Assessment practices in some language centers in public Universities of Ecuador and student proficiency outcomes*

Prácticas de evaluación de algunos centros de idiomas en universidades públicas del ecuador y resultados de competencia estudiantil

Evelyn Almeida** , Diana Castillo*** , Andrea Rosero****

ABSTRACT
This research paper describes the variety of the evaluation process in three language centers of public universities in Ecuador. The main objective is to explore the evaluation practices that are used, their effectiveness in the teaching-learning process and the results obtained in English language learning. Researchers conduct interviews with teachers, principals, and students. In addition, evaluation instruments such as tests and other evaluation materials were analyzed. The study reveals that the students who achieved the highest scores are those who study in a language center which uses a more structured evaluation system. Therefore, the data analysis shows that there is a close relationship between student performance and the evaluation systems applied in the different language centers.

RESUMEN
Este trabajo investigativo describe la diversidad en el proceso de evaluación en tres centros de idiomas de universidades públicas del Ecuador. El principal objetivo es explorar las prácticas de eva-

Keywords: alternative assessment, assessment system, language centers, standardized tests.

Palabras clave: centros de idiomas, evaluación

---

* Este artículo corresponde a una amplia línea de investigación realizada por las autoras.
** Ecuadorian Ph.D. (c) Language, Literacy, & Sociocultural Studies, Universidad Central del Ecuador, Universidad de las Fuerzas Armadas ESPE. Quito, Ecuador. evalmeida@uce.edu.ec
*** Ecuadorian. MSc. Didactics and Linguistics Applied to Foreign Languages Ecuadorian, Universidad Central del Ecuador. Quito, Ecuador. dacastillos@uce.edu.ec
**** Ecuadorian. MSc. Didactics and Linguistics Applied to Foreign Languages, Universidad Central de Ecuador. Quito, Ecuador. aarosero@uce.edu.ec
luación que se utilizan, su efectividad en el proceso de enseñanza aprendizaje y los resultados obtenidos en cuanto al aprendizaje del idioma inglés. Los investigadores realizan entrevistas a maestros, directores y estudiantes. Además, se analizaron instrumentos de evaluación como las pruebas y otros materiales de evaluación. El estudio revela que los estudiantes que lograron los puntajes más altos son los que estudian en un centro de idiomas que utiliza un sistema de evaluación más estructurado. Por tanto, el análisis de datos muestra que existe una estrecha relación entre el rendimiento del alumno y los sistemas de evaluación aplicados en los diferentes centros de idiomas.
Introduction

The need for information about the process of English language assessment applied in Ecuador has great interest because in some language centers it is common to see many students who learn English well enough to pass a proficiency test, but they are often unable to speak that foreign language fluently. The present study consists of an analysis of the English language assessment conducted in three language centers of public universities in Ecuador. Students had to take a mock Preliminary English Test (PET) to assess their English performance. The study includes interviews with students, teachers, and coordinators of each center. Finally, the written assessments provided by the language centers are analyzed.

The key questions that the researchers want to answer are: How are students promoted from one level to another? What kind of assessment process is conducted? Is the kind of assessment appropriated?

Literature

The assessment of foreign language proficiency is directly related to several factors such as the quality of teaching, the learning environment, resources like books, journals, type of tests applied, as well as the learners’ needs and desires, attitudes, aptitudes, and learning styles. One cannot only consider assessment as a number or measurement taken from a test neither be limited to assess just what is easy to assess, “the aim of assessment is primarily to educate and improve student performance” (Wiggins, 1998, p. 7). Assessment should provide feedback about students’ difficulties, achievements, and progress. It has close relation with instruction and learning outcomes therefore, “it is a strong contribution that benefits all the parts” (Hughes, 2003, p. 8) but they shouldn’t be used just as a measurement tool. Hamps-Lyons (2016) argued that language tests “are different in that they measure general language command and make no assumption about a relationship with the content of a teaching sequence” (p. 14). In short, tests are a measurement tool that only provide quantitative data, they assign scores but don’t identify or analyze the data like assessment does.

