CATERING TO ASSESSMENT NEEDS OF STUDENTS OF ENGLISH - CALL TO THE RESCUE?

by **Anita Buczek-Zawiła**
Pedagogical University of Cracow
Podchorążych 2, 30-084 Kraków, Poland
anita.buczek-zawila@up.krakow.pl

Abstract

The study focuses on the process of assessing (a micro-skill) goal attainment in EFL pronunciation course via measures which can foster different students' attitudes and self-perceptions. Standard (pen-and-paper) tests offer immediate evidence of success but they put heavy demands on students' cognitive, performance and stress-controlling skills. CALL-related techniques can be used as supplementary ones, even if technically assessing different sub-skills, *Kahoot* or *Moodle* quizzes can complement and re-orientate the assessment as well as the learning processes.

To investigate the impact of the diverse assessment measures a small-scale research was conducted among Year 1 students of the English Department at the Pedagogical University in Kraków. They are participants in a 90-hours-a-year pronunciation course, where one of the components involves mastering transcribing skills. The specific element of the course evaluated by standard and CALL-related measures in the study were the phonetic variants of the *-es* and *-ed* endings in English. Through analysis of test scores, coupled with the ideas obtained via semi-structured interviews, the study hoped to verify the claim that matters such as student comfort, instant individual feedback and personal safety are most efficiently handled by the *Moodle* quizzes. Apart from providing well-balanced scores, they offer the least-threatening, stress-free environments for learning and assessment, thus developing students' self-monitoring their progress.

Keywords: assessment; test anxiety; pronunciation instruction; pen-and-paper tests; computer-based tests

1. Introduction

Any learning process requires assessing goal attainment. Whether the focus is on performance or mastery goals, the measures applied to assess attainment of these can be diverse and can foster different students' attitudes and self-perceptions. For many students, the test situation on its own creates considerable anxiety which can badly affect their performance (Komorowska, 2002; Underhill, 1987). Therefore, it is deemed desirable to attempt working out assessment schemes and conditions that would both determine the relative success at learning a specific content or

skill and do so in a reliable and guided manner. Additionally, instructors would want to show the learners how to evaluate their progress in learning, to be able to diagnose the strong and the weak points independently of the teacher. Both teacher assessment and carefully introduced elements of student's self-assessment contribute to students' making more progress and taking more responsibility for their learning (Czetwertyńska, 2015; Komorowska, 2005; 2002; Underhill, 1987). It appears to be a necessity to reconsider the habitually used assessment practices in order to implement diversification and flexibility inherent in adaptive teaching to make assessment more comprehensive, more inclusive and more students'-needs-relevant.

English Studies pronunciation courses, apart from regular training in English oral skills, characteristically insist on mastering transcribing skills. The belief behind this requirement is that those who can record in a visual form certain contrasts or variants practiced in class will also be able to pronounce them – and vice versa (Werfel, 2017). It is seen as a feeding-breeding relationship. Assessing students' success in this skill typically incorporates pen-and-paper tests of transcription, where the test-takers' task is to apply the IPA characters appropriately. While this mode offers the teacher immediate evidence of success – or lack thereof – it puts heavy demands on students' cognitive, performance and stress-controlling skills. Students focus on the outcome and what can be done to ensure the positive final score, not on improved understanding, self-monitoring or generating solutions. Therefore, some more comprehensive and varied measures need to be implemented to provide assistance in the assessment process and to complement it.

Even if technically assessing different sub-skills, certain CALL-related techniques, such as game-based student response systems (*Kahoot!* for this study) (Licorish et al., 2017; 2018) or *Moodle* quizzes (Neill, 2018; Aperliński et al., 2013) can complement and re-orientate the assessment as well as the learning processes. These supplementary measures enable important factors to surface from the background. Student comfort, instant individual feedback and personal safety reveal themselves to be all-important and largely conducive to the ultimate success, at the same time contributing significantly to developing self-monitoring skills.

Therefore, the paper tries to examine the more adaptive and flexible assessment protocols with a view to cater for student diversity in a relevant manner. The typical, ever-present "test culture" largely ignores the need to adapt assessment activities and tasks to students' characteristics and personal needs within the affective domain, at the same time not taking notice of students' emotional well-being. The sole focus on measuring progress or achievement prevails. The study aims to investigate whether expanding task types and techniques can constitute a move towards a form of inclusive assessment consistent with demands of adaptive

teaching (Coll et al., 2000). Through analysis of test scores, both traditional and CALL-related, enriched by information obtained via a series of semi-structured interviews, the author hopes to determine whether incorporating the various assessment forms and CALL elements can indeed foster students' performance skills. Put differently, the question arises: Does CALL come to the rescue when we try to cater for students' assessment needs?

In order to reveal the preferable and comprehensive assessment forms of the study participants, a number of research questions have been formulated (section 4). Based on these, fed by ideas developed in the subject literature, a working hypothesis has been proposed to the effect that catering for students' needs when it comes to assessing the achievement of transcription skills ought to involve incorporation of carefully planned CALL-based activities.

2. The background: Transcription vs. speaking

It has been observed countless times that pronunciation errors cause the majority of communication breakdowns, both when we talk about communicating with native and non-native speakers, with, ultimately, pronunciation constituting possibly the greatest single barrier to successful exchanges. Speech is, after all, an activity which is carried on in numerous events. Lack of knowledge of English sounds, both in terms of recognition and production, can thus impede communication. Too much accentedness or distorted speech may give rise to misunderstanding, miscommunication and frustration, resulting in psychological nervousness in speakers (Lu, 2002, p. 37). Learning to understand and use phonetic symbols leads to effective remediation of these difficulties.

It is commonly believed that early integration of the basic phonetic symbols into the English language classroom, particularly so at the college or university level, is essential to the students' formation of accurate English pronunciation and intonation – the foundations of verbal language, helping to minimize interference from their mother tongue. As Lu (2002, p. 38) observes, phonetic symbols are not difficult to learn, particularly with teachers who know how to effectively use them to improve and further develop learners' pronouncing skills. Needless to say, using them requires a lot of practice before a strong command of the symbols is possible. It may be because learning to transcribe necessarily requires adults to shift the way they analyse words, ultimately to make them able to think beyond the print.

Phonetic transcription training leads to increased levels of explicit phonemic awareness (Werfel, 2017, p. 285). Phonemic awareness, the ability to analyse the sound structure of words at the level of individual speech sounds, has emerged over the past several decades as an important, but not the only, linguistic predictor of children's reading and spelling skills.

Kindergarten phonemic awareness skills are positively correlated to word decoding and reading comprehension across grade levels, from elementary school to high school (Werfel, 2017, p. 282). However, the further step, namely, thinking beyond print, requires explicit phonemic awareness skills. A higher level skill than phonemic awareness, explicit phonemic awareness, is a literate individual's ability to analyse the sounds of words separate from print. After breaking the code of letter-sound correspondence, individuals are influenced heavily by print in their analysis of words, even when instructed to analyse the sounds of words. Individuals with higher levels of explicit phonemic awareness can better manipulate the sound structure of words while ignoring the influence of print and, thus, likely would report that, for example, "witch" and "rich" contain the same number of sounds and that they rhyme.

Phonetic transcription training is the active ingredient of foreign language education that results in higher levels of explicit phonemic awareness. It is a useful learning technique for two different kinds of people: native speakers and EFL learners alike (Lecumberri and Maidment, 2000). Generally, getting familiar with it reinforces the idea that spoken and written representations of language are separate things. It makes people, especially native speakers of a language, more aware of what they actually say, rather then what they think they say. In the learning context, "transcribing texts helps to make one aware of the target one should be aiming for and of the pronunciation (or range of pronunciations) one can expect from native speakers" (Lecumberri and Maidment, 2000, p. 1). It therefore makes pedagogical sense to underscore the fundamental significance of sound discrimination and identification in the process of developing good foreign language pronunciation (Eckman et al., 2009), where transcribing skills offer great assistance. ICT affords various tools to be used for such training, including specific protocols implemented in the (Polish university-) popular learning managing system Moodle (Aperliński et al., 2013). *Conditio sine qua non* of good pronunciation is the ability to distinguish the foreign language phonemic inventory's sounds. A student who is unable to hear the difference between the sounds is unlikely to be able to pronounce them correctly. It is not possible to expect and demand proper sound reproduction of what we cannot extract auditorily (Eckman et al., 2009). However, it is equally clear that although perceiving the contrasts in the sound inventory of a foreign language is the basis for developing good pronunciation, it does not guarantee that. From the start, the main objective of learning pronunciation is to develop phonematic hearing, which is the ability to accurately recognise a foreign language's sounds. Again, at the university level, phonetic transcription sensitizes students to the existing contrasts.

