



Classroom assessment practices of English language instructors *

Meryem Özdemir-Yılmaz ^a † , Yonca Özkan ^b 

^a Cukurova University, School of Foreign Languages, Adana 01380, Turkey

^b Cukurova University, English Language Teaching Department, Adana 01380, Turkey

APA Citation:

Özdemir-Yılmaz, M., & Özkan, Y. (2017). Classroom assessment practices of English language instructor. *Journal of Language and Linguistic Studies*, 13(2), 324-345.

Submission Date:29/05/2017

Acceptance Date:11/08/2017

Abstract

Assessment of students is an essential part of instruction in both teaching and learning. With the recognition of alternative assessment methods, classroom assessment has gained attention focusing on learning of students. However, high-stakes testing turns classroom assessment into teachers' high stakes decisions, ignoring the development of learners. In the context of language teaching at tertiary level, school of foreign languages serves as a gatekeeper by deciding on whether new students are proficient enough to start their professional education. Therefore, these schools impose a proficiency exam whose impact is relatively high-stakes. Thus, this study aims to have a descriptive investigation of the classroom assessment practices of instructors by considering the purpose, methods, and procedures of assessment and compares the context between state and private universities. The data was collected through survey questionnaire which includes both multiple choice and open-ended questions. Based on the quantitative and qualitative data, the results depicted that uniformity regarding classroom assessment practices of the instructors was observed; however, to what extent this uniformity embraces formative assessment practices needs to be further explored. The study also implies instructors' need for such training that involves theoretical and practical aspects of classroom assessment at both pre-service and in-service level.

© 2017 JLLS and the Authors - Published by JLLS.

Keywords: English language learning/teaching; classroom assessment; language assessment; language instructors; formative assessment

1. Introduction

Student assessment plays a significant role in almost every teaching and learning context. Assessment for learning has recently been voiced in educational context along with the emergence of formative classroom assessment methods as an alternative to psychometric summative assessment (Assessment Reform Group, 1999; McDowell et al., 2011). The paradigm shift towards more constructive methodologies which focus more on student development might play a central role in adopting formative assessment methods in classes. As a result, teacher role in assessment has been

* A small part of this study was orally presented at 5th International Conference on Language Literature and Culture, Burdur, Turkey

† Corresponding author. Tel.: +90-322-338-6084/2921

E-mail address: mozdemir@cu.edu.tr

changing to be teachers as “agents of assessment” (Rea-Dickens 2004; Teasdale and Leung, 2000; Yin, 2010).

Although formative assessment seems to be promising in terms of learning about the progressive development of learners, it has been questioned for not producing reliable results as compared to summative assessment. "Teacher individuality" is echoed as one of the reasons because individual teachers have individual assessment practices, which hinders the reliability dimension of assessment (Shepard, 2000). Therefore, enabling commonality of assessment practices among teachers is one of the key points that should be considered. As highlighted by Teasdale and Leung (2000), undermining current assessment practices of teachers might be the first step of building a commonality of practice. Hence, this study aims first to describe classroom assessment practices of English language instructors teaching English at the preparatory classes of universities regarding their purpose in assessment, methods of assessment and assessment procedures. The context is limited to instructors' practices at schools of foreign languages due to their clear role description at tertiary level, which is increasing language proficiency of university students before they start their professional training at the departments. In spite of the shared mission of different universities, the way of fulfilling it may vary in different settings. Therefore, the second aim of the study is to carry out a comparison between assessment practices of instructors working at state and private universities. The type of university (whether it is state or private) is considered as a variable for this study since organizational setting may be diverse in state and private schools though their context-specific objectives in teaching presumably overlap to a great extent.

1.1. Literature review

As a result of the paradigm shift towards constructive methodologies, learner development in progress has been emphasized in assessment field through formative assessment. This shift has also changed the roles of teachers, which make teachers active in assessment. They become "agents of assessment" (Teasdale and Leung, 2000) who can learn about the development of learners in progress and intervene in the potential development of learners.

When considered from this aspect, formative classroom assessment sounds promising for teachers regarding its feeding into their practices to enhance their learners' development and for learners regarding its contribution to their development. Among various terminologies of formative classroom assessment, there is an apparent focus on the learning side of assessment such as learning-oriented assessment (Carless, 2007) or assessment for learning (McDowell et al., 2011). However, this novel and promising way of assessment is not without some theoretical and conceptual problems making it have a precarious position against psychometric testing which has clearly defined boundaries (Leung, 2004; Teasdale and Leung, 2000; Sadler, 1998).

One of these problems arising in classroom-based assessment is teacher individuality for certain. Considering that each teacher employs classroom assessment by acting on their personal beliefs, values, and assumptions, it is not easy to mention about a commonality of practice within the scope of classroom-based assessment (Borg, 2003; Leung, 2004). The studies conducted on classroom assessment practices of teachers confirm this argument. For instance, the results of Torrance and Pryor's (1998) investigation on incidents of classroom assessment of elementary school teachers implied that there are different perceptions and attitudes between teachers and even individual teachers have personal contradictories within themselves. In the similar vein, the results Davison's study (2004) depicted a continuum having "technical orientation" at one end and "assessor as God" orientation at the other end. By "technical orientation" Davison (2004) refers to teachers who regard assessment as strictly following to published criteria whereas he refers to teachers of whom

assessment judgments are based on unarticulated references by "assessor as God" orientation. Davison (2004) stated that classroom orientations of teachers have a place between these two extreme ends in a divergent manner. These individualized teacher practices of classroom assessment were analyzed by Yin (2010) through a case study in which underlying cognitions of two English teachers in English for Academic Purposes context were explored. The results of the study indicated that participating teachers held common cognitions regarding assessment, but these cognitions were individualized ones, which were reasoned by their individual differences in personal and professional background, approach, and experience. On the other hand, common cognitions were also identified, which were regarded as cognition as a basic pattern of thinking (Yin, 2010).