There are also essential aspects in the development of a foreign language learner to consider in the moment of assessing. According
to Wiggins (1998), assessment is of no value unless it is instructive to students and shows them what, how and when is necessary to change to achieve the goals. Moreover, the learner’s L1 plays an important role when assessing students. The cultural or linguistic background of the learner being tested is greatly influenced by his L1. How test-takers understand and interpret the test items can be influenced by their values, beliefs, experiences, teaching and learning styles, acquired in their L1 (Solano-Flores & Nelson-Barber, 2001). Consequently, culture is definitively a sensitive fact that needs to be taken in account during the assessment process to avoid bias issues. Another fact to consider is that “learners do develop in their proficiency, although few reach similar levels in a foreign language as they attained in their native language, at least under conditions of instruction” (Alderson, 2005, p. 1). This fact seems to be forgotten when assessing students, especially when they have to take standardized tests that are designed and scored based on native-like proficiency to prove their language competence in the foreign language. For instance, the Common European Framework expects C1 level candidates to produce and understand the target language at a native-like speaker level (Council of Europe, 2001). It must be kept in mind that language testing is a complex task that assesses our knowledge about a specific language but most important of all it “assesses how effectively, as test-takers we are capable of communicating with the rest of the world” (Plakans & Gebril, 2015, p. 1).

According to Kiely (2009) there are two types of language assessment: assessment for learning and assessment as measurement. The goal of assessment for learning is to enhance learning. This kind of assessment is “viewed as an integral part of the process of teaching and the development of learning opportunities” (p. 3). On the other hand, language assessment as measurement aims to determine the level of student or know exactly which content has the student learned. Standardized tests serve the purpose of assessment as measurement. In fact, standardized tests in English as a Foreign Language (EFL) settings are used for different purposes. One of them is to measure how much students have progressed in their competence in English. As Hughes (2003) mentions, British and American universities “need dependable measures of language ability” before accepting students from overseas (p. 4). For instance, examinations such as International English Lan-
language Testing System (IELTS), Test of English as a Foreign Language (TOEFL) and Certificate in Advanced English (CAE) are used “as predictors of the ability to succeed linguistically in the academic environment of an English-medium” (O’Sullivan, 2012, p. 79). Furthermore, many researchers and practitioners see the value in such tests and deem them essential in “diagnosing the students in order to improve learning” (Alderson 2005, p. 1). As Burger and Burger (1994) mention, “a standardized test is used in assessment because this instrument has established objectivity, reliability, and validity as well as ease of administration and scoring” (cited in Abedi, 2010, p. 1). However, assessment should not be based only on standardized tests, “a complete view of what has been achieved should include information from as many sources as possible” (Hughes, 2003, p. 5) including formal tests, and the different pieces of information from all sources should be consistent with each other.

Standardized testing cannot be taken out of the schools; what is suggested is that alternative forms of assessment like checklists, journals, logs, videotapes, audiotapes, folios, conferences, diaries, self-assessments, peer assessments, portfolios, should also be used to gain a fuller picture of student proficiency.

According to Brown and Hudson (1998) alternative assessments are beneficial to learners because they (a) require students to perform, create, produce or do something; (b) use real-world context or simulations; (c) are non-instructive in that they extend the day-to-day classroom activities; (d) allow students to be assessed on what they normally do in class every day; (e) focus on processes as well as products; (f) tap into higher level thinking and problem-solving skills; (g) provide information about both the strengths and weaknesses of students; (h) are multiculturally sensitive when properly administered.

Brown and Hudson (1998) also claim that alternative assessments “provide opportunities for students to demonstrate their abilities to use language for meaningful communication (in ways that are consonant with the particular curriculum in which they are studying)” (p. 655). Additionally, alternative assessment allows not just teachers but also students evaluate whether the objectives set at the beginning of the academic term have been fulfilled and if the students have actually acquired language skills.
Four types of assessment are commonly applied in the educational system: diagnostic, formative, summative and integrative (Jennings et al., 2013; Plakans & Gebril, 2015). When choosing a type of assessment for students it is important to remember that each of these types of assessments has different purposes, as will be explained below. A combination of them can give teachers a more accurate picture of student proficiency.

Diagnostic assessment for instance may offer the learners and teachers an insight into the students’ necessities and goals pertaining to a particular module or level (Jennings et al., 2013).

Summative assessment evaluates knowledge at the end of a term usually through a final cumulative exam (Plakans & Gebril, 2015). There is a great range of standardized language proficiency tests like the (IELTS) administrated by the British Council, TOFEL, or the tests provided by the Cambridge English Language Assessment like the one used in this research the (PET) commonly used to full fill this purpose. “Summative assessment, undermines the proper risk-taking and tolerance of ambiguity that formative assessment tries to develop. It takes courage to stick with complexity and forego the temptation of the short-term and superficial” (Blanchard, 2009, p. 89). Determination and brave attitudes are not highly prized to score well on summative assessment. Summative assessment seems to encourage students to look for the simple but safe answer instead of taking risks with complex answers.