Once the teacher made sure that the students had mastered the skill of recognising a given sound or sound groups, they could go to the next phase of training, that is, to practice correct

articulation, frequently trying to overcome L1 habits in the process. All imitative acts that are the quintessence of such production training are an act of will and the effect of the decision-making process. They include acceptance of a certain pattern, recognition of its characteristics as a specific implementation instruction, and consequently execution of this implementation through the control of the articulatory apparatus. If such actions become fixed, they may turn into desirable articulation habits (Isaacs, 2014). This is where the universally perceived primacy of perception in relation to speech production comes from (Escudero, 2007). Thus, the help offered in the form of phonetic transcription is seen as essential in fostering increased phonemic awareness and - in consequence - good speech habits.

It has been observed through years of classroom practice at the Pedagogical University in Kraków Year 1 English Pronunciation Classes that students' performance on tests on phonetic transcription, particularly those targeting the realisation of appropriate variants of two inflectional endings of English, the "-es" and the "-ed", is less than satisfactory. In personal communication, when discussing the individual results, students report on their difficulties with having to cope with multiple areas of expertise at the same time: having to recall and recognise the pronunciation of a whole item, to apply an appropriate variant of the ending and to record all that in the form of phonetic script, all in the allocated time limit and in classroom conditions, surrounded by their peers and closely monitored by the teacher. For years, the results obtained in these tests were significantly lower than those in other tests of transcription. The specific kind of test this study focuses on are achievement tests, which take a sample of a language element, here – the specific variants of two inflectional endings in English (the "-es" and the "-ed"), which has been covered in the course of study in order to determine how well the learners have mastered the specific element(s) (Underhill, 1987, p. 13). They also perform a subsidiary diagnostic function, allowing the teacher to know which aspect of the particular content was not successfully learnt.

It has been judged desirable to enrich the assessment process with techniques that can support students in their learning and create conditions that would eliminate the debilitating effect of classroom test-related anxiety. Two such solutions were introduced: *Kahoot!* and e-Quiz on the department e-learning *Moodle* platform. These two tools were selected in order to minimise the emotional inference and to emphasise the centrality of learning, within the broadly perceived interrelatedness between instruction, assessment and learning, as those that address the contextual, the instructional, the interactional, the elicitation and the affective dimensions straightforwardly.

3. Literature review

Ideally, for any meaningful learning to occur, assessment-elicited evidence of student learning should be gathered, as a result of which instruction is modified in response to feedback. Only recording current student achievement seems inadequate and incompatible with the ideas of assessment for learning (Cheng & Fox, 2017) and adaptive teaching (Coll et al., 2000). In terms of actual uses and practices employed by instructors, still mostly traditional forms of assessment, tests or quizzes, continue to be used, and even with more innovative procedures or modes, the focus is on traditional language elements or skills (Shohamy et al., 2008). Considering the relationship between the process of learning and the need to evaluate its relative success, the following regularity transpires: "Learning a language, like learning anything else, is essentially an individual achievement… But typically this private process takes place in the public context of the classroom, the individual is one of the group" (Castillo, in Turula, 2013, p. 255).

Assessment tools and methods are multiple and varied (see Cheng et al., 2004 and Cheng and Fox, 2017, for a review). Studies frequently concentrate on external, standardised, high-stakes assessment formulas, though the diversity of teacher practices is acknowledged, stressing the relationship between assessment and instruction (Cheng et al., 2004) as well as the test 'fitness for purpose' concept (Galaczi, 2010). Research has also acknowledged the challenges inherent in meaningful and efficient assessment as faced by teachers (Wach, 2013; İnceçay & İnceçay, 2010; Cummins, 2009).

With the constant spreading of computer-based testing (CBT), important features of the mode are brought to the foreground, such as greater flexibility in test item design, access to a larger repository of items, and faster turnarounds for score receipt (Backes & Cowan, 2018). Such a test mode also allows for testing accommodations, understood as changes to testing procedures (presentation of test materials, students' responses to test items, scheduling, test setting, etc.) that have little or no impact on the construct being tested and supporting the performance of students so that their academic content knowledge and skills are demonstrated to the greatest extent (Educational Testing Service, 2009). Computer-based tests can provide constructive diagnostic information to complement the language learning process, yet it is argued they should be used more selectively in other contexts, for example in high-stakes tests, where the factors of design, validity, technological familiarity and practicality are an issue (Dooey, 2008). The advantages of online testing, though outside the context of language learning, have been summarised in Candrlic et al. (2014), who say that an essential part of every LMS is a subsystem that enables online testing. Students solve tasks using a computer rather than writing their responses on paper. Standard LMS systems (such as Moodle, Sakai, Dokeos and

Blackboard) provide various options that make the process of creating tests or quizzes very flexible. As a result, instructors are able to construct tests that feature different types of questions, randomly generate tests from question banks, allow students to solve the test several times, etc. Their results show that online tests can replace traditional paper-based tests for students' knowledge assessment, but special attention should be paid to the test composition. They also report that, based on students' comments received during the testing activities, preference to use the keyboard rather than a pencil transpires more and more, so they would rather take online tests than paper-based tests, however, no specific reason for that preference was stated (Candrlic et al., 2014).

The testing mode has thus been viewed to play a major role. Shohamy et al. (2008, p. 19) underline that quizzes and tests were criticised on several grounds, for instance, for the fact that they produce anxiety and diminish performance, placing learners in (apparently) high-stakes situations, where one-time accuracy determines their future learning. Against the background of concerns for how paper-and-pencil testing fits into modern pedagogy, including contemporary language teaching, research is focused on, among others, the (non-existent) difference in scores between in-class testing and online testing in a communicative and proficiency-oriented program (VanPatten et.al., 2015); on the postulates to view computer-based and traditional speaking assessment as *complementary* and not competing perspectives, where technology is seen not as a replacement for standard methods, but as a new additional possibility (Galaczi, 2010); or on the relationship between testing mode (taking a test on computer versus on paper) and cognitive load (Prisacari & Danielson, 2017), where the results support the claim that online testing can be implemented in educational settings without imposing additional cognitive load on students. The opposite, namely increased cognitive load, yet resulting in better recall and positive impact on working memory performance during paper-based tests rather than computer-based tests, was reported in the study by Carpenter & Alloway (2018), who also point out that research on CALT aims to demonstrate the equivalence of scores between CBT and standard paper-and-pencil tests, so that the two modes are comparable. Stowell and Bennet (2010) argue that administering regular course exams in an online format would reduce test anxiety experienced at the time of the exam and improve exam scores. Online quizzes offer considerably more flexibility than more traditional paper-based tests as students can participate in them 24/7. It is by definition easier to arrange early or late online quizzes (Neill, 2018).

There is wide research available into the relationship between testing, motivation, anxiety and stress-related factors (Cheng et al., 2014; Coll et al., 2000; Stowell & Bennet, 2010; Vavla & Gokaj, 2013; Wu & Lee, 2017). The studies reveal significant relationships among attitudes,

motivation, text anxiety and test performance, all within the context of self-efficacy and pressure on self-esteem. Motivation and anxiety exhibit a mutually influential relationship and alongside other personal factors will influence students' overall performance (Cheng et al. 2014). This anxiety typically surfaces during certain cognitive performances for test-takers, for instance, when they compare themselves with others in the group, are concerned about the consequences of failing the test, experience low self-confidence or are generally excessively worried about being (publically) assessed (Wu and Lee, 2017). Additionally, difficulties may be experienced with adapting to the level and type of exercises, enforced under psychological pressure (Vavla & Gokaj, 2013). In this context, Yielder et al. (2017) discuss not only the impact of test format on strategies for test taking and learning in general, but also on students' well-being. In the latter aspects students described the effects tests had on their emotions, detailed as fear of failing, anxiety about negative marking, a feeling of lack of control, a worry of not learning enough. Such emotions of uncertainty and broadly-conceived-of stress had a detrimental and demoralising effect on their learning.

