BEYOND PAPER-AND-PENCIL TESTS: GOOD ASSESSMENT PRACTICES FOR EFL CLASSES

SANDY T. SOTO / EDER INTRIAGO PALACIOS / JOHNNY VILLAFUERTE HOLGUÍN

Beyond Paper-and-Pencil Tests: Good Assessment Practices for EFL Classes

Sandy T. Soto
Eder Intriago Palacios
Johnny Villafuerte Holguín

Coordinators

César Quezada Abad, Ph.D
Rector

Amarilis Borja Herrera, Ph.D
Vicerrectora Académica

Jhonny Pérez Rodríguez, Ph.D
Vicerrector Administrativo

COORDINACIÓN EDITORIAL

Tomás Fontaines-Ruiz, Ph.D
Director de investigación

Karina Lozano Zambrano, Ing.
Jefe Editor

Elida Rivero Rodríguez, Ph.D
Roberto Aguirre Fernández, Ph.D
Eduardo Tusa Jumbo, Msc.
Irán Rodríguez Delgado, Ms.
Sandy Soto Armijos, M. Sc.
Raquel Tinóco Egas, Msc.
Gissela León García, Mgs.
Sixto Chiliquinga Villacis , Mgs.
Consejo Editorial

Jorge Maza Córdova, Ms.
Fernanda Tusa Jumbo, Ph.D
Karla Ibañez Bustos, Ing.
Comisión de apoyo editorial

Contents

Dedication

To all the people who made the publication of this book possible.

To those EFL teachers who, day by day, give the best of their own for helping their students construct their knowledge and learn English in meaningful and effective ways.

Introduction

Individuals´ learning of a second or foreign language has been traditionally measured with paper-and-pencil tests. Unfortunately, such assessment practice prevents learners from demonstrating the skills gained throughout the teaching-learning processes and thus, their actual ability to use the target language effectively. It also limits learners from receiving positive feedback; which opens doors for them to improve their language skills.

The language teaching field demands that English as Foreign Language (EFL) teachers have a vast knowledge of the fundamental concepts and theories that surround the assessment of EFL learning. It also requires that professionals who teach a foreign language keep up to date with assessment tendencies that go beyond paper-and-pencil tests as is the case of authentic assessments.

Assessment practices that go beyond traditional paper-and-pencil tests provide students with opportunities to be assessed in mental stress-free environments. Teachers who promote this alternative form of assessment prompt learners to perform real-world tasks so that they can demonstrate their capability to apply essential knowledge and skills in creative and meaningful ways. In other words, teachers gain insights about how much students have grasped by

their actual ability to perform in a specific situation instead of the number right or wrong answers they have made on a test.

This book is composed of seven chapters intended to inform pre-service and in-service EFL teachers about good assessment practices that go beyond the bounds of tests that require learners to read questions and respond in writing. The first chapter of this book provides EFL educators with a menu of authentic assessments that can be implemented in their classrooms. It also builds a synopsis of assessment practices in Ecuadorian EFL classes and the educational policies that have been implemented to improve them.

The second chapter of this book digs into the use of role-plays as an alternative to assess students´ oral production. The chapter is built upon the results of two studies on the topic and some research conducted by its author, addressing the causes that affect learners´ willingness to speak English. The third chapter focuses on the assessment of one of the receptive skills in language learning, reading. This chapter offers a compilation of resources for effectively assessing reading comprehension in EFL programs; detailing how these resources intertwine with the reality of EFL settings.

The fourth chapter discusses a set of strategies that have been evaluated by the authors of this section through action research. Based on their experience, the authors explain how such strategies can be used as tools to gain insights, develop reflective practice, and improve students' outcomes as well as the teaching environment.

In the fifth chapter, the readers will learn about the importance of determining and how to diminish students' test anxiety. This chapter also addresses practical authentic assessment tools and scenarios that give language learners anxiety-free opportunities to demonstrate their knowledge.

The sixth and seventh chapters have been devoted to the use of technology for assessing language learners authentically. Chapter six analyzes the change from traditional pen

and paper tests to those that have incorporated technology. The authors review the evolution of Information and Communication Technologies based evaluation and assessment applications for English as foreign language learning and teaching, as well as their advantages and disadvantages, current developments, and future trends for technology-based assessment practices.

Finally, grounded in an action research intervention, chapter seven examines how the use of Literature Circles, Google Apps, and corrective feedback can help students improve learners´ English language level. Each chapter in this book offers EFL teachers with valuable information on good assessment practices. It is expected that the educators who read this work consider the suggestions provided here and implement them in their practice. We are sure that by doing so, these educators will give their students the possibility of being assessed authentically; it is to say, by what they can do instead of by how many items they get right on a test.

06

Chapter

The role of ICT in the evaluation and assessment of English as foreign or as a second language

José Luis Ramírez-Romero; Migdalia Rodríguez-Rosales; Héctor Salazar-Sorcia

Introduction

Computers have been used in language assessment since at least the 1960s. However, it was not until the emergence of personal computers, in the late 1970s and early 1980s, that their widespread incorporation to most educational institutions took place, especially in the developed countries (Godwin-Jones, 2001 & Davis, 1998).

In the specific case of English as foreign language learning and teaching (EFLLT), the incorporation of computers has triggered a change from traditional pen and paper examinations to others that have fully incorporated the latest technological advancements.

José Luis Ramírez-Romero holds a doctoral degree from the University of California, Los Angeles (USA). He is a tenured professor at the University of Sonora (Mexico) and has worked in several universities in Latin America and the USA. He has published numerous articles and book chapters and has coordinated four books in ESL related topics.

Migdalia Rodríguez-Rosales holds a PhD in Second Language Acquisition and Teaching from the University of Arizona. She is a full-time professor at the BA in ELT at the University of Sonora in Mexico where she teaches Applied Linguistics and ESP courses. Her main interests are corpus linguistics and online learning.