Formative assessment is used to assess learning. It provides “the ability of the learner to be able to respond to their assessment feedback in a positive form, preparing them for the next stage/phase of their program” (Jennings et al., 2013, p. 7). Formative assessment also, informs progress that teachers and students make throughout the learning process. Since the aim of this assessment is not to report results, it is assumed that it gives students more opportunity to experiment and take risks with more complex language. Formative assessment can be done by teachers or pupils. It takes place along the school year, it can assess “the quality of teaching, resources, learning environment ... attitudes, motivation, and readiness” (Blanchard, 2009, p. 1). As Plakans and Gebril (2015) mention, the goal of formative assessment “is to improve learning, rather than to judge individual achievement” (p. 25). Fur-
thermore, as Heritage (2020) mentions formative assessment requires constant professional learning from teachers so they could come up with a great deal of tools to support this type of assessment.

Integrative assessment “enables the learner to embrace their reflexive nature and captures the capabilities associated with lifelong learning skills” (Jennings et al., 2013, p. 7). While summative assessment is related to tests and exams, according to Harris and McCann (1994), informal assessment (typically formative assessment) is a way to collect information on the development of students in regular English Foreign Language classes. Informal assessment is performed without exams and usually relates to continuous assessment throughout the academic period. Instinctively, teachers assess their students in identifying their skill level for each English language skill: speaking, listening, reading, writing, and use of English. This informal, systematic way to assess students is highly effective. But, it should be kept in mind that habit formation of students affects the perception and response to formative assessment (Black & Wiliam, 1998). Therefore, it is very important to think about the overall educational development of the students in terms of language and attitudes toward learning language. Likewise, it is necessary to consider the ability of the students to take responsibility for organizing their own learning.

Whether it is diagnostic, summative, formative, or integrative assessment teachers must have clear that in general language testing needs to focus on issues relevant to learning and achievement of the class objectives and not just statistics or measurement. Going in the same path, when assessing students’ language proficiency, there are also different tools or types of language tests like diagnostic tests, progress tests or summative tests to choose from. According to Alderson (2005) a diagnostic test is “designed to identify both strengths and weaknesses in a learner’s knowledge and use of language” (p. 255) and this type of test is given at the beginning of the academic term to provide a baseline measurement of student English language proficiency.

The progress test is another type of language test which implies a continuous process used as a mean to assess the learning progress of the student and as a requirement of student evaluation requirements.
A progress test can be carried out periodically depending on the planning and teaching. It can be done after each class or after each learning objective and can be revised after each unit.

The final assessment involves the collection and evaluation of the entire learning process that is, all the contents seen in the course or level are tested and it is usually performed upon completion of a summative test. Its purpose is to ensure student success and to then promote students to the next level of English.

The literature discussed so far, allows readers to see that there are some other techniques or alternatives in assessment for assessing English language proficiency apart from standardized testing. For instance, Abedi (2010) discusses in detail the advantages of using performance assessment instead of standardized achievement test. For this author, despite standardized achievement tests now are more manageable for English language learners, “the outcomes of these assessments may not be useful in evaluating student learning and informing instruction” (p. 5). Furthermore, these tests do not offer students opportunities to present a more comprehensive picture of what they know and are able to do. The good news is that this gap can be fulfilled by performance test because they engage the students, give them a chance to demonstrate their knowledge and show information on students’ academic needs (Abedi, 2010).

Assessment must motivate students and this is better achieved by performance assessment. In this sense Darling-Hammond (2006) indicates performance assessments that require students to “write extensively, and demonstrate their learning in projects, papers, and exhibitions” (p. 655) can highly motivate students to achieve great levels of learning. Rising motivation in English learners is important because they need encouragement and support in their academic activities. On Abedi (2010) words, when designing performance tests, linguistic modifications need to be made, especially if we are dealing with English language students that are struggling with learning, this will help ensure a valid and fair assessment, and not to discourage the learners.

Teachers can make use of linguistic modifications to reduce the length of the task and make it easier to understand. For example, use more familiar words and eliminate grammatical complexities. Linguis-
tic modifications make the test more appealing to students, “without altering the knowledge and skills being measured” (Abedi, 2010, p. 17).