It is only natural that teachers often find students in a class showing much diversity in their needs and interests. They differ a lot in their motivation, prior knowledge and skills, learning styles, multiple intelligences, interests and backgrounds, to name but a few aspects. It is in this context that alternatives to traditional in-class, paper-based assessment are sought, and the solutions need to encompass innovative, possibly technology-related, but first of all student-centred protocols (Chapelle & Voss, 2017). Learning structure that is fun and interesting is more appealing to students than when it is done for the sole purpose of passing the test and meeting requirements.

Instructional games are gaining acceptance in the classroom as the e-learning merits of student engagement and immediate feedback are recognised. Within higher education, the use of such tools is often limited due to doubts regarding the scholarly merits of such activities, whereas when applied with caution they allow educators to tailor their instruction based on student scores and understanding on quizzes (Plump and LaRosa, 2017). *Kahoot!* is a popular eLearning tool that can easily be used to add vitality, student engagement, and meta-cognitive supports to higher education classrooms with a limited instructor or student training required (Plump and La Rosa, 2017). The format is believed to offer support for students to interactively answer quizzes in classrooms as part of a formative assessment regime (Licorish et al., 2018; 2017). As such it is a welcome innovation introduced during classes.

Aperliński et al. (2013) discuss students' opinion voiced in relation to Moodle or traditional paper transcription exercises. Among the pluses of paper exercises, points such as

quicker and more effective learning, more comfortable and natural writing, no need to carry a laptop and no websites to distract are mentioned. These are countered with the minuses: no (instant) feedback and difficulty with error correction as rewrite is needed. Online Moodle activities are praised for being fast and easy to use, for low effort demands, instant, accurate and clear feedback and for relative ease to correct errors (no rewrite needed). The problematic areas are the overall difficulty level being too low because of low effort requirements, and the fact that proper transcription form is quickly forgotten.

The present study was grounded in a course whose primary objective was improving the students' pronunciation of English. As the course progressed, the students developed their personalised attitudes towards English pronunciation as well as the abilities to self-rate their achieved levels, therefore attaching different degrees of importance to having good pronunciation. Consequently, motivational aspects and willingness to attain good pronunciation are not stable, but vary with experience and, often, learner characteristics (Waniek-Klimczak et al., 2013). Cognitive factors that are typically found relevant for more accurate pronunciation and phonological processing involve attention control allowing one to switch between tasks and mental sets, phonological short-term memory used to encode phonological units and their sequential order for further processing, and inhibition control – the mechanism that allows bilinguals to speak one of their languages while blocking the interference from the language currently not in use, so that learners can better avoid interference from their L1 phonological categories resulting in more target-like / less accented L2 perception and production (Mora and Darcy, 2017). Pronunciation learning regularly involves positive and negative emotions, which in turn may have great impact on their L2 learning and assessment. Errors may lead to adult learners' hypersensitivity and self-consciousness as well as lower self-image. When they feel that they are the focus of attention and under constant scrutiny of their classmates and the teacher, they may not perform their real competence, which points to a significant role of the affective domain in learning L2 pronunciation (Jedynak, 2013). Gabryś-Barker (2012) further emphasises that students may demonstrate preference to perform certain actions in a specific context, depending on perceived intrinsic pleasantness of the stimulus in specific favourable instructional conditions. Tasks which offer expectation of success promote high need of achievement and low fear of failure. All this is conducive to enhancing learners' self-esteem. Achievement goals fall into two categories: performance goals and mastery goals. A mastery goal orientation emphasises learning, understanding, improving, mastering new skills, and taking on challenges, they require varying evaluation methods, and making evaluation private. In contrast, a performance-goal orientation (also called an ego-involved orientation) emphasises comparison of students'

abilities, where student evaluation is public, performance is attributed to individual ability, and students who outperform others are rewarded (Cauley and McMillan, 2010). Consequently, the achievement of both types needs to be assessed. All this must be somehow integrated with the demand for concrete marks that frequently are required by the specific educational reality to certify to the course completion, though Shohamy et al. (2008, p. 19) observe that assigning grades was consistently among the least important reasons for assessing ventured by teachers. Such seems to be the case with the group targeted in the study.

4. Methodology

4.1. The aims of the study

This study aims at answering the following research questions:

1. Is there a significant difference in the scores obtained on traditional pen-and-paper tests of transcription as compared to the scores achieved in CALL based measures (Kahoot! and *Moodle* quiz)?
2. What are the students' opinions on regular pen-and-paper tests used to evaluate their transcription skills?
3. How welcome is the presence of CALL-related solutions as part of the assessment process?
4. Which CALL solutions are most conducive to both effective and student-friendly assessment, as judged by students' test scores and personal opinions?

To achieve the goals, the current study was conceptualized as a small scale mixed method study into the assessment preferences in the group of 1st year students in the English Department of the Pedagogical University of Kraków (PUK). The aim was to determine the set of possible formulas and contexts for meaningful, comprehensive and at the same time student-friendly assessment of elements of specific course content, compatible with the principle of catering for diversity.

4.2. Participants and the context

For the students in the English Studies Department at PUK a course in English Pronunciation, termed in the curriculum as (Practical) English Phonetics, is mandatory. It comprises 90 hours of class instruction and obligatorily includes training in IPA transcription with the aim to visually record English words and utterances.

The study focused on 4 groups of 1st Year students (N = 72) in the English Department at the Pedagogical University of Kraków, who as part of the degree participate in the 90-hours-course called Phonetics.

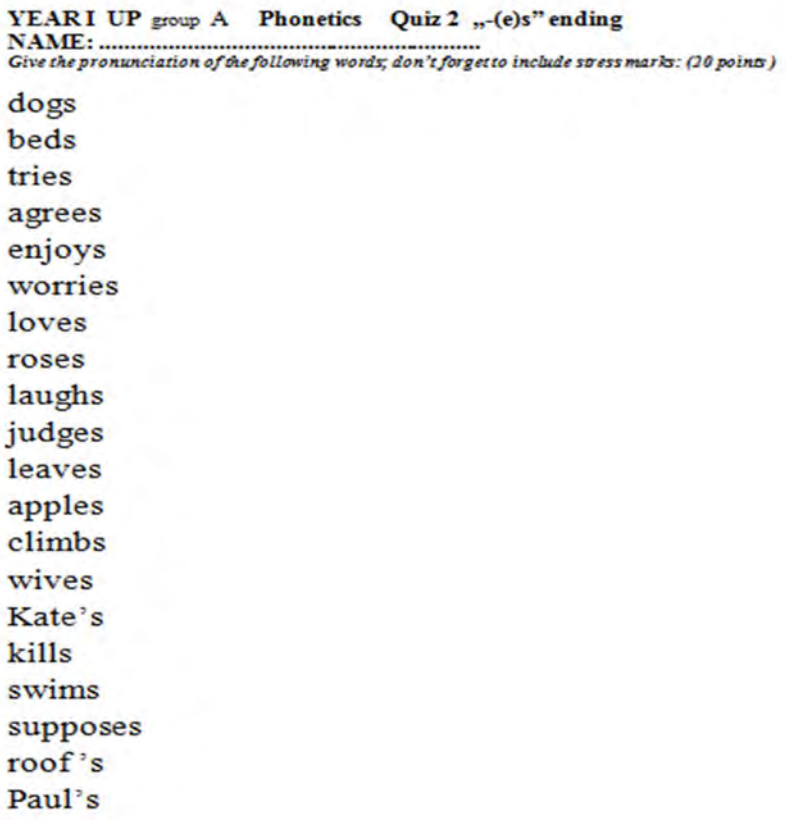
Table 1. The sample

Number of Year 1 groups	4
N (participants overall)	90
Number of Year 1 groups with CALL-related testing protocols	2
N (CALL-groups participants)	36
Overall number of test samples	216
Overall number of interviews	10

The central objective of the course was actual training in English pronunciation, where students develop their expertise through a variety of instructional modes in order to attain high-accuracy near-native-like pronunciation goals, disassociating themselves from their L1 background and abandoning, to the extent it is possible, those habits of pronunciation that contribute significantly to L1 (Polish, Ukrainian, Russian)-accented speech. An important part of the curriculum is learning to use (read and produce) the IPA phonetic transcription of English words and utterances. The belief behind this requirement is that those who can first recognise and subsequently record in a visual form the contrasting variants practiced in class, will also be able to pronounce them as a result of developing and thus raising their explicit phonemic awareness (Werfel, 2017).