Héctor Salazar-Sorcia is a graduate student at the University of Sonora (Mexico) and holds a B.A in English Language Teaching from the same university.

In this chapter, we will analyze that change and its implications for EFL teachers by reviewing the evolution of Information and Communication Technologies based evaluation and assessment applications for EFLLT, as well as their advantages and disadvantages, current main developments, and their future trends[1].

For that purpose, we have structured the chapter in six sections. In the first one, we will present a summary of the evolution of ICT based evaluation and assessment practices. In the second section, we will discuss the main advantages and disadvantages of these type of practices. The third section will focus on technology and language assessment research, while the fourth will portray the main recent developments in technology, open access software and tools currently available online which can be adapted to design and conduct evaluation and assessment in EFL/ESL courses, organized by abilities or specific purposes. In the last sections, based on the most current available information, we will discuss where the future of Language Learning (LL) evaluation seems to go, and finally, we will present our conclusions.

---

[1] Since there is no universally accepted definition of the term "Information and Communication Technologies" (ICT), in this chapter, we will understand ICT as a wide term that encompasses telecommunications (such as telephones and wireless signals), computers, software, and audio-visual systems merged in a unified system to access, store, transmit, and manipulate information. We will use the term "computers" only when referring to this specific device employed without an internet connection.

Evolution of computer and ICT based evaluation & assessment in Language Learning

Even though specific ICT based applications have been developed for evaluation and assessment, most applications have been closely linked to the larger field of ICT in LL, and therefore, have followed the same phases or threads. Warschauer & Healey (1998), two of the authors that have studied the evolution of ICT in LL more systematically, initially categorized such evolution in three phases: behavioral, communicative, and integrative[2].

The behavioral phase, covering the period from 1960 to 1970, basically relied on drill-and-practice computer-based exercises, where the computer was viewed as a mechanical tutor. The best example of this phase was the tutorial system PLATO (Programmed Logic for Automated Teaching Operations) that required a mainframe computer and terminals to run (Warschauer & Healey, 1998, p.57). For Godwin-Jones (2001), two of the best-known early applications of computers in language learning were CALIS, from Duke University, and DASHER, from the University of Iowa. Both applications were designed to support of grammar and vocabulary active drill and practice, rather than formal assessment, and "provided for varied feedback options and recognition/display of partially correct answers" (Godwin-Jones, 2001, p.9).

Between the late 1970s and the early 1980s, the communicative phase took place with the widespread launching of personal computers and among growing criticisms to the behavioral postulates, especially from cognitive theory which argues that learning implies a process of discovery, expression and development. The software developed during this phase encompassed text reconstruction programs and simulations (Warschauer & Healey, 1998, p.57).

[2]For Warschauer (2000) and (Motteram, 2013), there is a specific term that should be used when referring to the use of ICT in Language learning, namely: CALL which stands for computer assisted language learning. In this chapter, we will rather use ICT in LL because the latter term is a wider concept.

The integrative phase can be placed between the late 1980s and early 1990s. It was framed within the socio-cognitive approach, which placed greater emphasis on language use in authentic social contexts. In consonance with this approach, computer applications in Language Learning sought to integrate language skills (listening, speaking, reading, and writing) and technology more fully into the language learning process.

> In integrative approaches, students learn to use a variety of technological tools as an ongoing process of language learning and use, rather than visiting the computer lab on a once a week basis for isolated exercises (whether the exercises be behaviouristic or communicative). If the mainframe was the technology of behaviouristic CALL, and the PC the technology of communicative CALL, the multimedia networked computer is the technology of integrative CALL (Warschauer & Healey, 1998, p.58).

In 2000, Warschauer (2000) renamed the first phase and reconsidered the years for each period, placing the first phase, now "structural" from 1970 to 1980, the second one, communicative, from 1980 to 1990, and the third one, integrative, from 2000 onwards.

These phases, however, should not be understood as fixed ones, but rather as trends that even though were dominant at a given time, currently still coexist, and their concepts and applications are used interchangeably in every day practices.

More recent works from other authors (see for example Hubbart, 2009) predicted at least three major areas of development: social networks, mobile technologies, and virtual worlds. These areas are already part of most people´s everyday lives in developed countries. Thus, we could argue that CALL is in a new phase, and that due to its connection to these technologies, could be called: the virtual-social-mobile one.

Advantages and disadvantages of ICT-based evaluation and assessment

The advantages and disadvantages of ICT-based evaluation and assessment have been extensively discussed and studied by many scholars. In this section, we will summarize some of their main arguments and findings.

Advantages

Many of the advantages associated with the incorporation of ICT to the general field of Language Learning and Teaching can be extended to language evaluation and assessment, such as allowing students to work at their own pace, providing them with tasks appropriate to their own levels and giving them prompt feedback, and using multimedia to present authentic situations of language use (Dunkel, 1999 cited by Chapelle & Voss, 2016, p.3). In the particular case of language evaluation and assessment, other advantages frequently mentioned are facilitating, contextualizing, and enhancing the assessment of linguistic abilities (Winke & Isbell, 2017); conducting individualized analyses of learners' language, errors, and performance that can automatically provide feedback to students. Advantages also include generating reports with detailed information about each student to teachers (Chapelle & Voss, 2016); enhancing language assessment with the integration of automated test analysis, scoring, feedback, as well as conversational agents (Li, Schubeck, & Graesser, 2016); and, constructing assessment instruments that expand the possibilities for student learning beyond the traditional classroom (Chapelle & Voss, 2016, p.7).

In addition, according to the 2017 Horizon Report (Adams Becker et al., 2017) there is a growing interest in using data from learning environments to gather information about learning trajectories since learners' actions reveal their progress.