According to Blanchard (2009) a productive assessment offers learners opportunity to understand how to judge what they can achieve, understand how to take their learning further, help them become self-confident, and trust in their own learning abilities. Assessment is not only a way to check the students’ progress, but also to verify that the objectives proposed at the beginning of the cycle, semester, or term have been accomplished. It is also a way to assess what teachers have done correctly so they may continue doing what works and to determine what needs improvement. As Blanchard (2009) explains, “assessment interprets the evidence. How it is done has a powerful effect on learners and how they feel about learning” (p. 94). Assessments are not a way to discipline or penalize learners or even a way to castigate students for their means of studying. Rather, assessment of learning is done to gauge what teaching methods and learning strategies are working for the student and which methods and strategies need to be revised or ended.

Teachers must also embrace being interactive in order to “develop flexibility in their responses and methods. They tend not to stick rigidly to systems, but change their ways of handling issues that arise” (Blanchard, 2009, p. 37). Consistency and flexibility are key components to assessment conducted during the language learning process. (The ACER Centre for Global Education Monitoring, 2017).

It is important to emphasize that assessment is not a matter of developing and administrating a big test to students at the end of the cycle, but rather offering them opportunities to improve themselves with other types of assessments and appropriate feedback as well. However, this is not a call to reduce the challenges in the tasks. Instead, a diversification of the degree of structuring of the assessment and more authentic forms of assessment should be part of the assessment process. In the specific case of the English language, assessment of all language skills could be done through a written examination, a presentation, a dramatization of an oral examination, etc. To consider the scope of the assessment teachers must consider the complexity and type of assessment being performed, the available resources, time constraints for student testing, and the qualifications of the teachers.
Methodology

The present research was carried out in three public universities in Quito, Ecuador: Universidad Central del Ecuador (IAI), Centro de Educación Continua-Escuela Politécnica Nacional (CEC) and Universidad de las Fuerzas Armadas (ESPE). This study uses a mixed method approach, combining elements of qualitative research such as in-depth interviews and analysis of artefacts, and elements of quantitative research such as tests, quizzes, and evaluation sheets. The methods of collecting data consist of interviewing the principal, one teacher, and one student from each of the language centers. The questionnaire applied to the students, teachers and coordinators has a semi-structured question format with open-ended questions. The rationale behind the open-ended questionnaire was to generate qualitative data to support the quantitative data which was elicited from the test. The questionnaire was validated through expert-validation, and then reworded and irrelevant items were taken out. The qualitative data was analyzed using themes that were relevant to the research questions and presented in charts.

In Ecuador, the higher education system, requests their students to show knowledge of B1 level in a second language before the end of their careers. In the case of the three institutions being analyzed in this paper, their 6th level of English is the equivalent to B1. Therefore, all the students who took the 6th level of EFL in the three language Centers were considered to collect the quantitative data of this research.

Since the total number of the population (967) was actually known, a purposeful sampling strategy was used to establish the number of participants.

Once the sample group was established, the participants took the standardized mock PET (Preliminary English Test) or B1-level-test in relation to the Common European Framework of Reference. This test had three sections namely; reading and writing, listening, and speaking. It consisted of 5 reading tasks (35 questions), 3 writing tasks (7 questions), 4 listening tasks (25 questions) and 4 speaking tasks. The test evaluates the participants’ B1 or low-intermediate proficiency level of English. The test content was checked against the table of spe-
cifications. Good testing condition were provided by giving adequate time, answering all the questions before administering the main test, and preventing them from cheating.

All of the students participating in this study were learning English as their foreign language and their native language was Spanish.

Results
The overall results of the test analyzed with SPSS show that there is a similar performance from all the three language centers, Centro de Educación Continua CEC (A), Universidad de las Fuerzas Armadas ESPE (B) and, Universidad Central del Ecuador IAI (C). However, when reading in depth, data shows that students with the highest scores are in center A with 59,5 % average grade (Figure 1 & Table 1), while the Language Center B obtained a 51,8% and the Language Center C 47,5%.

Analysis of the data revealed that none of the centers reached a high level of performance of the international standards, since all of them are reporting below the average academic performance necessary for the students to pass the level. On Table 3 it is seen that all three centers require students to pass with 70% of the grade. However, since this research analyzed just the assessment process that is going on in the three centers, additional analysis would be required to further explore the reasons why the students performed so low on the PET exam.