One of the L1 habits to be eradicated in the course of instruction is the massive tendency towards Final Obstruent Devoicing, carried over from L1 onto English. Students work on developing an understanding of the notion of voicing in general, on its functions in terms of contrasting meaning and on the problems inherent in dropping L1 interference in the form of producing voiceless consonants at the end of words. Curriculum writers judged the two English inflectional endings: the multifunctional "-es" and the (mostly) regular praeterite "-ed" marker to be used in this context, first as phonological awareness raisers, and second as items combining several aspects of the language phonology, morphology and (in a broad sense) grammar. Such item practice goes hand in hand with the proposed ban on compartmentalization of language skills practice. Thus, during the course, students spend several classes practicing the variant-to-environment applications, working on handouts specifically designed to bring out the multitude

of functions and contexts and produce desirable contrasts. Their achievement is then measured through transcription-based tests. These are followed by a face-to-face oral pronunciation test, where the two inflectional endings in their respective variants feature prominently, both in individual words and in longer utterances.



YEAR I UP group A **Phonetics Quiz 2 „-(e)s" ending**
NAME: .....
Give the pronunciation of the following words; don't forget to include stress marks: (10 points)

dogs
beds
tries
agrees
enjoys
worries
loves
roses
laughs
judges
leaves
apples
climbs
wives
Kate's
kills
swims
supposes
roof's
Paul's

Figure 1. Pen-and-paper quiz

4.3. Design and procedure

For many years the pen-and-paper tests described above were considered relatively straightforward and easy for students in terms of low cognitive load required to perform well. And yet, in personal communication, when discussing the obtained scores, this view was challenged. As a result, in two student groups of students (N = 36) additional assessment techniques were implemented. The "-es"-ending assessment was supplemented with a *Kahoot!* quiz run in class, followed by a regular class based written test of transcription, to be finally complemented by a *Moodle* quiz. For the "-ed"- ending, only the Moodle quiz was used, to the total abandonment of the pen-paper-test. Sample quiz formats for the online formulas are shown below. In the remaining groups no CALT measures were introduced.