Apart from the individual differences in practices of classroom assessment, context is another concern factor influencing these practices. The status of English, for instance, was identified as a contextual factor that influences teachers' assessment practices (Cheng, Rogers & Hu, 2004). Variances among teachers were reported across different settings of Canada, Hong Kong and Beijing where the status of English differs (Cheng et al., 2004). Moreover, in the abovementioned study of Davison (2004), the difference among teachers' classroom assessment practices was observed according to the cities they were working in (which are Melbourne and Hong Kong). This difference was attributed to the presence and/or absence of mandated standards because teachers who held more standard views on assessment are from Melbourne where the government provides criteria on assessment. On the contrary, teachers from Hong Kong were found to have diverse practices as result of an absence of mandated standards (Davison, 2004).

The standards based assessment at the organizational level commonly involves summative testing via an exit test conducted at the end of the teaching process. The stakes of this kind of tests are high which in turns influence the acts of teachers and learners in the process. Therefore, incorporating such tests might be regarded as another influential factor in determining classroom assessment practices of teachers, which is called as the washback effect (Messick, 1996). As a result, this effect of testing should be considered painstakingly as the higher the stakes, the more frequently teachers adopt the format of exit proficiency tests (Cheng et al., 2008).

As indicated, previous studies depict the individual and contextual complexity of classroom-based assessment, which confirms Leung's (2004) concern for its need of epistemological and empirical conceptualization. However, as asserted by Teasdale and Leung (2000), the situation of classroom-based assessment should not be described as a hopeless case, rather action should be taken starting from the questioning of the current practices of classroom assessment. One recommendation for this case of classroom-based assessment has been expressed by Leung (2004) by which he calls for some context-specific sets of principles, values, and knowledge guiding and informing the practice of assessment, whereby a community of practice could be formed leading consistency in assessment practices.

Drawing on the arguments in the relevant literature, this study first aims to describe classroom assessment practices of English language instructors teaching English at the preparatory classes of universities regarding their purpose in assessment, methods of assessment and assessment procedures. The context is limited to instructors' practices at schools of foreign languages due to their clear role description at tertiary level, which is increasing language proficiency of university students before they start their professional training at the departments. In spite of the shared mission of different universities, the way of fulfilling it may vary in different settings. Therefore, the second aim of the study is to carry out a comparison between assessment practices of instructors working at state and private universities. The type of university (whether it is state or private) is considered as a variable for this study. The organizational setting may be diverse in state and private schools although their context-specific objectives in teaching presumably overlap to a great extent.

1.2. Research questions

The following research questions guide this study:

1. What assessment purposes, methods and procedures do English language instructors report utilizing in their classroom assessment practices?
2. Is there a difference between state and private university English language instructors in terms of classroom assessment practices?

2. Method

2.1. Context of the Study

The study was conducted in the context of English language teaching at tertiary level in Turkey. In Turkey, higher education is mandated by Council of Higher Education (CHE) which regulates and monitors the education at tertiary level. According to CHE, the aims of foreign language education involve “teaching the basic rules of foreign language that the students study, to improve their foreign language vocabulary knowledge, to ensure that they can understand what they read and listen, and to ensure that they can express themselves orally or verbally” (CHE, 2008). Meanwhile, the CHE obliged universities to offer preparatory classes to students if they are enrolled in departments which serve courses in a foreign language or courses both in Turkish and in a foreign language (2008). The students enrolling to these departments are supposed to pass the proficiency exam or else participate in the preparatory classes in one semester, and if they become unsuccessful, they participate in the preparatory classes one more semester at the end of which another proficiency exam takes place. The presence of proficiency exam at the beginning of the university education and the requirements of attending semester-based proficiency exams are stated explicitly in the regulations of CHE (2008). Also, the weekly and per semester course hour load are specified as having minimum 20 hours a week and 260 hours per semester (2008). However, no regulations are mandated about the content of the proficiency exam, no standardization is provided for the weight of language skills in both exams as well as teaching objectives, and no criteria is watched for the evaluation process of proficiency exams. In practice, schools of foreign languages which are responsible for offering foreign language preparatory classes follow these regulations (which is a must) but are quite flexible in adopting any teaching and assessment policy. For instance, in some universities, the language courses are offered in a traditional manner while a modular system is adopted in others. Therefore, it is difficult, if not impossible, to mention about standardization across the language teaching practices of the school of foreign languages which is not at least regulated by CHE.

2.2. Participants

As a result of convenient sampling, 70 English language instructors working at the school of foreign languages of different universities participated in the study. The elicited personal and professional information presents the portrait of participants. That is, out of 70 participants, 43 of them are female, and 27 of them are male, most of whom have the teaching experience between 6-10 years (40%). Most of the participants are graduates of English Language Teaching Department (81%), and the majority of them hold a master's degree (59%). Considering the type of the institution, 44 of the English language instructors work at state universities as compared to 26 instructors who work at private universities. Further, the average size of instructors' classes is 24 students, and their weekly work load is 22 hours in average. Lastly, 33 instructors stated that they had training on assessment. When further investigated, the type of the training that instructors exposed to is limited to BA level

courses or organizational orientation programs for test office members with few exceptions of participated workshops and webinars on assessment.

2.3. Data Collection Tool and Data Analysis

The data collection tool is the adopted version of survey questionnaire developed by Cheng et al. (2004). It comprised of five parts: *a) personal/professional information, b) assessment purpose, c) assessment methods, d) procedures of assessment and e) open-ended questions* (please see Appendix). The questionnaire was modified to the context under study and piloted among a small number (n:10) of English language instructors. 35 minutes in average was required to complete the questionnaire. The questionnaire was disseminated to instructors via *Google Forms®* except for the small portion of them with whom the questionnaire was delivered face to face. The reason for choosing *Google Forms®* was its feasibility and practicality enabling the researchers to access more participants in the context of the study.

As for the data analysis, SPSS 15.0 was utilized to code and analyze the numerical values within the questionnaire. The first aim of the study was to describe the case of the classroom assessment practices in English language education context at Turkish universities; therefore, descriptive statistics were the main method of the quantitative data analysis. As the second aim of this study was a comparison between state and private universities, the chi-square test was utilized. The analysis of comparison was carried out with a non-parametric test due to dichotomous nature of variables (Cheng et al. 2004). Meanwhile, the responses to open-ended questions were analyzed through content analysis of which results were triangulated with the quantitative data analysis results.