Table 1
Chart of frequencies from the three language centers

<table>
<thead>
<tr>
<th></th>
<th>Total B</th>
<th>Total A</th>
<th>Total C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>51,82</td>
<td>59,52</td>
<td>47,47</td>
</tr>
<tr>
<td>Median</td>
<td>51,00</td>
<td>64,00</td>
<td>46,00</td>
</tr>
<tr>
<td>Mode</td>
<td>41,00</td>
<td>64,00</td>
<td>46,00</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>13,66</td>
<td>17,02</td>
<td>6,30</td>
</tr>
<tr>
<td>Variance</td>
<td>186,696</td>
<td>289,988</td>
<td>39,80</td>
</tr>
<tr>
<td>Sum</td>
<td>1192,00</td>
<td>1369,00</td>
<td>1092,00</td>
</tr>
</tbody>
</table>

a. There are different modes. The lower value is taken

Source: Self-made.
Nevertheless, this does not mean that none of the students passed the PET exam. If the mean and standard deviation obtained is analyzed in this study (Figure 2), it can be clearly seen there is a linear relationship between the mean and the standard deviation of the scores obtained by the students.
The analysis of the mean and standard deviation shows that the three centers are different, in fact, students from center A got better scores, since the means is higher (59.52%). However, this center had the largerest standard deviation (17.2). This means that the range and spread of grades scored by the students on the tests were wider.

It could also be seen that, although, students from the center C had the lowest average, with a mean of (47.47%) this center had the most consistent grades, since the standard deviation is (6.3). This data shows that the quantitative data is totally coherent with the following founding’s gathered out of the qualitative data.

After collecting the written tests and assessment evidence from the EFL teachers from the three language centers, documents were analyzed using as parameters: the number of evaluations by level, the content of the different tests, and how each English language skill (use of language, speaking, writing, listening, and reading) was evaluated in the different institutions.

The results portray a variety of patterns that show large differences on the assessment process of the three centers (Table 2 and 3). It is clear that the center A which scored higher on the PET exam, has already established the number of evaluations per level, the content of the instruments, and the frequency for them to be applied. Members of center A integrate the teaching content and the assessment process, it has clear parameters already set to evaluate the oral productions, and an assessment committee has been assigned. This center has made a great effort to establish a procedure regarding assessment policies. Out of these findings, it can be concluded that having clear evaluations policies and procedures can help in attaining better results in the tests.

On the other hand, it is seen that the center that scored the lowest in the standardized test has an ambiguous assessment process which depends mainly on the teacher in charge of the class. The data analysis also showed that teachers applied a different number of evaluations and most of them were multiple choice formats.
Table 2

*Analysis of artifacts (written tests and evaluations)*

<table>
<thead>
<tr>
<th>Parameters</th>
<th>C</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of evaluations by level</td>
<td>Number of evaluations depends on the level and teachers. The center is trying to implement standardized tests for middle and end of each level.</td>
<td>The number of evaluations has a maximum of twelve evaluations between unit tests and mid-term and final exams</td>
<td>There are five evaluations within 8 weeks. Assessments are every two weeks.</td>
</tr>
<tr>
<td>Content of the tests</td>
<td>The content of the examination depends on the teacher conducting the tests but considering the content of the textbook.</td>
<td>The unit tests are based on the guidebook (<em>English Result</em>). The end of level tests are based on KET, PET and FCE level.</td>
<td>The examination content is strictly related to the text used. The book used is <em>Interchange</em>.</td>
</tr>
<tr>
<td>Use of the language</td>
<td>Emphasis on grammatical activities depending on the contents of the textbook. Usually multiple choice.</td>
<td>A variety of activities to evaluate the grammatical part. There are multiple activities to develop. It emphasizes grammatical activities depending on the contents of the textbook. Usually multiple choice.</td>
<td>Evaluations focus mostly on grammar and vocabulary. They are usually multiple choice and open questions.</td>
</tr>
<tr>
<td>Writing skills</td>
<td>There is a written assessment in the mid-level and one at the end of level.</td>
<td>There is written assessment in the mid-level and one at the end of level. This consists of one question, which has to be developed.</td>
<td>The written portion of the assessment depends on the level. The extension of the question can be a paragraph or an essay.</td>
</tr>
</tbody>
</table>
### Speaking
There is no parameter set of how oral assessments and the students. Much depends on the discretion of the teacher. Some teachers make students prepare exhibitions, reading reports, or interviews. Speaking assessments are performed by oral presentations. The number of minutes per presentation varies depending on level. Two English teachers who are teaching at the same level evaluate the evaluation. After each presentation, the teachers ask questions on the subject to describe the understanding of language.