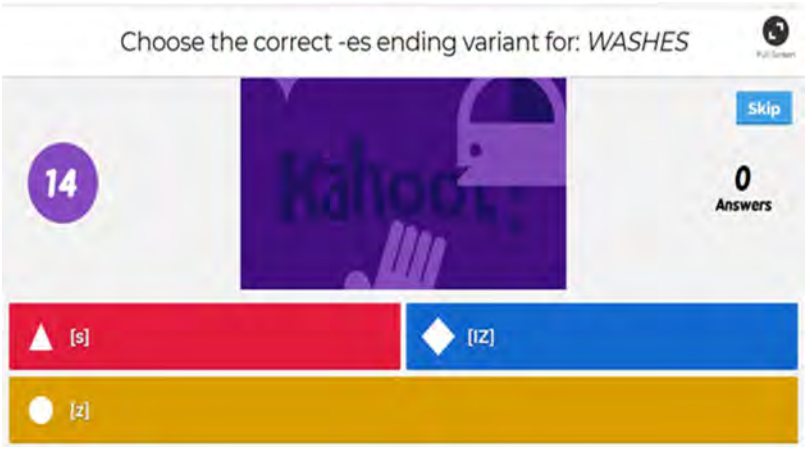


Figure 2. *Kahoot!* – the question format

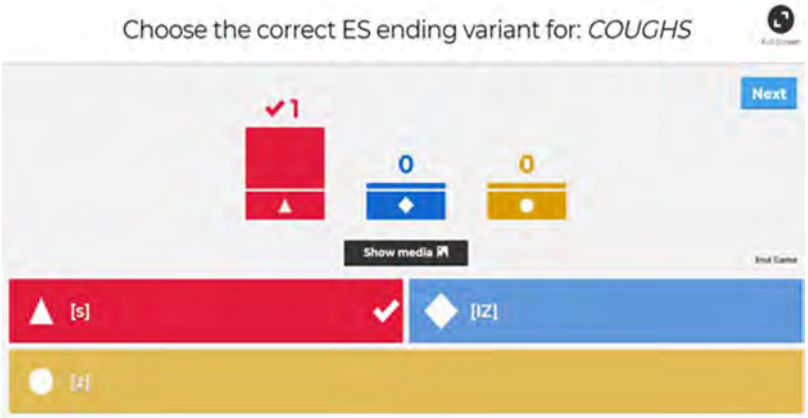


Figure 3. *Kahoot!* – the main screen feedback

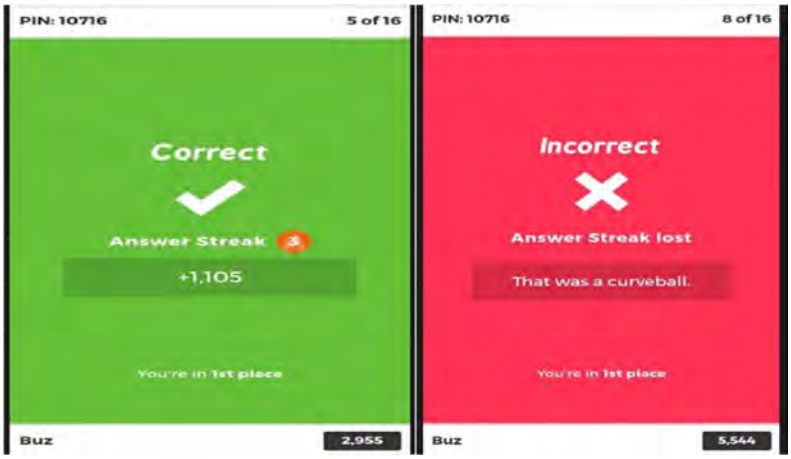


Figure 4. *Kahoot!* - individual feedback on a mobile device

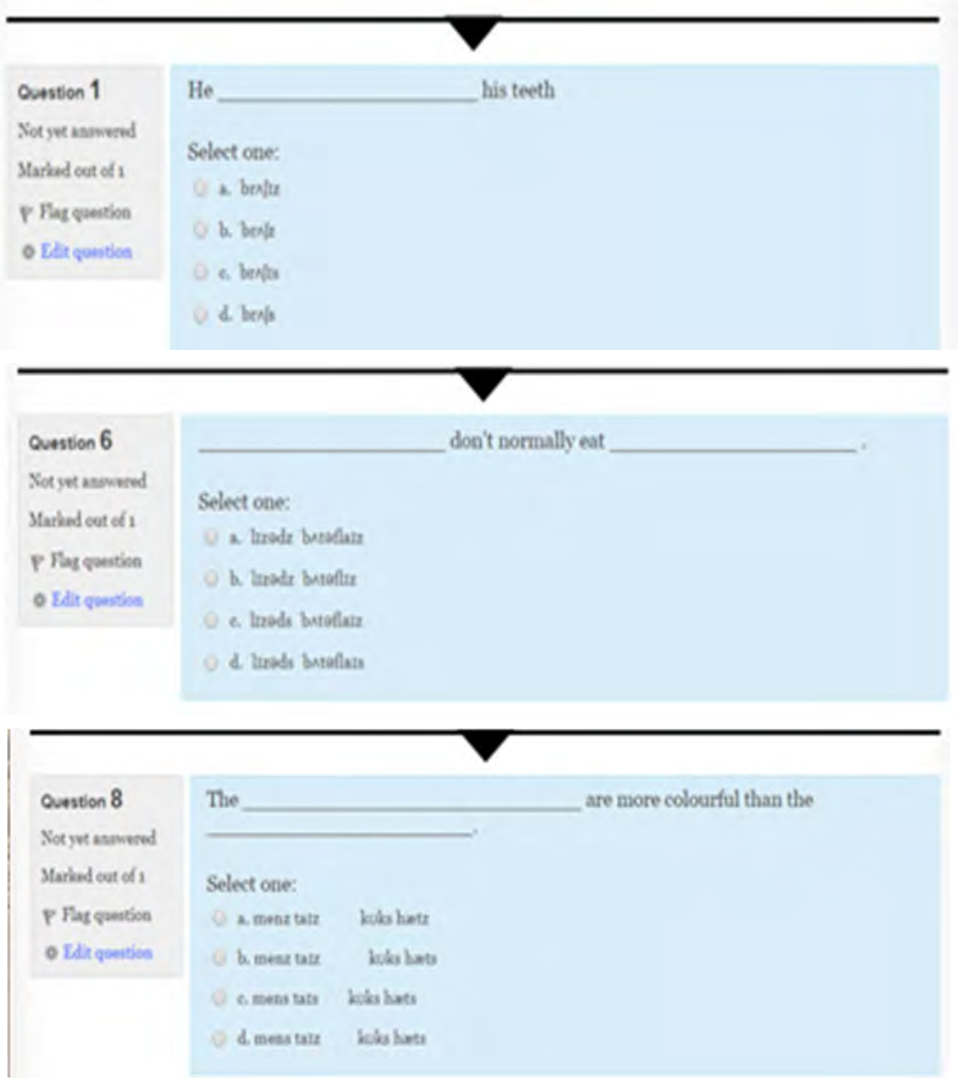


Figure 5. *Moodle* - the format: from simple one-item-questions to progressively more complex multi-items ones

Figure 6. *Moodle* - the quiz summary for an individual student

4.4. Data collection tools and procedures

The instruments of data collection included written test results (the quantitative part) and the semi-structured interviews (the qualitative part). The data were gathered between December 2018 and February 2019. Pattern-coding strategy (Miles et al. 2014) was employed to analyse the interviews data qualitatively.

4.5. The finding: test results

The following section outlines the scores obtained in all the assessment formulas across the student-participants. The order of presentation is as follows: first, the scores on the *Kahoot!* quiz on the "-es" ending variant is shown, then the pen-and-paper "-es" ending written test on transcription results are shown. These are followed by the scores obtained on the *Moodle* quizzes on the "-es" and the "-ed" endings in the two groups (N = 36) (Table 2). For comparison, scores of randomly selected students (from the general population taught) obtained in the groups with no added CALL-related techniques are presented for both the endings (Table 3).

Table 2. Scores in written and CALL tests

N=36	**Kahoot "-es"**		**"-es" pen-and-paper**		**"-es" Moodle quiz**		**"-ed" Moodle quiz**	
	Score (16)	%	Score (20)	%	Score (10)	%	Score (10)	%
Correct (average)	12.222	**76**	17.406	**87**	9.333	**93**	9.305	**93**
SD	2.21		2.05		0.98		0.75	
Number of failed tests	10		4		1		2	
Chi-square validation	**13.006**							
P-value (p<.05)	**.004624**							

Table 3. Pen-and-paper test scores in no-CALT supported groups

N=36	**"-es" pen-and-paper**		**"-ed" pen-and-paper**	
	Score (20)	%	Score (20)	%
Correct (average)	16.5	**82**	16.62	**83**
SD	2.81		2.46	
Number of failed tests	9		8	

The results show, on the whole, that course participants in their majority have successfully acquired relevant transcription skills when it comes to mastering the correct application of the appropriate variant of the two inflectional endings. The passing grade for the tests was generally rather high, 75% for the *Kahoot!* quiz (12 points out of 16) and 80% for the remaining test

formulas. The percentage of failed attempts is acceptable in all cases, the overall discriminability allowed for appropriate proportions of all marks.

One could argue that the tests are not really comparable, since they require slightly different skills of the participants. The *Kahoot!* and the *Moodle* quizzes follow the multiple-choice format and focus on recognizing the correct set of IPA symbols and the appropriate ending variant. The pen-and-paper test requires students to actually produce the transcription themselves, which is seen as a special difficulty in comparison to the other too, as it calls for a heavier cognitive and memory load on the part of the students (Mora and Darcy, 2017). On the other hand, the *Moodle* quizzes get progressively more complex, they start with a single word-item to be recognised, ending up with as many as four distinct forms within one question space. This supports the assumption that the facilitation in the form of a multiple-choice recognition task is compensated for by the increased relative complexity of the task itself. All three formats are strictly time-limited, with more leniency in this respect observed with classroom-run written tests. The results were always discussed with each individual, who can study the test, compare it to the expected value and analyse the feedback given. Notably, the types of feedback were different for each of the techniques, with *Kahoot!* quiz offering the least personalized feedback, though the instant one, just like in the case of a *Moodle* quiz. The feedback, albeit individualised, was significantly delayed in the case of pen-and-paper transcription test.

In order to arrive at a deeper understanding of the students' performance in the assessment procedures, as well as to discover their opinions on the forms most conducive to their learning a series of semi-structured interviews was carried out. The next section discussed the ideas emerging from them.

4.6. The findings: interviews

From the total of 36 participants from the two groups who did both traditional pen-and-paper tests as well as *Kahoot!* and *Moodle* quizzes, a selection of 10 students was sampled to forward their opinions and ideas in the form of a semi-structured interview, carried out during the author's office hours. The students were not selected randomly, it was the deliberate intention of the author to have a spectrum of opinions coming from high, average and low achievers. Therefore purposive and non-probability sampling was judged desirable. The interviews were run after all forms of the "-es" - ending testing were done, but before the "-ed" ending was tested. Before the end of the semester, the interviewees as well as any other participants could venture their opinions during the final individual oral task test, carried out in the face-to-face situation with the teacher. The carefully-orchestrated selection of students was the consequence of a popular view that

different types of students work best in different instructional contexts and need different kinds of feedback. Low achieving students require highly specific immediate and explicit feedback, while high-achievers benefit more from delayed feedback (Cauley and McMillan, 2010). Also, in a group there will always be students who are reticent to demonstrate their knowledge publically, for fear of being labelled attention-seeking by their peers (Licorish et al., 2018). One such student, a high-achieving 'tall poppy', who scored maximum on all tests, demonstrated precisely that sort of attitude in class, and was therefore asked to give the interview. Apart from them, two more high-achievers, four average students and three low-achieving participants were asked to contribute their ideas.

The interviews were structured so as to disclose students' opinions and encourage them to forward their own ideas for assessment that supports their learning. Some of the questions were fed by responses offered by the author, and basing on views found in subject literature (Jedynak, 2013; Gabryś-Barker, 2012; Mora and Darcy, 2017; Plump and LaRosa, 2017). Others were more open, so as to avoid intentional answering and the statistical "halo-effect", where essentially the respondents answer in the way they intuitively feel the researcher expects them to answer. They were grouped into several major areas, following the pattern coding strategy (Mills et al., 2014):

a. identifying the procedural differences between the three testing formulas;
b. rating the significance of factors such as task type, task time limit, teacher involvement, student engagement, cognitive load, attention focus, the testing context, the elicitation techniques;
c. the problems with each of the techniques, with the division into inherent and individual;
d. the affective and the interactional dimension of assessment and their relatedness to learning.

The data extrapolated from the interviews are discussed below, where a selection of the students' opinions are tabulated for ease of reference. These present but a fraction of the opinions expressed, they can, however, be treated as representative for the views held by this particular group of respondents.

Table 4. Semi-structured interviews responses

The feature	"-es" pen-and-paper	Kahoot! "-es"	"-es" Moodle quiz
What is required of students	Producing transcription of individual items with "-es" ending added	Recognizing the required variant of the ending, appropriately to the phonetic context	Recognizing the correct transcription of individual words incorporating the ending variant appropriate in the given phonetic context
General context	Classroom, public	Classroom, public	Familiar, private
Task formula	Clear, unchanging	Clear, unchanging	Clear, unchanging
Time limit	Set but negotiable with the teacher if need be; teacher leniency observed	Strictly defined by the program	Defined by the platform, non-negotiable
Aids allowed	Printed table with IPA characters	None	Printed table with IPA characters
Teacher intervention	Possible when needed, but generally minimal	None	None
Difficulty level	Consistent, stable, non-varying	Consistent, stable, non-varying	Progressively more challenging and demanding, more items to be handled at once
Attention focus	Can vary	Essential	Essential but can vary
Feedback formula	Delayed, individual, informative	Immediate, simple (right or wrong)	Immediate, individual, can be insufficient
Potential problems 1: technicalities	"nothing to write with"; "pencils not allowed", "an easier group"	Slow internet connection, slow mobile device, BYOD principle	Slow internet connection; platform breakdowns;
Potential problems 2: individual	Insufficient command of the symbols	Slow reaction time, following the screen action	Insufficient command of the symbols; attention fluctuation
Anonymity	Some	None	Full
Competitive spirit	None	A lot	None
Interaction pattern	With the material, not the group	With the class	With the material, very individual
Relatedness to further learning	Regular class assessment practice; feeds self-study protocols; enhances skill development	Not much; a welcome break in the class flow; possibly a revision tool; affects class dynamics	Additional procedure to classroom based ones; feeds self-study protocols; enhances skill development

Test anxiety	High	Some, related to competition, activity not perceived as a test situation	Intermediate

When asked specifically to evaluate the three techniques as forms of assessment, the respondents presented rather balanced and mature views, certifying to a lot of reflective critical thinking about the matters discussed. Table 5 summarizes their opinions (underlining mine).