3. Results

3.1. Purposes of Assessment

In the first part of the questionnaire, the aim was to determine the classroom assessment purposes of English language instructors. In Table 1 below, the overall descriptive results of assessment purposes were reported along with the descriptive results of assessment purposes of instructors working at state and private universities. Chi-square test was operated for comparison of private and state university instructors' classroom assessment purposes, apart from only one item not meeting the prerequisite of the test assuming "at least 80% of the expected frequencies should be 5 or larger" (Morgan et al., 2011).

Table 1. Assessment purposes of instructors

Purpose of assessment	overall (%)	state universities (%)	private universities (%)	Significance of comparison
<i>Student-centered</i>				
To gain information about my students' progress in the targeted domain	72,9	72,7	73,2	0,975
To give feedback to my students as they progress through the course	67,1	63,6	73,1	0,416
To diagnose strengths and weaknesses in my students	54,3	54,5	53,8	0,955
To motivate my students to learn	45,7	45,5	46,2	0,955

To determine the final grades for my students	40	38,6	42,3	0,762
To make my students work harder	35,7	40,9	26,9	0,238
To prepare my students for standardized tests they will need to take in the future	31,4	29,5	34,6	0,659
To formally record growth in learning	21,4	20,5	23,1	0,796
<i>Instruction</i>				
To diagnose strengths and weaknesses in my own teaching and instruction	45,7	43,2	50,0	0,580
To plan my instruction	34,3	34,1	34,6	0,964
To group my students for instruction purposes in my class	11,4	13,6	7,7	-----
<i>Administration</i>				
To provide information to central administration (e.g. school, administration, higher education institution (CHE/YOK))	31,4	34,1	26,9	0,533

As seen in Table 1, the assessment purposes were categorized into three: *student-centered purposes*, *instructional purposes*, and *administrative purposes*. According to the descriptive results, the student-centered purpose of “*gaining information about my students’ progress in the targeted domain*” had the highest percentage among all the instructors (72,9%). It was also the highest ranked purpose among both state (72,7%) and private (73,2%) university instructors. 67,1% of instructors ranked the student-centered purpose of “*giving feedback to my students as they progress through the course*” second. It was kept in number two by both state university instructors (63,6%) and private university instructors (73,1%). As for the instructional purposes, the purpose of “*diagnosing strengths and weaknesses in my own teaching and instruction*” was the top ranked purpose (45,7%); and for the administrative purposes, not many instructors had the purpose of “*providing information to central administration*” (31,4%). Among all the purposes of assessment, the least indicated one by instructors was “*to group my students for instruction purposes in my class*” (11,4%). Statistically significant difference was not found between state and private university instructors regarding any purpose statements in the questionnaire ($p>0,05$) which indicates that the context is not an influential factor impacting assessment purposes of instructors.

The descriptive results were supported by the content analysis results which were reported in Table 2 below. The constructs were categorized under the themes associating the purpose statements included in the questionnaire; however, some new themes (purposes) were elicited in the content analysis.

Table 2. The results of content analysis

Purpose of assessment	Frequency of Constructs(f)
<i>Student centered</i>	
To make students monitor their own learning in progress	35
To diagnose strengths and weaknesses in my students	33

To motivate my students to learn	15
<i>Instruction</i>	
To diagnose effectiveness of my own teaching and instruction	42
To plan my instruction/to improve my instruction	28
<i>Administration</i>	
To give feedback to administration about the extent how much goals and objectives are achieved	20

The results of the content analysis were found to adhere to the assessment purposes indicated in the descriptive analysis. While instructors ranked instructive purpose of “*diagnosing strengths and weaknesses in my own teaching and instruction*” fourth (45,7%), it is the most frequent construct elicited in the content analysis. However, the theme of purpose was modified in “to *diagnose effectiveness of my own teaching and instruction*” because instructors commonly stated that they were evaluating the effectiveness of their instruction rather than having a detailed reflection on weak and strong points of their instruction. This notion is quite explicit in the quotes of the instructors below:

It [classroom assessment] shows whether the instruction has been effective. (Instructor 19 (I19))

S/he added:

It [classroom assessment shows] whether our teaching methods have been effective or not.

Instructor 9 (I9) indicated:

It [classroom assessment] gives feedback about whether your way of teaching works or not.

The discrepancy between descriptive results and content analysis results may be due to this slight but remarkable difference in the understanding of instructors about the instructional purpose of classroom assessment. Furthermore, another theme of purpose is “*to make students monitor their own learning in progress*” which is associated with the descriptively top-ranked purpose of “*gaining information about my students’ progress in the targeted domain.*” The difference in the statement might indicate a further insight that instructors not only gain information about their students' progress but use this information to engage students in the process. At this point, the results appear to be overlapped.

Overall, the results revealed that English language instructors have both student-centered and instructional purposes in their classroom assessment practices. These purposes are much more related to a diagnostic purpose, one of the purposes of assessment. However, how this diagnostic data is reflected in teaching and learning is the main concern of alternative classroom assessment. Although instructors stated that they *plan and/or improve their instruction accordingly* (f: 28), only 34,3% of instructors indicated that they do so. In essence, the purpose of classroom assessment is expected to be practical in the learning process by going beyond being diagnostic in nature.

To understand how instructors use this diagnostic information about students' progress may emerge in the form and quality of feedback that they deliver in the classroom. Taking this assumption into consideration, the methods that instructors use while delivering feedback were inquired in the procedures of assessment part of the questionnaire. The responses provided in the form of feedback that instructors use were reported in Table 3.