### Reading comprehension
The assessment depends on the level of the students. There is one activity to assess reading comprehension. The assessment depends on the level. There are 3 or 4 sub reading comprehension activities.

### Listening
There are two activities for listening. These will depend on the level of the students. The assessment takes into account 3 or 4 sub listening skills. The evaluation is multiple choice or open questions.

### Three oral evaluations
Evaluation has 5 components: complexity, student understanding, intelligibility, fluency and content. It has two interviews and projects.

---

Source: Self-made.

As part of this research, the principal, one teacher, and one student from each of the language centers were interviewed. The Table 3 below shows some similarities and differences found out of these interviews regarding the levels of English each center has, the number of hours, the methodology and evaluation system they have. Additionally, the interviewees talked about the opinions about the different types of assessment, the advantages, and disadvantages of assessments.

There are a lot of consistencies between the data provided by the interviewees and the data found on the artifacts analyzed above. It was confirmed that center A is the place with the most structured assessment.
process. Furthermore, it has a great diversification of the assessment instruments, including more authentic situations like written examinations, presentations, and oral examinations. It is also worth mentioning that authorities, teachers and students find more advantages than disadvantages on having a defined assessment process. For instance, all of them mentioned that having a predictable assessment process gives students more opportunities to improve their performance.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>C</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels offered and number of hours</td>
<td>Four levels are offered with each level having 60 class hours. However, students must complete a platform.</td>
<td>Offers eight levels, each level consists of 80 class hours. Students register for one semester with include two levels.</td>
<td>It offers seven levels for obtaining B1. The student has to cover 80 hours per level. Additionally, they offer advanced courses in writing, reading and conversational English.</td>
</tr>
<tr>
<td>Teaching methodology</td>
<td>There is no established methodology for teaching English in the Language Center.</td>
<td>Uses a holistic approach to learning. For teaching English, they use a communicative method.</td>
<td>Communicative Language Teaching, or CLT, is paired with the book Interchange.</td>
</tr>
<tr>
<td>Evaluation system</td>
<td>There is no defined evaluation system for students and for teachers. They are working towards unifying exams</td>
<td>There is an evaluation system. One unit per week is covered which has an oral and written evaluation. At least two weekly grades are needed</td>
<td>There is an evaluation system that is well defined. Scores are weighted according to the degree of complexity of the examination. There is a computer system exclusively for the center for their evaluations, assistance, monitoring the student, the teacher, and the coordinator.</td>
</tr>
<tr>
<td>Supervision of evaluation instruments</td>
<td>The same teacher of each class performs the supervision of the assessment.</td>
<td>There is a coordinator and a supervisor who oversees the assessments that are made. Additionally, there is a group of teachers responsible for developing the assessments each semester.</td>
<td>The academic team coordinator supervises the assessment in the center.</td>
</tr>
<tr>
<td>Parameters to Pass each level</td>
<td>The student must pass with at least 28 points. There are two grades in the system, the first 14 and second 14. There are two tests mid-term and final. The four language skills are assessed.</td>
<td>Students need to meet the parameters of the Common European Framework. The end of level tests are standardized tests depending on KET, PET and FCE level. The students have a minimum score of 14 to pass the semester. That is if the student passes or approves two levels, or otherwise fails two levels.</td>
<td>Students pass with 70% of the grade. Evaluations are strictly related to the content of the book in this case <em>Interchange</em> of Cambridge University Press.</td>
</tr>
<tr>
<td>Grades</td>
<td>It depends on the teacher. Recently the center began giving students a mid-term and a final exam.</td>
<td>Teachers have two grades per unit at least. This will depend on the teacher as well.</td>
<td>There are five evaluations within 8 weeks. Assessments are every two weeks.</td>
</tr>
<tr>
<td>Speaking assessment</td>
<td>Oral evaluations should have a section, but each teacher is free to take oral assessments. The type of evaluation also depends on the teacher such as an exhibit, interview, etc.</td>
<td>A committee of teachers prepares a list of issues in accordance with the student preparation level and the students prepare an oral presentation. Time depends on the level of exposure.</td>
<td>Three oral evaluations. Evaluation has 5 components: complexity, student understanding, intelligibility, fluency and content. It has two interviews and</td>
</tr>
</tbody>
</table>
After introductions, teachers (2) made questions to students. The grade of these exhibitions is agreed between the two teachers appointed. Interviews and projects. There is a change of teachers for evaluations for greater objectivity.