Table 5. The tests evaluation

"-es" pen-and-paper	● heavy demand on memory and recall skills ● writing things down takes time ● many details to take into account: the symbols, stress marks, variants etc. ● assesses many aspects at the same time ● external control of all circumstances ● *high level of test-induced stress and anxiety* ● undoubtedly one's individual work; *reflecting the level of course element attainment* ● *detailed though delayed feedback in personal communication with the teacher* ● "if you do it slowly, everybody in class can see that" ● teacher monitoring constantly and able to react to potential problems
Kahoot! "-es"	● "a waking up activity"; ● introduces variety and diversity, ● welcome break in classroom routine*;* changed dynamics ● *high level of stimulation and excitement*; ● external control of all circumstances; ● "enjoyed by those who can't live without some form of competition" ● improves students' engagement and satisfaction ● *A revision activity - yes; form of assessment - no* ● *A "lifebuoy" function in assessment, tipping the scales in favour of mildly doubtful performance* (the tall poppy's suggestion)
"-es" Moodle quiz; **"-ed" Moodle quiz**	● focuses on recognizing transcription and the variants - *a much lower cognitive load, still heavy memory load* ● done individually within the time limit and time window defined by the teacher ● technical possibility for cheating by sharing answers; *therefore they cannot be the only form of assessment, could be assigned relative less worth to eliminate the incentive for cheating* ● quick and quiet feedback, regular, though not as detailed as for the class written test ● not every class material lends itself nicely to this form of assessment ● my own machine, my own space, my own comfort => *personal safety* ● nobody physically monitors the activity; no hovering to see where I am in the task ● improve student engagement and satisfaction ● *much lower test-related stress or general anxiety* ● *not so very easy as I had originally thought* ● *the clear task instructions are all important*

5. Discussion and implications

At a superficial level, especially when comparing the task formulas and demands, what could reasonably be expected is the overwhelming student preference for lighter, somewhat entertainment-oriented, involving, competition-related forms of assessment. A closer inspection, however, reveals some very thoughtful considerations are surfacing.

It is only natural that undergraduate students welcome the use of the *Kahoot!* game. It bases on students bringing and using their own mobile devices (the BYOD practice) and, frequently, their own high-speed internet. With the growing user expertise and new potential of mobile devices also for working (indirectly) on one's pronunciation or general speaking skills, effective use of them in classroom practice is preferable to trying in vain to ban them from the class completely (Niewiadomski, 2010). One result that is perhaps the most striking in relation to *Kahoot!* quizzing technique is the reported lack of anonymity. Numerous sources (Licorish et al., 2017; 2018; Plump and LaRosa, 2017) underline this aspect as one of the benefits of using GSRS tasks in the classroom, fostering deep and enriched participation. Because of the game's high-competitiveness, anonymity was not enforced in class, actually students revealed their nicknames voluntarily and wanted to be known for their achievement, ready to risk demoting when answering incorrectly. They were not, however, judged on the spot for their responses, correct and incorrect.

Judging from the opinions expressed in the interviews, the participants felt strongly that *Kahoot!* could be used for revision as a useful tool for course content, before actual assessment. That is in accordance with the findings of Licorish et al.(2017; 2018), who pointed out to *Kahoots!*'s potential for repeating course content in a novel way, facilitating remembering key concepts and allowing a deeper understanding of the relevant issues, later reinforced and enriched in class discussions and more in-depth thinking. Equally important is the specific feature of *Kahoot!*, namely that as its organizing ground rule it assumes the central role of fun and entertainment, which on the whole positively impacts learners' motivation, enriches learning experience and enhances active participation. This contributes to the learners' emotional and psychological well-being. As such it can also positively affect the attitude learners develop towards the specific language aspect studied, which in the long term must contribute to higher attainment and increased target-seeking (cf. Waniek-Klimczak et al., 2013). For those in class who are less competition-oriented, it clearly changes the patterns of class dynamics, a desirable occurrence anyway. On the whole, it is likely to contribute to creating an environment that is most conducive to deep learning and thus providing students with the much needed tools to adopt

diverse learning strategies in their study as assessment (Coll et al., 2000). The most interesting and most innovative suggestion, though, is the idea to use *Kahoot!* quiz results in favour of those students who represent borderline cases assessment-wise. Coming from a person who does not need that sort of crutch, being a top scorer on all the course assessment procedures, this is seen as a reflection of deep critical thinking and a sign of mature attitude to managing one's own learning. The idea was openly presented to the class and accepted by popular vote as an element of classroom procedures.

Pen-and-paper tests as a form of mainly summative assessment have been a regular fixture in many classes. Students are accustomed to them, see their role as evaluative measures and realize the cognitive and memory load associated with them (Carpenter & Alloway, 2018; Shohamy et al., 2008). However, if the only feedback students receive is a final grade (e.g., for a unit of instruction, midterms, finals, or external tests), they cannot see how their efforts improve skills, which may lower expectations for success in the future (Ellis, 2008). Furthermore, the evaluative comments and judgments of ability that are prevalent in comparisons can be debilitating for students. The test-induced anxiety can be somewhat manipulated by using fair tests, with unambiguous, familiar item types and tasks (Turula, 2010). When students know these are coming, they understand their role in realistic assessment of their performance and goal-attainment as well as receive simple and straightforward instructions, the tension and anxiety inherent in the test situation can be lowered. The interview responses revealed that students understand the common preconceptions and recognize symptoms of test-related anxiety. They still, however, underscore the lack of psychological and physical comfort when taking them. It is related to the fact that what happens with in-class assessment is that a private process is accommodated in and largely influenced by the public domain in which it takes place (Turula, 2013). Because of this public dimension, it requires some possible accommodations (Educational Testing Service, 2009) in a dialogue with the learner, who especially in this public context may experience a number of negatively affecting emotions, generally subsumed under the term "language anxiety".

The students then may and do experience any combination of the following (Turula, 2013): lack of confidence in oneself as a learner, uneasiness, frustration, self-doubt, apprehension and tension which are specifically related to language learning situations. It is regularly context-sensitive, influenced by learners' perception of this environment and arises in response to certain language learning situations: assimilation of knowledge and skills, evaluative procedures, having to perform publicly in the tasks, whether of a test-like nature or not. When all these occur in the sphere of pronunciation training, the effects can be amalgamated, since learner's individual

factors play an important role in acquiring the target language phonology. These comprise not only learner anxiety but can keep auditory information in working memory and focus attention on the relevant cues (Mora and Darcy, 2017). All these are arguments in favour of adapting classroom testing procedures and yet to continue some testing protocols while looking for student-friendly complementary measures.

When the factors described above – fun and engagement elements, competitiveness, test related stress – are cumulatively taken together, the significant role of the affective domain in L2 pronunciation instruction becomes evident. It therefore makes pedagogical sense to search for supplementary assessment techniques that would compensate for the variable anxiety and cognitive control influence on students performance. One such additional measure was the introduction of GSRS *Kahoot!* quiz. Another could be relocation of a portion of the course assessment process into the virtual environment of a *Moodle* platform.