Table 3. The form of feedback provided to students

Source	overall (%)	state universities (%)	private universities (%)	Significance of comparison
<i>During the course</i>				
Verbal feedback	90	88,6	92,3	-----
Written comments	62,9	68,2	53,8	0,230
Total test score	35,7	34,1	38,5	0,712
Conference with student	25,7	25,0	26,9	0,859
Checklist	15,7	15,9	15,4	-----
A letter grade	10,0	11,5	9,1	-----
Teaching diary/log	5,7	6,8	3,8	-----
<i>Final report</i>				
Total test score	58,6	50,0	73,1	0,058
Written comments	52,9	65,9	30,8	0,004
Conference with student	24,3	27,3	19,2	0,448
Checklist	14,3	11,4	19,2	-----
Teaching diary/log	5,7	6,8	3,8	-----
A letter grade	20,0	18,2	23,1	0,621

As it is evident in Table 3, the instructors frequently preferred to provide *verbal feedback* during the course (90%). Also, *verbal feedback* was leading source of feedback given during the course by both state (86,2%) and private university instructors (92,3%). The secondly ranked form of feedback delivered during the course was *written comments* (62,9%) by both state and private university instructors. As for the form of feedback delivered as a final report, *total test score* was mostly preferred one among instructors (58,6%). However, state and private university instructors had different rankings for the feedback sources of *total test score* and *written comments*. That is, private university instructors indicated that they used *total test score* as a feedback source of final report (73,1%) while *verbal comments* were the most frequently preferred source of feedback by state university instructors (65,9%). Moreover, the results of chi-square test indicated a significant difference between state and private university instructors regarding their use of *written comments* as a feedback source for a final report ($p > 0,05$). Other than this feedback source, no significant difference was observed in the items for which chi-square test could be operated.

Verbal feedback delivered during the course seems to be the de facto source of feedback among instructors. This finding is in parallel with the general purpose of instructors which is the diagnosis of strong and weak points of the students, which means that instructors do not only diagnose and give a mark to the students but inform them verbally about these points. This case is in line with the main idea of alternative assessment; however, other than the source, the content of the feedback is essential for the progress of students in the learning process. Therefore, the points that a teacher should consider while delivering feedback during the course were asked to instructors and the answers were analyzed through content analysis of which results were reported in Table 4.

Table 4. The content of feedback

Feedback	Frequency of Constructs (f)
Manner of giving feedback	42
Constructive feedback	9
Process-based feedback	7
Channel of feedback	6
Skill-based feedback	5
Motivation for students	4
Clear (understandable) feedback	3
Taking students' proficiency levels in consideration	3
Alignment with proficiency test	2

The themes elicited through content analysis for the content of feedback are observed in Table 4. The most underlined theme was “*manner of feedback*” along with 42 constructs. The instructors frequently mentioned about the way of delivering feedback rather than the intent of giving feedback or information about what kind of feedback is useful for students' progress. Some quotes taken from instructors' comments are provided below:

[A teacher should consider] not hurting the students while giving feedback and prompting the students to be more successful. (I5)

The teacher should make clear that the feedback is not for judging the student; it is for guiding the student into being an efficient learner. I myself prefer to start with positive sides in my assessment than [with] negative [sides]. (I38)

[A teacher should consider] using an appropriate language and starting with positive aspects [while delivering feedback]. (I7)

Utmost care should be taken to avoid offending the students while aiming to provide them an opportunity for introspection. (I17)

It is clear that manner of giving feedback is a significant characteristic for instructors. Although it is important to have a positive and appropriate attitude while giving feedback, the instructors were expected to focus more on the content of feedback which would contribute to the development of students' language proficiency. Only nine constructs were elicited for themes suggesting that feedback should be constructivist and only seven for the process-based feedback.

3.2. Procedure of Assessment

In the procedure of assessment part, the instructors were also asked to respond the questions for their choices of test items and the time spent on assessment. The results of these two questions were reported below.

Table 5. Sources of assessment items and time spent on assessment

Source	overall (%)	state universities (%)	private universities (%)	Significance of comparison
<i>Instructor</i>				
Items written by myself	42,9	38,6	50,8	0,353
<i>Other instructors</i>				
Items prepared together with other teachers	45,7	34,1	65,4	0,011
<i>Print sources</i>				
Items prepared by test office of the institution	77,1	79,5	73,1	0,533
Items from textbooks	65,7	68,2	61,5	0,572
Items from mandated syllabuses/curricula proposed by Higher Education Institution (CHE/YÖK)	8,6	9,1	7,7	-----
<i>Internet</i>				
Items extracted from the Internet	44,3	40,9	50,0	0,459
<i>Time spent on assessment</i>				
%5-or less	1,4	2,3	0	
%10-%15	18,6	11,4	30,8	
%20-%30	35,7	38,6	30,8	-----
%40-%50	34,3	34,1	34,6	
%50 and more	10	13,6	3,8	

Table 5 demonstrates that instructors were inclined to utilize print sources most of the time. Both state and private university instructors indicated that they mostly used “items prepared by test office of the institution” (77,1%). As their second preference for test items, overall results showed that the instructors used “items from textbooks” (65,7%). Although “items from textbooks” as a source of test items were remained in the second row by state university instructors (68,2%), the second preference of private university instructors was “items prepared together with other teachers” (65,4%). This result indicates that private university instructors used these items much more frequently than state university instructors. Also, this difference between instructors was found to be statistically different ($p>0,05$).

The distribution of time spent by instructors on assessment is also shown in Table 5. That is, 35,7% of the instructors indicated that they spent their time on assessment between 20% and 30% of their total working time during a term. This time span is valid for 38,6% of the state university instructors. However, the majority of private university instructors (34,6%) stated that they spent their time on assessment between 40% and 50% of their total working time during a term. Although the descriptive results highlight that private university instructors spend more time on assessment, no statistically significant difference was found for state and private university instructors regarding their allocated time on assessment during a term.

Overall, the results demonstrated that instructors rely more on their test offices and published sources for their classroom assessment practices and they mostly spend less than 50% of their total time on assessment. The reason for this case may reflect instructors’ reliance on test offices for

assessment which is a result of standardization imposed by administration through test offices. When instructors' self-evaluation regarding their assessment practices was asked, the broad themes of "yes (I feel myself effective in assessment), no (I do not feel myself effective in assessment), no idea, and undecided" themes were elicited in content analysis accompanied by the themes elicited with regard to the reasons behind these evaluations. The results were reported in Table 6 below.