Furthermore, ICT use multiple modalities to create simulations, or to create or recreate real-life situations that can be used to evaluate students. For instance, learners can easily record their speaking or pronunciation practices and their interactions with other learners; furthermore, the information recorded can be used to assess their progress. They can also display their work online in a blog, and the texts produced can be assessed.

Other advantages linked to more specific applications, are presented below:

- Computer testing: Brown (1992, p.48), states that "the advantages of using computers in language testing can be further subdivided into two categories: testing considerations and human considerations." Among the first ones, he argues that computers are more accurate at scoring selected-response tests and at reporting scores than human beings are; they allow testers to target the specific ability levels of individual students and they can therefore provide more precise estimates of those abilities. Other advantages include the fact that the use of different tests for each student minimizes any practice effects, studying for the test, and cheating; and diagnostic feedback can be provided quickly to each student. Among the human considerations, he mentions that the use of computers allows students to work at their own pace; they can complete computer-run tests in less time than traditional paper-and-pencil tests; and they experience less frustration than on paper-and-pencil tests because they will be working on test items that are appropriate for their own ability levels. For Brown, students may find that Computer Assisted Language Testing (CALT) examinations are less overwhelming than paper-and-pencil tests, because the questions are presented one at a time on the screen rather than in a test booklet with hundreds of test items; and many students like computers and might even enjoy the testing process.

- Online testing: some of the advantages of online testing over traditional paper-and-pencil testing this type of testing are related to automatic grading, making it more efficient since there is rapid correction and feedback, and less expensive. Moreover, multimedia prompts (i.e. videos) can be used giving the test a more 'real' feeling. Another advantage is that online testing can be adaptive, and this can facilitate rapid diagnosis. (García Laborda, 2007, p. 8 cited by Chapelle & Voss, 2016, p.2).

- Online assessment: it can help teachers to more efficiently evaluate surface features such as spelling and grammar, to perform automated scoring, and to offer students individualized feedback on their writing (Godwin-Jones, 2008, cited by Chapelle & Voss, 2016, p.6).

- Game-based assessment: it allows teachers to assess complex problem-solving processes and outcomes in a digital game-based learning environment that can be highly attractive to students (Zourou, 2014, cited by Chapelle & Voss, 2016, p.7).

- Mobile-Assisted Language Assessment: in a study by Chen, Hsu and Doong (2016) it was found that students moderately developed self-regulation through mobile learning and assessment, and some mobile applications provided alternative learning opportunities for them.

Disadvantages

Brown (1992) presented a list of disadvantages associated with the use of ICT or computers, for language testing. In the following paragraphs, we list those that in our opinion, might still be valid nowadays, at least in many schools.

For him, the disadvantages of using computers in language testing can also be divided into two categories: physical considerations and performance considerations. Among the physical considerations, he highlights the following: computer equipment may not always be available and the amount of material that can be presented on a computer screen is still limited. Regarding the performance considera-

tions, he argues that the differences in the degree to which students are familiar with using computers may lead to discrepancies in their performances on computer-assisted or computer-adaptive tests and might cause computer related anxiety.

Additional disadvantages mentioned by other authors are related to specific types of ICT applications. In a study about Mobile-Assisted Language Assessment (MALA), results indicated mixed attitudes from the learners towards MALA regarding fairness and lack of authentic communication (Samaneh & Samaneh, 2016).

Technology and Language Assessment Research

In their review about technology and language assessment research, Chapelle and Voss (2016) identify three main themes: computer-adaptive testing, automated writing evaluation (AWE), and the comparison of computer- assisted language testing (CALT) and non-computerized testing. These themes will be developed briefly in the following sections.

Computer-adaptive language testing

Computer-adaptive language testing refers to the ability of CALT technology to allow for the interaction between the input and the answers given by test takers. Depending on test takers' performance on an item, the computer selects the next item to show, based on an algorithm defined by the test developer. The Educational Testing Service (ETS), a private educational testing and assessment organization that administers international tests including TOEFL, reports the use of item response theory (IRT) to provide for computer-adaptive language testing (Carlson & von Davier, 2013). In general terms, IRT follows a statistical analysis which estimates items' difficulty and other parameters. When the item is tested and a level of difficulty is assigned, it is tagged and used in tests. The advantages of this type of testing include the possibility of personalizing the test to evaluate the learners with items appropriate for their level; consequently,

there are many different versions of the tests, and test time is used more wisely since learners spend time in tasks more tailored to their level.

Computer-adaptive testing is mainly used in formal international proficiency tests. In the classroom day to day scenario, instructors can develop quizzes in LMS platforms such as Edmodo, Moodle, or Blackboard; however, the option of adaptive testing is not yet available. Perhaps an update to include the adaptability bonus would be feasible in years to come.

Automated Writing Evaluation (AWE)

Automated Writing Evaluation (AWE) or Automated Essay Scoring (AES) is technology that has been under development since the 1960s (Chapelle & Voss, 2016). It was conceived not merely to assess second language (SL) or foreign language (FL) learners. This type of evaluation requires discourse analysis and parsing among other features, and it is based on natural language processing (NLP). Currently, TOEFL iBT uses an automated writing evaluation engine called e-rater. This engine evaluates writing quality in terms of grammar, usage, mechanics, discourse structure, among other features. This tool is used to support human rating to assess this section of the test. The ETS also offers Criterion, an online essay evaluation system for schools which also uses e-rater and includes peer review, teacher comments and reports. Reviewers of this tool, Lim and Kahng (2012) consider that although it is objective, it fails to evaluate content and argumentations. In a later study of the use of Criterion, Li, Link and Hegelheimer (2015) reported that draft revisions improved. A similar tool MY Access! developed by Vantage Learning, has been perceived to be useful in the drafting and revising process as well, according to research findings by Chen and Cheng (2008) and Grimes and Warschauer (2010). However, students needed their teacher and classmates' feedback as well. Thus, their recommendation is to integrate AWE with a clear pedagogical design.