The course followed a module-like organization, therefore any number and arrangement of online quizzes can be created, at least partly in place of a mid-semester series of written tests. Students were permitted to see responses and answers immediately after submission, but only general feedback and overall mark was viewable after this window is closed. This is consistent with the general advantage of computer-based-tests, namely that when administered in relevant situations they provide nearly instant constructive diagnostic information to complement the language learning process (Dooey, 2008). Quiz extensions can be offered only to students under exceptional circumstances, without the necessity to arrange additional times and venues for class test. All this makes computer-based tests a valuable element of dynamic and adaptive classroom assessment (Backes & Cowan, 2018; Neill, 2018). Although quiz security remains a concern for the teacher (Neill, 2018), the incorporation of them seems inevitable. These measures are also welcome by students themselves, who classified this form of assessment as one of possible testing techniques, reflecting on the unsuitability of every chunk of the course to be e-evaluated. They understand the pedagogical concerns of the teacher when it comes to possible cheating and suggest different score value to be allocated to the online quiz within the whole course assessment. At the same time they appreciated the lowered cognitive load associated with this test formula (Aperliński et al., 2013; Prisacari & Danielson, 2017; Vavla & Gokaj, 2013), still noticing the progressive increase in question items difficulty. Test mode effects are evident in the overall better scores achieved, though the difference is not overwhelmingly in favour of out-of-class assessment (VanPatten et.al., 2015). However, what they seem to view as the biggest advantage and a single factor most contributing to their success on the test is the positive influence of personal safety and familiar, anxiety-reducing environment of doing the test.

Emotional security, interaction with the material only, generally lowered anxiety all appeared to positively impact test-takers' performance (Yielder et al., 2017; Wu and Lee, 2017). Personal, safe environment, without the teacher hovering over, without looking to the peers around is seen as conducive to raised self-esteem and better achievement as a result, to the exclusion of factors such as phonetic aptitude and high motivation (Jedynak, 2013; Turula, 2013; Gabryś-Barker, 2012).

The participants demonstrated great sensitivity to personal and contextual circumstances in the strife to obtain successful outcomes in their learning process (Gabryś-Barker, 2012). In all three measures taken to assess goal-attainment, slightly different task types were employed. In the CALL-related ones, the tasks themselves may have had a more facilitative power and create a positive attitude in students. However, since positivity works best when it precedes the given learning situation and a learner is willing to actively engage in it, it is not surprising that the experience of online *Moodle* quiz was uniformly viewed as conducive to greater than usual as well as more pleasant achievement. Since the potentially threatening elements were eliminated, the situation-specific anxiety was sufficiently low to stimulate desired performance rather than promote defensive or withdrawal strategies. After the positive appraisal of first of these quizzes (the "-es" ending), the expectation of success on the second one (the "-ed") was high, magnified by low fear of failure and newly discovered and developed need for achievement (Stowell & Bennet, 2010). The added value of personal well-being and safe personal environment are understood as crucial to the general outcome of assessment.

It is believed that there is an interdependence between learners' self-perception and their success in achievement-related situations. The online form of assessment in the form of a *Moodle* quiz eliminates the undesirable influence of fear of failure (it can be done again) (Yielder et al., 2017), poor self-image and low self-esteem – there is no one to judge the student, potential advantage of extroversion/introversion, inhibition – the outer world is made absent. The technology here is seen as a welcomed complement to standard class assessment methods, a new additional possibility (Galaczi, 2010). It has been stressed before (Carpenter & Alloway, 2018; Turula, 2013) that VLE participants are equally anxiety-sensitive when compared to traditional learning contexts, being subject to the same real world driving forces and constraints, all related to offline affect-related problems. However, when it comes to e-assessment, it appears that the present study offered compelling evidence to the contrary: the specific personal-safety-guaranteed environment, coupled with conditions supporting interaction with the material only rather than with other course participants or the teacher, constitutes a favourable testing scene and thus encourages students to perform at their best.

6. Conclusion

Varying the forms of assessment by incorporating technology-related solutions provides multiple opportunities for classroom success in the language aspect trained. That, in turn, boosts the self-esteem and self-confidence of all students, particularly those with a long history of well-established language anxiety (Turula, 2010; 2013).

E-assessment has certain obvious merits (Chatzigavriil et al., 2015): it enhances the quality of assessment in general; meets students' expectations, keeps up with educational standards and trends while for prospective teachers and/or managers it has the potential for experiential learning by doing training. One can also notice some inherent problems, to name but a few: the reliability of available tools is critical yet still questionable while the still common lack of institutional culture works as a barrier to e-assessment.

While extensive (and effective) application of technology throughout the assessment and learning cycles has yet to be reached, its increasing presence reveals a clear inclination towards innovative practices in assessment. Building a clear consensus as for the goals, objectives and merit of (formative and summative) e-assessment to all entities involved in the process is thus of fundamental importance.

Our findings indicate that the specific type of (test) task and its cognitive demands might advantage certain groups of learners to obtain better performance for reasons unrelated to their phonological skills / pronunciation ability. With no significant difference in the test scores, with mature and sound opinions expressed by participants in relation to the CBTests and the standard ones, it may be concluded that a feasible approach will thus be to entail using various kinds of tasks and test modes (pen-and-paper as well as online) that would help compensate for individual differences in cognitive control and stress-handling. It has become evident that the instructor needs to consider the possible consequences of task demand, individual differences in cognitive control and the general safety-promoting context when setting the necessary assessment protocols. Likewise, it is important to use various types of explicit feedback strategies to compensate for potential differences. In blended learning environments this means taking into account individual learner differences and needs as well as group dynamics, but also in addition a number of new-media-related constraints on e-learning and e-assessment.

We are fully aware that the presented research cannot form the basis for drawing final conclusions. The study was largely done in an action-research mode, implemented in a small atomic group. The results obtained have the purpose of helping the teacher work towards a more effective and more learner-centred course organization. It was done also with the purpose of

involving the students into the process of active learning and active evaluation, not just passive test-takers (Dzierzgowska, 2011).

Acknowledgement

I would like to thank the two anonymous reviewers for their constructive comments on my paper.

References

Aperliński, G., Łodzikowski, K., & Weckwerth, J. (2013). Doing phonetic transcription on Moodle. *Conference: 3rd International Conference on English Pronunciation: Issues & Practices*. Retrieved 12 April 2020 from https://www.researchgate.net/publication/348280751_Doing_phonetic_transcription_on_Moodle.

Backes, B., & Cowan, J. (2018). Is the pen mightier than the keyboard? The effect of online testing on measured student achievement. *CALDER Working Paper* No. 190. Retrieved 11 January 2021 from https://files.eric.ed.gov/fulltext/ED583636.pdf.

Chapelle, C. A., & Voss, E. (2017). Utilizing technology in language assessment. In: E. Shohamy, I. G. Or & S. May (Eds.), *Language Testing and Assessment* (pp. 149-161). Cham: Springer.

Carpenter, R., & Alloway, T. (2018). Computer versus Paper-Based Testing: Are they equivalent when it comes to working memory? *Journal of Psychoeducational Assessment 37*, 1-13.

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House*, *83*(1), 1-6.

Chatzigavriil, A., Tarini, F., & Malte, W. (2015). *E-assessment practice at Russell Group Universities.* The London School of Economics and Political Science, Learning Technology and Innovation, London, UK. Retrieved 10 April 2019 from https://core.ac.uk/download/pdf/35437417.pdf.

Cheng, L., & Fox, J. (2017). *Assessment in the Language Classroom*. London: Palgrave.

Cheng, L., Klinger, D., Fox, J., Doe, C., Jin, Y., & Wu, J. (2014). Motivation and test anxiety in test performance across three testing contexts: The CAEL, CET, and GEPT. *TESOL Quarterly*, *48*, 300-330.

Cheng, L., Rodgers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, *21*(3), 360-389.

Coll, C., Barberà, E., & Onrubia, J. (2000). Catering for student diversity in assessment practices. *Journal for the Study of Education and Development*, *23*(90), 111-132. Retrieved 14 January 2021 from https://www.academia.edu/29889899/La_atenci%C3%B3n_a_la_diversidad_en_las_pr%C3%A1cticas_de_evaluaci%C3%B3n_1_Catering_for_student_diversity_in_assessment_practices.

Cummins, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics*, *29*, 90-100.

Czetwertyńska, G. (2015). Ocenianie jako nauczanie i uczenie się. *Języki Obce w Szkole*, *2015/03*, 4-9.