Table 6. Self-assessment of instructors

Self-assessment of instructors	Frequency of Constructs(f)
<i>Yes</i>	
giving feedback	13
having knowledge of assessment	5
use of various assessment tools	2
to be able to decide on the level of ss	2
having a focus on the process	1
knowing needs of students	1
being objective	1
<i>TOTAL</i>	25
<i>No</i>	
lack of training	11
system requirements	4
impact of end-of-year test	4
not believing in assessment	3
lack of time	2
<i>TOTAL</i>	24
<i>No Idea</i>	2
<i>Undecided</i>	5

As it is explicit in Table 6, the instructors who felt themselves to be effective in assessing the students associated their effectiveness to the notion of giving feedback to their students (f:13). The statements below were extracts from instructors' words:

It is difficult to answer, but I think I am effective in giving feedback as my students can build on the feedback that I give them. (I20)

I think I'm an effective assessor I try to give feedback orally and written all the time. And I look in detail. (I1)

[...] giving clear feedback. [students can understand what I mean easily]. [There should] not [be] destructive comments, feedback should be constructive (I9)

Moreover, some other instructors indicated in their statements that they did not feel as effective assessors as they lacked training on assessment (in their statements), *they are doing what system requires* (f:4) and feeling *the impact of end-of-year test* (summative testing) (f:4). These statements were like:

[...] the system does not allow us [instructors] to use in-class assessments such as projects, portfolios, blogs, drama...etc as we have an obligatory pre-planned syllabus to implement and a pre-determined exam system that never changes. (I10)

No, because I am not competent enough about effective ways of conducting classroom assessment. (I23)

I try hard to be an effective assessor however the curriculum and the nature of end-of-year test limit my assessment practices. (I8)

Overall, under the light of emerging constructs in the content analysis, the main reason why instructors feel as effective assessors is their reliance on their ability to give feedback. On the other hand, instructors who felt ineffective reported this fact was highly related to their limited training on assessment and their need for more training on assessment.

3.3. Skill Assessment and Skill Assessment Methods

In order to have a profound understanding of how instructors conduct the assessment on skills of English, their methods of skill assessment were also investigated. However, before having a detailed investigation of the assessment methods, the most and least assessed skill which was integrated into their classroom assessment practices was asked to the instructors, and the descriptive results of their answers were presented in Figure 1 below.

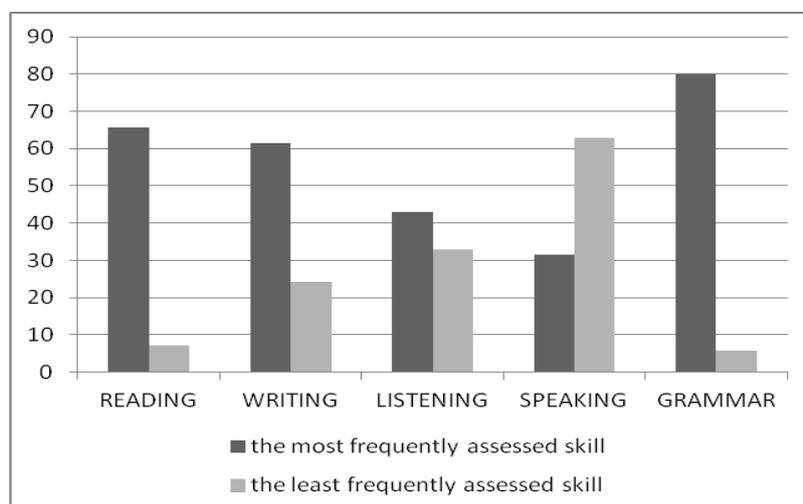


Figure 1. The most and least frequently assessed skills by instructors in classroom assessment practices.

As explicitly seen in Figure 1, English language instructors mostly stated that they frequently assessed grammar of English in their classroom assessments (80%) followed by reading (%65,7) and writing (61,4%). The least frequently assessed skill, on the other hand, is speaking (62,9%) along with listening skill (32,9%). Chi-square test was also operated for the difference between state and private university instructors regarding their most and least assessed skill in their classroom assessment practices; however, no significant difference was found ($p > 0,05$).

Other than the most and least assessed skill in the classroom, which methods instructors utilized in their classes while assessing one of these skills were under investigation of which results are presented in Table 7 below.

Table 7. Assessment methods of reading, writing, listening and speaking

ASSESSMENT METHODS	overall (%)	state universities (%)	private universities (%)	Significance of comparison
<i>READING</i>				
<i>Instructor-made</i>				
Teacher made tests that may involve different types of items (e.g. true/false questions, fill in the blanks, comprehension questions)	77,1	79,5	73,1	0,533
<i>Student-conducted</i>				
Student summaries of what is read	45,7	45,5	46,2	0,955
Oral interviews/questioning	38,6	40,9	34,6	0,601
Learner portfolio/e-portfolio	11,4	13,6	7,7	-----
Peer assessment	10,0	9,1	11,5	-----
Self assessment	10,0	2,3	23,1	-----
Learner Journal	2,9	4,5	0	-----
<i>Non-instructor developed</i>				
Standardized reading tests	47,1	50,0	42,3	0,533
<i>WRITING</i>				
<i>Instructor-made</i>				
Teacher made tests that may involve different types of items (e.g. true/false questions, matching questions, paragraph/essay writing)	51,4	52,3	50,0	0,854
<i>Student-conducted</i>				
Learner portfolio/e-portfolio	57,1	65,9	42,3	0,054
Self assessment	24,3	29,5	15,4	0,182
Peer assessment	15,7	20,5	7,7	-----
Learner journal	8,6	9,1	7,7	-----
<i>Non-instructor developed</i>				
Standardized writing tests	45,7	43,2	50,0	0,580
<i>LISTENING & SPEAKING</i>				
<i>Instructor-made</i>				
Teacher made tests that may involve different types of tasks (e.g. taking notes, retelling)	51,4	54,5	46,2	0,497
<i>Student-conducted</i>				
Oral presentations	58,6	59,1	57,7	0,909
Oral interviews/dialogues	48,6	56,8	34,6	0,073
Oral discussion with each student	52,9	52,3	53,8	0,899
Peer assessment	20,0	20,5	19,2	0,902
Public speaking	18,6	20,5	15,4	-----
Oral reading/dictation	17,1	15,9	19,2	-----
Self assessment	11,4	13,6	7,7	-----
<i>Non-instructor developed</i>				
Standardized speaking test	41,4	36,4	50,0	0,263
Standardized listening test	42,9	40,9	46,2	0,668

Assessment methods were divided into three main groups which were instructor-made tests, student-conducted tests and non-instructor developed tests (Cheng et al. 2004) for each skill of language. Speaking and listening assessment has been regarded in unity regarding methods that are

used by the instructors. The results depicted in Table 1 shows that English language instructors frequently used instructor-made methods (77,1%) in reading assessment followed by “standardized reading tests” (47,1%) grouped under non-instructor developed tests. In reading assessment, both state (79,5%) and private university instructors (73,1) indicated that they utilized instructor-made tests most of the time. Second rankings of instructors were slightly different where private university instructors stated to use “student summaries of what is read” (46,2%) on contrary to state university instructors’ preference for “standardized reading tests” (50,0%). However, no significant difference was found for any items signaling any difference between state and private university instructors' assessment methods of reading ($p>0,05$).