Other advocates for AWE are Shermis et al. (2016). For them, a main advantage of AWE is the ability to provide formative assessment in the form of feedback, giving the learners the opportunity to improve their texts. The Intelligent Academic Discourse Evaluator (IADE) (Cotos, 2009) and more recently the CyWrite, are two examples of AWE tools created by researchers from Iowa State University which target non-native English writers. Their interest is not only summative assessment as in standard tests, but also to provide formative assessment and inform research about writing.

Automated Speaking Evaluation (ASE)

Speech recognition tools face the challenge of processing speech that varies greatly from person to person. Providing responses to spoken utterances is even more complex since in addition to processing speech, the software decides upon a possible path to follow. Speech recognition tools are currently considered mainstream and are even accessible from mobile devices (i.e. Siri from apple) and as has happened with other tools that were not designed specifically for language learning, language instructors have used them to provide language learners with L2 practice.

Automated Speaking Recognition (ASR) is another area of constant growth that will continue to develop. For this type of tool, natural language processing (NLP) advances in technology are essential. They include a database of representations of sounds from a great number of native speakers; the computer then compares and recognizes the sounds produced by learners to provide a response and even elicit further communication. Thus, the feedback is in the form of an evaluation of what was said; that is, what the computer understood.

As it was mentioned previously, ETS is a global leader in the administration of proficiency tests such as TOEFL. The TOEFL iBT is the internet based test that assesses speaking as well as reading, listening and writing. ETS's Speech Rater scores spontaneous responses. It is "based on NLP and speech-pro-

cessing algorithms are used to calculate a set of features that define a 'profile' of the speech on a number of linguistic dimensions, including fluency, pronunciation, vocabulary usage, grammatical complexity and prosody." ("Automated Scoring of Speech," 2017). ETS adds that human raters review the scoring to validate it and that they are working on the inclusion of more extensive NLP features to analyze discourse.

Computer-assisted feedback in the form of audiovisual displays of pitch and intonation has been used to provide pronunciation practice. Hincks (2002) summarized research studies about the use of this signal analysis software with positive results. She observes that even though the tool is helpful, feedback provided by instructors is imperative. In this type of pronunciation practice the learners can see the display and intonation of the utterance to be practiced as well as their own, comparing how similar or different they are and then have the possibility of trying it again (i.e. Visi-Pitch).

In addition to the tools developed to help language learners practice their speaking skill before taking a high stakes test such as the one described previously, there are several software applications free of charge that can be used in a classroom setting scenario. A recent study by Li et al. (2017) about the use of the ASR device IVI (iFlytek Voice Input) to improve pronunciation in Chinese English learners yielded positive results. In this study learners would read a text aloud and the app would transcribe it. They could then visualize their pronunciation mistakes. They were then asked to work on the sounds that presented problems and repeat their practice after a week. A drawback that they observe is the fact that the tool lacks an option to check the pronunciation of the words.

Developments in Computer Assisted Language Testing per Language Skill

Nowadays any of the four basic language skills can be assessed using a computer. Technological advancements and innovations have enabled educators to assess all four skills whereas only receptive skills (listening and reading), were feasible in years past.

Speaking Evaluation Tests and Tools

ICT and language assessment have been associated as far back as the 1980s (Suvorov & Hegelheimer, 2014) and ICT usage has only continued to evolve. Regarding speaking assessment, institutions have access to standard language tests that can provide feedback on students' speaking abilities. Some of these tests are offered by renowned institutions in the language testing field and they include, in addition to the already mentioned TOEFL IBT® (Test of English as a Foreign Language Internet-Based Test), the BULATS (Business Language Testing Service) Speaking Test, the BEST™ (Basic English Skills Test), the Versant™ English Test, and the PTE Academic™ (Pearson Test of English) (Suvorov & Hegelheimer, 2014). Nevertheless, these tests have a considerable cost and are not available in most locations, which make them impractical for the average language educator. Fortunately, there are many commercial and open-source tools to evaluate and assess speaking abilities that are available to English teachers everywhere.

Regarding the commercial options, there are some speaking evaluations tools that teachers can use. Some of the most recent options include tools such as the Speech Rater Engine in the TOEFL Online Practice Test by ETS and the iSpraak Online Pronunciation Feedback web application.

As it was mentioned, the Speech Rater Engine is a system to evaluate spoken production of speakers who will be taking a standardized speaking test such as the TOEFL® test. This system´results have shown a close correlation to the results of human-based evaluation of speaking (Bat & Yoon, 2015).

Another commercial option for educators is the iSpraak Online Pronunciation Feedback web application. This application was developed by Dan Nickolai at Saint Louis University to integrate a formative assessment tool with emergent speech recognition and speech synthesis technologies (Adams, et al., 2017). This application works through the Google Chrome web browser and although initial use is free to use, extended use of the application requires a subscription.

Duolingo is a widely known and popular application for language learning. English teachers can use this application to grade students' oral proficiency since the application scores the students' oral skills and pronunciation. Additionally, this application offers the Duolingo English Test. This test offers a trial version which is free, and a paid version that includes a certificate. The Duolingo English Test is relatively new but some authors suggest its usefulness in improving students' oral abilities as well as confidence (De Castro, Da Hora, & Pinto, 2016). Also, according to Ye (2014), it can help students improve their TOEFL scores.

