form over meaning, and grammar rules over language in use. Naturally, the following question can 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 the connection between learning aims (improved proficiency and intelligibility), teaching methods (moving from audiolingualistic pedagogy to CLT principles), and assessment (monitoring of learners' progress and language acquisition whilst providing timely feedback). Perhaps, some input both for pre- and in-service teacher training which takes into consideration the rationale offered should be adopted for improved learner outcomes.

References

- 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 (Eds.), *Language Testing Reconsidered* (pp. 21-39). Ottawa: University of Ottawa Press.
- Angelo, T., & Cross, P. (1993). *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco: Jossey-Bass Publishers.
- 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.
- Blood, I. (2011). Diagnostic Second Language Assessment in the Classroom. *TESOL & Applied Linguistics*, 11(1), 57-58. Retrieved from academiccommons.columbia.edu
- 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, & et al. (Eds.), *Tertiary Assessment & Higher Education Student Outcomes: Policy, Practice & Research* (pp. 79-90). Wellington, New Zealand: Ako Aotearoa.
- 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. doi: <http://hdl.handle.net/10362/125772>
- Luoma, S. (2004). *Assessing Speaking*. Cambridge: CUP.
- 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 (Eds.), *Multiple Perspectives on Interaction: Second Language Research in honor of Susan M. Gass* (pp. 135-156). London: Routledge.
- Purpura, J. (2004). *Assessing Grammar*. Cambridge: Cambridge UP.
- Rodgers, T., & Richards, J. (2001). *Approaches and Methods in Language Teaching*. Cambridge: CUP.
- Shohamy, E. (1997). Second Language Assessment. In R. Tucker, & D. Corson (Eds.), *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 (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: OUP.