Regarding writing assessment methods, overall results indicate that English language instructors preferred to use student-conducted tests of learner portfolio/e-portfolio in the first place (57,1%). However, state university instructors preferred to use learner portfolio/e-portfolio more frequently (65,9%) than private university instructors (42,3%), and this difference is found to be significantly different ($p=0,05$). Other than the use of learner portfolio/e-portfolio, private university instructors informed that they use instructor-made tests (50,0%) and non-instructor developed standardized writing tests (50,0%) as primary methods of writing assessment. Instructor-made tests are also the second preference of state university instructors (51,4%) following student-conducted tests of learner portfolio/e-portfolio.

Among the assessment methods of listening and speaking, the leading method emerges as student-conducted tests of oral presentations (58,6%). It is the top-ranked assessment method among state (59,1%) and private university instructors (57,7%). An agreement among all instructors was observed for the second-ranked method of student-conducted tests of oral discussion with each student (52,9%). However, in the third place, there is a difference found between state and private university instructors. That is, state university instructors stated that they preferred instructor-made test (51,4%) which private university instructors did not use as frequently as non-instructor developed standardized speaking tests (50,0) and non-instructor developed standardized listening tests (46,2%). This difference between state and private university instructors for the instructor-made tests is significantly different ($p<0,05$).

4. Discussion

Along with the understanding that individuals construct their own meaning in the learning process, more individualized modes of teaching has been observed in educational practice. Based on the premises of constructive theory, formative assessment which provides judgments to the learning of students in the process has therefore been popular in education (Assessment Reform Group, 1999; McDowell et al., 2011). However, the reflection of formative assessment in the classes comes out with some problems especially in the settings where high-stake testing dominates the educational policy. Therefore, this study aimed to examine the practice of formative assessment in classes in the setting of schools of foreign languages where proficiency test has been regarded as high-stake by the stakeholders. The examination was conducted from the view point of practitioners as their construction for the meaning of formative assessment was assumed to influence their practices.

The two-fold intent of the study was to describe current assessment practices of English language instructors in the first place and to investigate any contextual difference between instructors working at state and private universities. A multi-dimensional description was conducted by exploring assessment purposes, assessment procedures and assessment methods of instructors. The results indicated uniformity among instructors regarding these multi-dimensions of assessment; however,

some complexities are raising a need to question to what extent these practices of instructors align with the idea of formative classroom assessment.

First of all, it is found out that instructors in the tertiary level context of the study had student-centered purposes of assessment as their prior purposes; however, when further investigated, the instructional purpose of diagnosing effectiveness of their teaching was also indicated as a goal given priority by the instructors. Whether student-centered or instructional, it is observed that the purposes of instructors were diagnostic which is a function of assessment (Black and William, 1998); yet, how instructors incorporate this diagnostic data into their practices to move their students' learning further was undermined by focusing on instructors' feedback. The quality of feedback which feeds forward the development of learners is a keystone concept in formative teacher assessment (Sadler, 1998). The results signal that verbal and written feedback are the most preferred channels of feedback during the course, and written comments along with total test score are the most preferred ones among instructors for their final report regarding students' performances. No matter the source of feedback, the instructors cumulatively informed that they provided feedback to their students; however, the information gathered for the content of the feedback did not go beyond from being focused on the manner of feedback. On the contrary to the supposed traces of formative assessment such as constructive feedback and/or process-based feedback, instructors frequently highlighted the emotions of their students during feedback sessions. Though the undertaken manner is important to produce higher motivation for students (Sadler, 1998), it is noteworthy to observe such a scarcity for the constructive side of feedback among instructors in the context of the present study.

Furthermore, in assessment procedure dimension, instructors reported that they frequently used print sources in their assessment practices such as items from textbooks and items prepared by the test offices of the institutions. The inclusion of items prepared by the test offices of the institutions either signals the disposition of instructors to assess their students in the format of exit (proficiency) test (Cheng et al., 2008) or the organizational influence on instructors (Yin, 2010) in determining the format of their classroom assessment practices. On the other hand, the instructors might be inclined to use ready-made assessment modes so as to minimize their workload burden (Carless, 2007). The reasons for the instructors' disposition of adopting ready-made items might vary; still, it is a direct indicator of formal formative assessment (Carless, 2007) if not summative assessment.

Another result of the study shows that productive skills especially speaking were neglected in-class assessment of the instructors probably as a result of grammar dominant exit tests. As assessment methods, student-oriented methods were deployed in writing, listening and speaking while instructor-made methods were visible in all skills. The dominance of grammar assessment and instructor-made methods in assessment evidently reflects washback effect (Messick, 1998) of external test in this context.

In sum, the dominance of high stakes testing was observed in the setting under investigation. Therefore, it may be discussed that the formative assessment that the instructors incorporate is not so much formative as it means in theory. Moreover, the cognition of teachers might hinder their acts to be formative as they believe that they are preparing students to the proficiency exam in the end. The lack of theoretical base undermining the acts of instructors in the classroom might be another concern to be discussed (Torrance & Pryor 1998).