Dooey, P. (2008). Language testing and technology: Problems of transition to a new era. *ReCALL*, *20*(1), 21-34.

Dzierzgowska, I. (2011). *Jak uczyć metodami aktywnymi*. Warszawa: Fraszka Edukacyjna.

Eckman, F. R., Iverson, G. K., Fox, R. A., Jacewicz, E., & Lee, S. (2009). Perception and production in the acquisition of L2 phonemic contrasts. In M. A. Watkins, A. S. Rauber & B. O. Baptista (Eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production* (pp. 81-95). Cambridge: Cambridge Scholars Publishing.

Educational Testing Service, (2009). *Guidelines for the Assessment of English Language Learners*. Retrieved 11 January 2021 from https://www.ets.org/s/about/pdf/ell_guidelines.pdf.

Ellis, R. (2008). *The Study of Second Language Acquisition.* 2nd edition. Oxford: Oxford University Press.

Escudero, P. (2007). Second language phonology: The role of perception. In M. Pennington (Ed.), *Phonology in Context* (109-134). London: Palgrave McMillan,.

Gabryś-Barker, D. (2012). Explaining affectivity in second/foreign language learning. In D. Gabryś-Barker (Ed.), *Readings in Second Language Acquisition* (pp. 175-200). Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Galaczi, E. (2010). Face-to-face and computer-based assessment of speaking: Challenges and opportunities. In L. Araújo (Ed.), *Computer-based Assessment of Foreign Language Speaking Skills* (pp. 29-51). Retrieved 12 January 2021 from http://publications.jrc.ec.europa.eu/repository/bitstream/111111111/15037/1/lbna24558enc.pdf.

Isaacs, T. (2014). Assessing pronunciation. In A. J. Kunnan (Ed.), *The Ccompanion to Language Assessment* (pp. 140-155). Oxford: Wiley-Blackwell.

İnceçay, G., & İnceçay, V. (2010) A case study on needs assessment of English language teachers. *Procedia Social and Behavioral Sciences*, *2*, 317-321.

Jedynak, M. (2013). Affectivity in learning L2 phonology/phonetics – The role of self-concept in successful acquisition of English pronunciation. In D. Gabryś-Barker & J. Bielska (Eds.), *The Affective Dimension in Second Language Acquisition* (pp. 60-73). Bristol: Multilingual Matters.

Komorowska, H. (2005). Testowanie językowych osiągnięć uczniów – trudności i zagrożenia. *Języki Obce w Szkole, 2005/06*, 41-65.

Komorowska, H. (2002). *Sprawdzanie umiejętności w nauce języka obcego. Kontrola - Ocena - Testowanie*. Warszawa: Fraszka Edukacyjna.

Lecumberri, M. L., G. & Maidment, J. A. (2000). *English Transcription Course*. London: Arnold.

Licorish, S. A., George, J. L., Owen, H. E., & Daniel, B. (2017). "Go *Kahoot!*" Enriching classroom engagement, motivation and learning experience with games. In W. Chen, et al. (Eds.), *Proceedings of the 25th International Conference on Computers in Education* (pp. 755-764). New Zealand: Asia-Pacific Society for Computers in Education.

Licorish, S. A., Owen, H. E., Daniel, B., & George, J. L. (2018). Students' perception of *Kahoot!*'s influence on teaching and learning. *Research and Practice in Technology Enhanced Learning*, *13*(9). Retrieved 16 March 2019 from https://doi.org/10.1186/s41039-018-0078-8

Lu, D. (2002). Phonetic symbols: A necessary stepping stone for ESL learners. *Forum*, *40*(4), 36-39.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative Data Analysis. A Methods Sourcebook*. Third edition. Los Angeles: SAGE.

Mora, J. C., & Darcy, I. (2017). The relationship between cognitive control and pronunciation in a second language. In T. Isaac. & P. Trofimovich (Eds.), *Second Language Pronunciation Assessment: Interdisciplinary Perspectives* (pp. 95-120). Bristol: Blue Ridge Summit: Multilingual Matters / Channel View Publications. Retrieved 16 march 2019 from http://www.jstor.org/stable/10.21832/j.ctt1xp3wcc.10

Neill, J. (2018). Online quizzes and exams with Moodle. Wikiversity. Retrieved March 9, 2019, from

https://en.wikiversity.org/w/index.php?title=User:Jtneill/Teaching/Online_Quizzes_and_Exams_with_Moodle&oldid=1862062

Niewiadomski, K. (2010). Zastosowanie nowoczesnych telefonów komórkowych jako pomocy dydaktycznej w nauce fonetyki języka angielskiego. In M. Pawlak & E. Waniek-Klimczak (Eds.), *Mówienie w języku obcym – sukcesy i porażki uczenia się i nauczania* (pp. 277-285). Poznań-Kalisz-Konin: Wydział Pedagogiczno-Artystyczny UAM w Poznaniu & PWSZ w Koninie.

Prisacari, A. A., & Danielson, J. (2017). Computer-based versus paper-based testing: Investigating testing mode with cognitive load and scratch paper use. *Computers in Human Behavior*, *77*, 1-10. Retrieved 16 January 2021 from https://doi.org/10.1016/j.chb.2017.07.044

Plump, C. M., & LaRosa, J. (2017).Using *Kahoot!* in the classroom to create engagement and active learning: A game-based technology solution for elearning novices. *Management Teaching Review*, *2*(2), 151-158. Retrieved 17 March 2019 from https://doi.org/10.1177/2379298116689783

Shohamy, E., Inbar-Lourie, O., & Poehner, M. E. (2008). *Investigating Assessment Perceptions and Practices in the Advanced Foreign Language Classroom*. (Report No. 1108). University Park, PA: Center for Advanced Language Proficiency Education and Research.

Stowell, J., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research*, *42*, 161-171.

Turula, A. (2010). *Teaching English as a Foreign Language: From Theory to Practice and All the Way Back*. Częstochowa: Wydawnictwo Wyższej Szkoły Lingwistycznej.

Turula, A. (2013). Affect in VLEs: Anxiety and motivation in blended EFL teacher training. In D. Gabryś-Barker & J. Bielska (Eds.), *The Affective Dimension in Second Language Acquisition* (pp. 254-267). Bristol: Multilingual Matters.

Underhill, N. (1987). *Testing Spoken Language*. Cambridge: Cambridge University Press.

Wach, A. (2012). Classroom-based language efficiency assessment: A challenge for EFL teachers. *Glottodidactica*, *39*(1), 81-92.

Waniek-Klimczak, E., Porzuczek, A., & Rojczyk, A. (2013). Affective dimensions in SL pronunciation: A large-scale attitude study. In D. Gabryś-Barker & J. Bielska (Eds.), *The Affective Dimension in Second Language Acquisition* (pp. 124-137). Bristol: Multilingual Matters.

VanPatten, B., Trego, D., & Hopkins, W. (2015). In-class vs. online testing in university-level language courses: A research report. *Foreign Language Annals*, *48*(4), 659-668. Retrieved 16 January 2021 from n/a-n/a. 10.1111/flan.12160.

Vavla, L., & Gokaj, R. (2013). Learner's perceptions of assessment and testing in EFL classrooms in Albania. *Mediterranean Journal of Social Sciences*, *4*(11), 509-515.

Werfel, K. L. (2017) Phonetic transcription training improves adults' explicit phonemic awareness: Evidence from undergraduate students. *Communication Disorders Quarterly*, *39*(1), 281-287. Retrieved 16 March 2019 from sagepub.com/journalsPermissions.nav

Wu, J., & Lee, M. C. (2017). The relationships between test performance and students' perceptions of learning motivation, test value, and test anxiety in the context of the English benchmark requirement for graduation in Taiwan's universities. *Language Testing in Asia*, *7*(9). Retrieved 16 January 2021 from https://doi.org/10.1186/s40468-017-0041-4

Yielder, J., Wearn, A., Chen, Y., Henning, M. A., Weller, J., Lillis, S., & Mogol, V. (2017). A qualitative exploration of student perceptions of the impact of progress tests on learning and emotional wellbeing. *BMC Medical Education, 17*(148). Retrieved 15 January 2021 from https://doi.org/10.1186/s12909-017-0984-2.