There are other free and open-source applications created for educational purposes that can be used for language learning for assessment purposes. The Voki Ed application allows teachers and students to create talking animated avatars using their own voices. The Voki Ed application does not have speech recognition tools. However, it allows users to create an animated avatar that resembles their appearance, and then to record a spoken message using the recording feature. Voki Ed is free to download, but most of its features require users to buy the full version or to make in-app-purchases. Nevertheless, the advantages of using Voki Ed are many, such as the fact that students may access it through a computer or a handheld device. In addition, it appeals to different personalities, and since students are able to share their videos, they can also get feedback from their classmates (Yona & Marlina, 2014).

The EduSynch platform is a website and application that allows students to practice their language skills in preparation for a TOEFL exam and similar English Language Stan-

dardized tests. Educators can take advantage of its free speaking evaluations and professionally developed interface. In addition, if students so desire, the EduSynch platform also offers oral and written examinations grades by professionals for a fee.

Finally, there are applications that were not originally intended for language teaching, but are experiencing a considerable growth in language teaching and assessment. Among these are web platforms such as VoiceThread, the WeSpeke web application, and speech-coaching applications like Orai.

VoiceThread (Stannard & Basiel, 2013) is one of the platforms with the fastest rising popularity in spoken formative assessment in English language learning. It is cloud-based and thus does not require the installation of any software. It offers educators a platform where they can communicate using video and audio messages as well as text. The interface allows for group interaction and contains an analytics section that enables teachers to monitor the activity of their courses. Unfortunately, VoiceThread is a platform that requires a license to be used and thus it is not an option for educators who do not have financial support from an institution.

A second rising option is the WeSpeke platform. This platform was created with the purpose of language learning and speaker interaction. The WeSpeke interface allows learners to interact with other speakers around the world with written messages and audio recordings and it is completely free to use. It also has courses in which students can enroll to further practice their skills and it is a platform where teachers and students can practice their speaking abilities outside the classroom (Mora, 2016).

Finally, the Orai application is another rising star in speaking practice and assessment. Orai works with different speaking tasks, such as tongue-twisters, repetition, and text reading. The app can be used by both native and non-native English speakers — machine learning is used to interpret a range of accents, and while it is still in its early stages, it has shown promising results (Simon-Lewis, 2017). The app is still recent

(it was launched in April 2017) and it is only available for IOS devices at the moment, but it is likely to migrate to android as well in the near future.

Evaluation Tests and Tools for Writing

The evaluation of written language can be aided by a variety of technologies. In addition to the standardized tests for writing abilities mentioned earlier there are also the COMPASS® ESL Placement Test to evaluate students' essay writing abilities (Suvorov & Hegelheimer, 2014) and the DIALANG test to determine students' approximate writing proficiency according to the Common European Framework of Reference (Winke & Isbell, 2017). Nevertheless, like in the case of speaking evaluation, these tests are available for a fee and not accessible in all locations, which might be impractical for the average language educator.

Another commercial tool to evaluate and assess writing abilities is also available to English teachers. The WriterPlacer by Accuplacer® is offered from The College Board. This tool asks students to write an essay which is then graded by an automated system. This system is used to evaluate students' writing skills in their native language as well as in English.

The Grammarly web add-on is a tool that allows the users to check their grammar, spelling, punctuation, sentence structure and to detect plagiarism. Grammarly is an automatic editing tool like Word from MSOffice. This add-on, although recent, has been reported to be well-received by many higher education students (Cavaleri & Dianati, 2016) and to help in the learning of certain grammar structures such as the passive voice (Qassemzadeh & Soleimani, 2016). Finally, though Grammarly has a free add-on option for users, it is much more limited in use compared to the premium version, since it includes fewer assessment options.

Regarding free and open source options for the evaluation and assessment of writing abilities there are several options as well, including the Paper Rater website, the WhiteSmoke grammar checker, and the Language Tool Proofreading Service.

The Paper Rater website is an online platform where students can submit their written work and have it analyzed. The website checks the text's grammar, spelling, word choice, style, and vocabulary, and it also looks for plagiarized text. The text simply must be copy-pasted into a text box in the website and the feedback is generated automatically. Additionally, the Paper Rater website checks the texts according to the grade of the person submitting the paper. Although Paper Rater is mostly used to check students' writing skills it is also a popular tool to check for plagiarism (Masic, Begic, & Dobraca, 2017).

Another non-commercial alternative to assess writing is the WhiteSmoke grammar checker. This software checks spelling, grammar, punctuation and plagiarism. It is compatible with MS Office Word and it works in several operating systems as well. This site offers a free trial and users can purchase the full version. There is also a web version of the software as well as a mobile application. In addition to the features mentioned earlier, this mobile application seems to improve students' self-assessment (Qazzemzadeh & Soleimani, 2016).

Finally, the Language Tool proofreading service is another free assessment option. Language Tool is an open source option that allows users to identify grammar and style issues such as verb tenses, concordances, use of transitions, spelling errors and word choices. Though this option has more limitations than the other two mentioned above it is the only one that is genuinely free and it is compatible with many open source software options.

Reading and Listening Evaluation Tests and Tools

Reading and listening comprehension evaluation tools are more varied than their counterparts and they have been around for much longer. Educators may find reading and listening comprehension material with just a simple google search. However, good quality and appropriate material is not as easily found. Fortunately, there are several ICT tools

to help assess reading and listening comprehension, and unlike the tools mentioned for speaking and writing, these assessment tools include more non-commercial options. It is also important to note that there are very few commercial ICT tools that focus solely on the receptive skills, since these can be easily integrated into other assessment tools and applications, and anything that a commercial tool offers also can be obtained with a free or open-source ICT tool.

Among commercial ICT tools to evaluate reading and listening skills there are several applications that allow teachers to assess comprehension. The Test Prep Review website, by Mometrix®, is a standardized test repository. It offers educators and students access to a plethora of test simulators such as the Accuplacer test, SAT, Pearson Placer, and COMPASS test among many others. This website offers complete practice tests with answer keys to be used by individual students and as class activities. The only disadvantage is that it does not offer automated review on the site. However, since it provides a great variety of tests, it remains a popular option among educators.