5. Conclusions

Contrary to the findings in the literature, there is no significant difference found between the different context of state and private university context for most of the dimensions of assessment

under investigation (Cheng et al. 2004; Davison 2004). Hence, it is concluded that organizational context does not differ much for the instructors working in the same country of which tertiary education is mandated by a higher institution because the differences found in the literature are comparative investigations among countries (Cheng et al. 2004; Davison 2004). Besides, the redundancy of similarities among instructors contradicts with the current literature asserting individual differences among teachers about assessment practices (Davison 2004; Torrance & Pryor 1998; Yin 2010). Although this case sounds as desired uniformity among instructors for the conceptualization formative assessment (Leung 2004), there are complexities about how much these assessment practices of instructors align with formative assessment. It is probable for instructors to act on their past experiences of summative assessment which could not be altered through professional education they have been exposed to (Borg, 2003; Yin 2010).

6. Pedagogical Implications and Suggestions

The first pedagogical implication of this study is the integration of practical formative assessment courses (along with theoretical courses) in pre-service teacher education curriculum. For instance, specific tasks which focus on formative classroom assessment might be added to the tasks of the field experience course that pre-service teachers need to attend. Finally, both theoretical and practical in-service training programs, especially on formative assessment should be provided for teachers in field so as to develop their assessment practices. The viewpoint towards assessment should be changed into its unity with teaching practices (Lantolf, 2011); therefore, the unified understanding of theory and practice should be incorporated both in pre-service and in-service teacher education programs.

In order to build such curriculum for pre-service and in-service teacher education program, further studies might be conducted to determine specific needs of teachers regarding formative assessment. Further, teachers' underlying cognitions might be studied with appropriate methodologies in order to have a better idea about their understanding of formative assessment in their belief systems.

Acknowledgements

This study was funded by Cukurova University Research Fund under the Project Number SDK-2016-7468. We also would like to thank to English language instructors who participated in this study.

References

- Assessment Reform Group (1999). *Assessment for learning: Beyond the black box*. Cambridge, England: University of Cambridge.
- Black, P., and Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5, 7–74.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109.
- Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44, 57–66.
- Cheng, L., Rogers, T. and Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21(3), 360–389.

- Cheng, L., Rogers, W. T., and Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education*, 33(1), 9-32. doi: 10.1080/02602930601122555
- Council of Higher Higher Education (2008). *Regulation on principles to obey in foreign language education in higher education and teaching in foreign language*. Retrieved from: http://www.yok.gov.tr/web/guest/icerik/-/journal_content/56_INSTANCE_rEHF8BIsfYRx/10279/18036
- Creswell, J. W. (2009). *Research Design Qualitative Quantitative And Mixed Method Approaches*. California, USA: Sage Publications.
- Creswell, J. W. (2014). *Research Design Qualitative Quantitative And Mixed Method Approaches..* California, USA: Sage Publications.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21(3), 305–334.
- Lantolf, J. P. (2011). The sociocultural approach to second language acquisition: sociocultural theory second language acquisition, and artificial L2 development. In D. Atkinson (Ed). *Alternative approaches to second language acquisition* (pp. 24-47). New York, USA: Routledge.
- Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice and change. *Language Assessment Quarterly*, 1(1), 19–4.
- McDowell, L., Wakelin, D., Montgomery, C., and King, S. (2011). Does assessment for learning make a difference? The development of a questionnaire to explore the student response. *Assessment & Evaluation in Higher Education*, 36 (7), 749-765. doi: 10.1080/02602938.2010.488792
- Messick, S. (1996). Validity and washback in language testing. *ETS Research Report Series*, 1, 1-18.
- Morgan, G.A., Leech, N.L., Gloeckner, G.W., and Barrett, K. C. (2011). *IBM SPSS for Introductory Statistics*. New York: Routledge.
- Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, 21(3), 249–258.
- Sadler, D.R. (1998). Formative assessment: Revisiting the territory. *Assessment & Evaluation in Higher Education*, 5(1), 77–84.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Teasdale, A., and Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? *Language Testing*, 17(2), 163–184.
- Torrance, H., and Pryor, J. (1998). *Investigating formative assessment; teaching, learning and assessment in the classroom*. Buckingham: Open University Press.
- Torrance, H., and Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27(5), 615-631.
- Yin, M. (2010). “Understanding classroom language assessment through teacher thinking research.” *Language Assessment Quarterly*, 7(2), 175-194.

Appendix A. Assessment Practices Questionnaire

Dear Participant;

This questionnaire includes questions about your classroom assessment practices. It is a part of a study on assessment practices of English language instructors. Your responses will be kept confidential and will not be used for any other purposes. Thank you very much for your time and cooperation.

Personal/Professional Information

Your Sex:

- a. Male b. Female

Your Age:

Your Teaching Experience:

- a. 0-5 years b. 6-10 years c. 11-15 years d. 16-20 years e. 21 and more

The Type of Your Institution:

- a. state university b. private university

Your Educational Background. What is the department you have graduated?:

- a. English Language Teaching Department
b. English Language and Literature Department
c. English Linguistics
d. American Culture and Literature
e. English Translation and Interpreting
f. Other (please specify):

Your Educational Degree:

- a. B.A. b. M.A. c. Ph.D. d. Post Doc

You are teaching at: (multiple choices are accepted)

- a. A1-A2 Level c. B1 -B2 Level c. C1 -C2 Level

Your average size of class includes students.

Your weekly work load is hours.

Have you ever taken a training on assessment?

- a. Yes b. No

If yes, could you describe your assessment training in brief.

Please answer the questions below.

1. What is the most frequently assessed dimension of language in your classroom assessment practices?

- a. reading b. writing c. listening d. speaking e. grammar f. vocabulary

2. What is the least frequently assessed dimension of language in your classroom assessment practices?

- a. reading b. writing c. listening d. speaking e. grammar f. vocabulary

3. What is the purpose of your classroom assessment practices? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. To group my students for instruction purposes in my class
b. To gain information about my students' progress in the targeted domain
c. To plan my instruction
d. To diagnose strengths and weaknesses in my own teaching and instruction
e. To give feedback to my students as they progress through the course

- f. To motivate my students to learn
- g. To make my students work harder
- h. To prepare my students for standardized tests they will need to take in the future (e.g. Proficiency Test, Test of English as a Foreign Language (TOEFL), International English Language Testing System (IELTS), YDS, e-YDS)
- i. To diagnose strengths and weaknesses in my students
- j. To formally record growth in learning
- k. To determine the final grades for my students
- l. To provide information to central administration (e.g. school, administration, higher education institution (YOK))
- m. None
- n. Other (please specify):