| Responsible for the assessment | Previously each teacher was responsible for making their own assessments; however, now teachers are standardizing at least the midterms and final level. Other tests are at the discretion of the teachers. | There is an evaluation committee who prepared the exams and tests before the start the semester. This committee is responsible for correcting, analyzing and reproducing the new exams. There is a bank of tests if necessary. | The coordination group is responsible for the whole process of evaluation. There is a group to develop tests and quizzes. In each cycle, they change, and new tests are performed. There are four editors, a person who collects exams and two coordinators who view the macro process. |
| Assessment advantages | There is academic freedom and teachers can develop their own tests according to the needs of students. | There are several grades obtained in each unit. This helps students to have more opportunity to improve their performance. There is a continuous and standardized assessment. | The first advantage is the standardization of examinations. The tests are based on the curriculum. It is reflected in the results obtained by students. The evaluation form is a product of several grades. The assessment obtained in each unit helps students to have more opportunity to improve their performance. There is a continuous and standardized assessment. |
Assessment disadvantages

| Assessment disadvantages | There is no defined evaluation system. Tests are filtered. There may be cases where students are evaluated on content not viewed in class. | The performance of students studying at the university runs low and that focus more on their areas of expertise not in English. | As standardized tests, there is a vulnerability for plagiarism. |

Source: Self-made.

Conclusions

Out of the previous analysis, it is possible to conclude that there is a great difference regarding the evaluation system of the three institutions, including the number of evaluations applied to students was also different in each center. Language Center A has a more organized evaluation system including its own computer system for testing purposes which is used to deliver some formative or summative exams and provide immediate feedback to students. Compared to center A evaluation system, the Language Center B has a less complex but useful evaluation system. The Language Center C does not have an evaluation system defined to date.

Out of the evidence provided in this project, there is a direct relationship between the system of evaluation of Language Schools of Public Universities and performance of the students. The more defined and rigorous assessment is, the higher the level of students’ achievement is.

This finding supports the claims made by Struyven, Janssens and Dochy (2002) that, in general, having a well-defined assessment process allows students to know, when, where, what and how they are going to be evaluated. As one of the interviewed students mentioned this helps to lower the anxiety when taking the evaluations.

When considering results, particularly the overall low performance across all the language centers, examining the number of classroom hours can be of great importance. The number of hours students must be exposed to language learning shouldn’t be less than that recommended by the Common European Framework for Languages, which estimates that learners typically take 350 to 400 hours to reach a B1 level (Coun-
cil of Europe, 2001). Center C has the lowest number of levels and the lowest number of hours per level too, which is 60 hours and most of the students are required only four levels, or 240 hours; meanwhile centers B and A have levels of 80 hours and students are required 7 or 8 levels (560 or 640 hours) before taking the B1 exam. These two centers provide many more hours of instruction than that suggested by the Common European Framework, although the proficiency achieved in the test-taking performance is still lacking. The Language Center with the lowest number of instructional hours has the lowest results on the PET exam.

The kind of assessment done in each language center also differs. Center A has a standardized assessment system where the evaluation committee gives 90% of the evaluation. Center B has standardized tests and examinations, but teachers can perform other evaluations it deems necessary. Regarding Center C, it is beginning to standardize tests for mid-term and final.

It was observed that all center managers agree on the fact that for obtaining English proficiency, seven or eight levels of English are needed depending on the hours they are given. However, by institutional decisions, a language center is offering courses of four levels. This might be one of the reasons that has caused the level of students to be lower in comparison to the other language centers.

**Recommendations**

It is recommended that all language universities use an evaluation system to measure knowledge of English language students. Additionally, assessment must be comprised of more authentic forms of assessment, such as portfolios, interviews, journals, project work and self or peer assessment and not just standardized tests. The learning assessment must be strictly related to institutional objectives and language centers in addition to the content of the textbook. The number of hours for each level should be in accordance with the number of hours presented by the Common European Framework.

**References**

scale.stanford.edu/system/files/performance-assessments-english-language-learners.pdf