Several non-commercial options to assess and monitor the reading progress of students are available with a simple online search in a web browser. One of them is the website Exam English. Here, the users can find multiple choice exams and obtain feedback about their selections.

Authoring Tools

Software designed for purposes other than language learning has being used to develop exercises to evaluate language learning (i.e. Hot Potatoes). Chapelle and Voss (2017, 2016) argue that technology applications should allow instructors to design their own tests using accessible authoring tools. In other words, these tools should be easy to use, flexible and allow for customization. Some commercial and free authoring tools to consider for the evaluation and assessment of language learners are described in this section.

Quizlet is currently considered one of the most popular web-based and free mobile application that provides the tools to create flashcards, games and quizzes. Interestingly, it was created by Andrew Sutherland when he felt the need to learn vocabulary for a French course. Even though it does not include speech recognition features or automatic writing evaluation tools, users, students or instructors can easily create quizzes and share them for free. Developers also included a listening feature and now users can listen to text as well. Quizlet offers subscriptions for a fee for teachers and students to include other features such as personalized instruction.

Learning management systems (LMS) such as Moodle, Blackboard, Schoology, Edmodo, and the like, have a tool included to create quizzes tailored to the course needs. These tools tend to be comprehensible and include help features to assist users in the design of their assessment. However, if a LMS is not used in the course and instructors and students have access to online resources, instructors can use Google Forms to create a quiz for free. They would have to enter the answer key and use the Flubaroo plugin to store grading in a spreadsheet, send each student their grade, and then be able to see who obtained a low score, among other features.

Where does the future seem to go?

Technology trends have an impact in general in education and consequently in ELT. The 2017 Horizon Report published by the Media Consortium since 2002, presents research about the technological developments that have changed and will continue to change education in a period of five years. According to this report, adaptive learning technologies and mobile learning are already being used. The "internet of things" which consists of enabling devices to transmit information and capture and analyze data (i.e. Apple watch, Fitbit) and Next-Generation LMS which will have to include modern user experience (i.e. mobile, social, gamified, personalized to user's needs) will probably be adopted in two

to three years and artificial intelligence and natural user interfaces (i.e. speech recognition, touchscreen interfaces, eye-tracking) will take four to five years to be integrated into the mainstream. The technologies that, according to Adams Becker et al. (2017), we consider currently have and will continue to have a strong impact on ELT and consequently, on evaluation and assessment, are the following: adaptive learning technologies, mobile learning, Next-Generation LMS, and perhaps, Natural User Interfaces (NUIs).

Adaptive Learning Technologies

Adaptive learning systems, also known as intelligent tutoring systems (ITSs) consider students' learning styles, strengths and weaknesses to propose the learning activities and exercises each student should perform. This way, learning profiles are studied to design and implement specific instructional environments.

An example of this type of technology is EduSynch, a free adaptive training tool for English language proficiency exams such as TOEFL, IELTS, TOEIC, and Cambridge. With the help of this site, students take mock tests and their real-time performance is tracked. This information is then used to take corrective action in the classroom and the type of practice students need outside of the classroom as well, to help them improve their language proficiency scores.

Mobile learning

Mobile technologies, such as mobile smartphones or tablets for language learning and teaching, are becoming more mainstream than emergent. Nowadays apps of all kinds are offered through Google Play and App Store, and anyone with internet access can be connected, which has helped in making them become a necessary product. Many language learning apps can be found. However, these mainly offer vocabulary practice through drilling exercises with immediate evaluation. According to Godwin-Jones (2017), even though these devices offer more advanced communication oppor-

tunities, these are usually not utilized. Mobile assisted language learning (MALL) apps such as Duolingo and Memrise, however, are integrating multimedia and exploiting the connectivity feature by giving the opportunity to practice with a partner, another user from any part of the world who is also interested in learning. Perhaps a way to asses students is by having them reflect on the reaction of their partners; if they were understood or if they had to recast their message. The British Council also offers apps for smartphones (https://learnenglish.britishcouncil.org/en/apps) and will probably redesign them as technology improves. Perhaps the future will lean more towards taking advantage of features that give users the possibility to create their own game or Augmented Reality (AR) projects using online software such as LearnAR and ARIS   (Godwin-Jones, 2016). This gives instructors and students ownership and more control of their own teaching and learning.

   Instagram and Snapchat are apps that can be used in language learning even though their roles are as social media applications (Rosell-Aguilar, 2016). Similarly, Youtube's affordability for uploading and sharing videos from individuals and media corporations has been part of its attraction and thus is also widely used in language teaching. According to Lidsky (2017) the YouTube app has more mobile views than the desktop version. They rely on algorithms that analyze users' preferences. Additionally, they are working on immersive video which allows for Virtual Reality features that have had an impact on education (i.e. Discovery VR YouTube channel). Language learning examples can be seen in the LearnEnglishinVR YouTube channel.

   Smartphones and tablets are becoming more affordable and their features have been improving. It is likely that they will continue to be devices used by language learners. As Godwin-Jones (2017) points out, it is the instructors' duty to take the opportunity to use them to encourage language learning. Game-like applications used currently in these devices provide instant evaluation and designers are constantly improving the features and exploiting technology affordances.

Next-Generation LMS

Learning management systems (LMS) or Virtual Learning Environments (VLEs) such as Moodle, Desire2Learn, Edmodo, Schoology and Blackboard are now commonplace in higher education and more recently in K-12 as well. Through these systems students have access to course material and they are also able to submit assignments, take quizzes, check grades, interact with instructors and peers and so forth. The idea of a New Generation of LMS comes from the need to integrate more flexible features that allow for evolving students' needs and more opportunities for formative assessment, among others. Ideally, these LMS would not only be administrating students' information but also tracking students' behavior and a variety of assessment types to optimize teaching. Acrobatiq is one platform developed by the Carnegie Mellon University's Open Learning Initiative that allows for adaptive learning technology to provide a more individualized learning experience.