4. What are the methods you use to assess your students in READING skill? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. Read aloud/dictation
- b. Oral interviews/questioning
- c. Teacher made tests that may involve different types of items (e.g. true/false questions, fill in the blanks, comprehension questions)
- d. Student summaries of what is read
- e. Learner journal
- f. Learner portfolio/e-portfolio
- g. Peer assessment
- h. Self assessment
- i. Standardized reading tests
- j. None
- k. Other (please specify):

5. What are the methods you use to assess your students in WRITING skill? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. Teacher made tests that may involve different types of items (e.g. true/false questions, matching questions, paragraph/essay writing)
- b. Learner journal
- c. Peer assessment
- d. Self assessment
- e. Learner portfolio/e-portfolio
- f. Standardized writing tests
- g. None
- h. Other (please specify):

6. What are the methods you use to assess your students in LISTENING and SPEAKING skills? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. Oral reading/dictation
- b. Oral interviews/dialogues
- c. Oral discussion with each student
- d. Oral presentations
- e. Public speaking
- f. Teacher made tests that may involve different types of tasks (e.g. taking notes, retelling)
- g. Peer assessment
- h. Self assessment
- i. Standardized speaking test
- j. Standardized listening test
- k. None
- l. Other (please specify):

7. Which of the following represents your primary source(s) for test items and other assessment procedures? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. Items written by myself
- b. Items prepared together with other teachers
- c. Items from textbooks
- d. Items from mandated syllabuses/curricula proposed by Higher Education Institution (YÖK)
- d. Items prepared by test office of the institution
- e. Items extracted from the Internet
- f. Other published test items
- g. Other (please specify) :

8. When you give feedback to your students DURING THE COURSE, how do you provide that feedback? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. Verbal feedback
- b. Checklist
- c. Written comments
- d. Teaching diary/log
- e. Conference with student
- f. Total test score
- g. A letter grade
- h. Other (please specify):

9. When you give a FINAL REPORT to your students, how do you provide that information? Please choose the options which apply best to your practices. (Multiple choices are accepted)

- a. Checklist
- b. Written comments
- c. Teaching diary/log
- d. Conference with student
- e. Total test score
- f. A letter grade
- g. Other (please specify):

10. Approximately what percentage of total time you spend on assessment during a term? The total time includes preparing for an assessment, collecting assessment information, scoring the responses, and reporting assessment results. Please include time spent both at your university and at home. Please choose ONE of the options which applies best to your practices.

- a. 5%
- b. 10%
- c. 15%
- d. 20%
- e. 30%
- f. 40%
- g. 50%
- h. more than 50

11. In your opinion, what is the importance of classroom assessment practices for?

- a) instruction quality:

b) instructors:

c) learners:

d. institution:

12. In your opinion, which points should a teacher consider while delivering feedback during the course?

13. In your opinion, what is the effect of end-of-year test (e.g. proficiency test) on your classroom assessment practices?

14. Could you describe yourself as an effective assessor? If yes, which qualities make you an effective assessor in your opinion? If no, please specify the reason.

That's the end of the questionnaire. Thank you very much.

İngilizce okutmanlarının sınıf içi ölçme ve değerlendirme uygulamaları

Öz

Öğrencilerin değerlendirilmesi hem öğretme hem de öğrenme süreçlerinde eğitimin önemli bir parçasıdır. Alternatif ölçme yöntemlerinin tanınmasıyla birlikte, öğrenci öğrenmesi üzerine odaklanan sınıf içi ölçme yöntemine ilişkin ilgi artmıştır. Ancak, sonuçlarının öğrencilerinin geleceği ile ilgili önemli olduğu sınavlar neticesinde, öğretmenlerin sınıf içi değerlendirmeleri bu sınavlara dayalı olarak değişmiş ve öğrencilerin gelişimini göz ardı eden bir hal almıştır. Üniversite seviyesinde dil öğretimi bağlamında, yabancı diller yüksekokulları, üniversiteye yeni kayıt yapan öğrencilerin dil yeterliklerini ölçerek bölümlerinde eğitimlerine başlayıp başlayamayacaklarını belirleyen ilk eşik olarak nitelendirilebilir. Dolayısıyla, yüksekokulların uyguladıkları dil yeterliği testlerinin öğrencinin geleceği üzerindeki etkisi oldukça büyüktür ve bu testin dil öğreticilerinin sınıf içi ölçme yöntemlerini etkilediği düşünülebilir. Bu bağlamda, bu çalışma, okutmaların sınıf içi ölçme yöntemlerinin amacını, yöntemini ve prosedürünü belirlemeyi amaçlayan betimleyici bir çalışmadır. Ayrıca devlet ve özel üniversitelerde çalışan okutmaların arasında sınıf içi ölçme uygulamaları bakımından bir fark olup olmadığı araştırılmıştır. Nicel ve nitel verilere dayandırılarak ortaya çıkan sonuçlar, okutmaların çalıştıkları bağlam değişmeksizin sınıf içi ölçme uygulamalarının benzer olduğunu göstermiştir, ancak bu benzerliğin ne ölçüde süreç değerlendirmesi içerdiği araştırılmalıdır. Ayrıca çalışmanın sonuçları, okutmaların ölçmenin teorik ve pratik boyutlarını içeren bir eğitime ihtiyaç duyduğunu göstermiş; bu ihtiyacın da hizmet öncesi ve hizmet sonrası eğitimlerle karşılanması önerilmiştir.

Anahtar sözcükler: İngiliz dili öğrenimi/öğretimi; sınıf içi ölçme; dilde ölçme; dil okutmanları; süreç değerlendirmesi

AUTHOR BIODATA

Meryem Özdemir-Yılmaz is a PhD student in the Department of English Language Teaching and lecturer at Çukurova University, Adana, Turkey. Her research interests are language assessment, dynamic assessment, in-service and pre-service teacher education.

Yonca Özkan is an associate professor in the Department of English Language Teaching at Cukurova University, Adana, Turkey. She teaches undergraduate and graduate course in teacher education and second/foreign language teaching methodology. Her research focuses on pre-service language teacher education, technology integration into language teacher education, and language assessment.