Natural User Interfaces

According to the 2017 Horizon Report the constant development of natural user interfaces (NUIs) will impact education. Learners' expectations are constantly evolving and NUIs can provide gesture sensing technology, speech recognition and haptics or the way of applying touch (tactile) sensation and control to interaction with computer applications. According to Underkoffler (2010), an interface designer and inventor who has been consulted for the development of science-fiction movies, NUIs will be included in next generation computers. The NUIs are already being used in smartphones and the potential of having this pocket computer with NUIs capabilities will change the way people have access to resources. Prototypes of educational games designed with NAUs such as Kinect have been tested and proven to be effective (Shapi'I & Ghulam, 2016). In a language learning scenario, automatic speech recognition (ASR) technology is already being used to evaluate pronunciation quality (Neri, Cucchiarini, & Strik, 2003; O'Hear, 2010). Duolingo is an app

already including this feature. NUIs will continue to improve and have the potential of fostering new ways of presenting information, new ways of learning and thus, new ways of evaluating and assessing language learners.

Conclusions

The incorporation of ICT in language assessment has followed a similar path to that of language learning in general, which has meant a growing incorporation of ICT based applications in assessment and evaluation practices. This incorporation has brought many benefits to the field of assessment. One of the most important has been the implementation of more personally oriented practices. So, it is reasonable to assume that, as new technologies emerge, language assessment will also continue to evolve and adopt the new developments to move towards more personalized assessment practices.

For instance, the possibility of carrying pocket-sized computers such as smartphones and tablets with internet access is no longer just a dream and students are now able to use the tools by themselves to create their own study guides (i.e. quizlet) or access online practice exams. In the near future, ICT applications and advances will grow at an even faster pace and will improve the tools currently used in language evaluation.

Even though limitations of automated writing evaluation (AWE) and automatic speech recognition (ASR) technology are still prevalent, since technology is constantly evolving, these tools are expected to become more accurate and reliable. Furthermore, they will be integrated with adaptive learning technology and game-like applications for a more individualized experience that will bring about new ways of learning and evaluating. Thus, the need of more specialized professionals who understand the new roles of technology in language learning and assessment is still prevalent. We hope this chapter contributes to that understanding.

References

Adams Becker, S., Cummins, M., Davis, A., Freeman, A., Hall Giesinger, C., & Ananthanarayanan, V. (2017). NMC Horizon Report: 2017 higher education edition. Austin, Texas: The New Media Consortium.

Automated Scoring of Speech. (2017). Retrieved from

https://www.ets.org/research/topics/as_nlp/speech/

Bat, S. & Yoon, S. (2015). Automatic assessment of syntactic complexity for spontaneous speech scoring. *Speech Communication.* 67, 42-57.

Brown, J. D. (1992). Using computers in language testing. Cross Currents, 19, 92-99.

Carlson, J.E. & von Davier, M. (2013). Item response theory. In ETS R&D Scientific and

Policy Contributions Series. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.2013.tb02335.x/epdf

Cavaleri, M. & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Learning & Language.* 10(1), 223-236.

Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in Language Learning & Technology. Language Learning & Technology, 20(2), 116–128. Retrieved from http://llt.msu.edu/issues/june2016/chapellevoss.pdf

Chapelle, C.A., & Voss, E. (2017) Utilizing technology in language assessment. In: Shohamy, E., Or, I., & May, S. (Eds.), *Language Testing and Assessment. Encyclopedia of Language and Education* (3rd ed.). Cham: Springer.

Chen, Y.L., Hsu, C.C. & Doong, J.L. (2016). Self-regulation assessment of mobile game-based language learning. In G. Chamblee & L. Langub (Eds.), Proceedings of Society for Information Technology & Teacher Education International Conference (pp. 882-886). Savannah, GA, United States: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/p/171788/

Chen, C.-F. E. & Cheng, W.-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology,* (12)2, 94–112. http://dx.doi.org/10125/44145

Cotos, E. (2009). Designing an intelligent discourse evaluation tool: Theoretical, empirical, and technological considerations. In C. A. Chapelle, G. H. Jun, & I. Katz (Eds.), *Developing and evaluating language learning materials* (pp. 103-127). Ames, IA: Iowa State University.

De Castro, A., Da Hora, S., & Pinto, H. (2016). Duolingo: An Experience in Language Teaching. Journal of Educational and Instructional Studies in the World, 6(4) 59-63.

Godwin-Jones, R. (2017). Smartphones and language learning. *Language Learning & Technology,* 21(2), 3–17. https://dx.doi.org/10125/44607

Godwin-Jones, R. (2016). Augmented reality and language learning: From annotated vocabulary to place-based mobile games. *Language Learning & Technology,* 20(3), 9–19. http://dx.doi.org/10125/44475

Godwin-Jones, R (2001). Language testing tools and technologies. *Language Learning & Technology,* 5(2), 8–12. http://dx.doi.org/10125/25121

Grimes, D. & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *Journal of Technology, Learning, and Assessment,* 8(6).

Hincks, R. (2002). Speech recognition for language teaching and evaluating: A study of existing commercial products. Proceedings of ICSLP, 733-736.

Kessler, G. & Hubbard, P. (2017). Language Teacher Education and Technology. In Chapelle, C. & Sauro, S. (Eds.), *The handbook of technology and second language teaching and learning* (PP. 278-292). doi:10.1002/9781118914069.ch19

Li, H., Shubeck, K., & Graesser, A. C. (2016). *Using technology in language assessment: Contemporary second language assessment.* London, UK: Bloomsbury Academic.

Li, J., Link, S. & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing, 27*, pp.1-18.

Li, M., Han, M., Chen, Z., Mo, Y., Chen, X & Liu, X. (2017). Improving English pronunciation via Automatic Speech Recognition Technology. 2017 International Symposium on Educational Technology (ISET), Hong Kong. DOI: 10.1109/ISET.2017.58.

Lidsky, D. (2017). Six ways YouTube is primed for the future (and four areas that need work). [online] Fast Company. Retrieved from https://www.devicedaily.com/pin/six-ways-youtube-is-primed-for-the-future-and-four-areas-that-need-work/

Masic, I., Begic, E., & Dobraca, A. (2017). Plagiarism detection by online solutions. *Studies in health technology and informatics.* 238(n/a), 227-230.

Mora, J. (2016). Workshop: Training for the teachers of English at CEIC, in Alajuela, on web 2.0 tools that can be used with their students in order to promote the use of English outside of the classroom, Focusing on the Oral Production (Speaking Skill). *Revista Ensayos Pedagógicos,* 11(2) 173-201.

Motteram, G. (2013). Developing and extending our understanding of language learning and technology. In Motteram, G. (Ed.), *Innovations in learning technologies for English language teaching* (pp.175-192). London, UK: British Council. Retrieved from http://www.teachingenglish.org.uk/sites/teacheng/files/C607%20Information%20and%20Communication_WEB%20ONLY_FINAL.pdf

Neri, A. Cucchiarini, C. & Strik, W. (2003, Aug). Automatic speech recognition for second language learning: How and why it actually works. In ICPhS-15, pp.1157-1160.

O'Hear, S. (2010). Babbel introduces speech recognition to aid language learning. [online] TechCrunch. Retrieved from https://techcrunch.com/2010/06/23/babbel-introduces-speech-recognition-to-aid-language-learning/

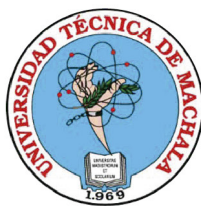Orai App. (2017). We are Orai. Retrieved from https://www.oraiapp.com/about-us/

Qazzemzadeh, A. & Soleimani, H. (2016). The impact of feedback provision by grammarly software and teachers on learning passive structures by Iranian EFL Learners. *Theory and Practice in Language Studies,* 6(9), 1884-1994.

Rosell-Aguilar F. (2017). State of the App: A taxonomy and framework for evaluating language learning mobile applications. *CALICO Journal,* 34(2), pp. 243-258

Samaneh, T. & Samaneh, K. (2016). Mobile-assisted language assessment: assessing speaking. *Computers in Human Behavior,* 64 (November), pp. 409-413.

Shapi'l, A. & Ghulam, S. (2016). Model for educational game using natural user interface. *International Journal of Computer Games Technology,* 2016 (2016). Retrieved from: https://www.hindawi.com/journals/ijcgt/2016/6890351/

Shermis, M.D., Burstein, J., Elliot, N., Miel, S. & Foltz, P.W. (2016). Automated writing evaluation: An expanding body of knowledge. In MacArthur, C. A., Graham, S., Fitzgerald, J. (Eds.) *Handbook of writing research* (Second ed.). NY: The Guilford Press.

Stannard, R. & Basiel, A. (2013). A practice-based exploration of technology enhanced assessment for English language teaching. In Motteram, G. (Ed.) *Innovations in learning technologies for English language teaching.* London: British Council. Retrieved from http://www.teachingenglish.org.uk/sites/teacheng/files/C607%20Information%20and%20Communication_WEB%20ONLY_FINAL.pdf

Suvorov, R., & Hegelheimer, V. (2014). Computer-assisted language testing. In The Companion to Language Assessment. John Wiley & Sons, Inc. DOI: 10.1002/9781118411360.wbcla083

Underkoffler, J. (2010). Pointing to the future of UI. [online] Ted.com. Retrieved from https://www.youtube.com/watch?v=b6YTQJVzwlI

Warschauer, M. (2000) CALL for the 21st Century. IATEFL and ESADE Conference, 2 July 2000, Barcelona, Spain.

Warschauer, M., & Healey, D (1998) Computers and language learning: an overview. *Language Teaching* 3(1/2): 57–71. Retrieved from http://hstrik.ruhosting.nl/wordpress/wp-content/uploads/2013/03/Warschauer-Healey-1998.pdf

Winke, P. M. & Isbell, D. R. (2017). Computer-assisted language assessment, language, education and technology. *Encyclopedia of Language and Education,* pp 1-13.

Ye, F. (2014). Validity, reliability, and concordance of the Duolingo English Test. Retrieved from https://s3.amazonaws.com/duolingo-certifications-data/CorrelationStudy.pdf

Yona, S., & Marlina, L. (2014). The use of voki website in teaching speaking on oral descriptive text for junior high school students. *Journal of English Language Teaching,* 3(1), 235-242.

Redes

Redes es la materialización del diálogo académico y propositivo entre investigadores de la UTMACH y de otras universidades iberoamericanas, que busca ofrecer respuestas glocalizadas a los requerimientos sociales y científicos. Los diversos textos de esta colección, tienen un espíritu crítico, constructivo y colaborativo. Ellos plasman alternativas novedosas para resignificar la pertinencia de nuestra investigación. Desde las ciencias experimentales hasta las artes y humanidades, Redes sintetiza policromías conceptuales que nos recuerdan, de forma empeñosa, la complejidad de los objetos construidos y la creatividad de sus autores para tratar temas de acalorada actualidad y de demanda creciente; por ello, cada interrogante y respuesta que se encierra en estas líneas, forman una trama que, sin lugar a dudas, inervará su sistema cognitivo, convirtiéndolo en un nodo de esta urdimbre de saberes.