

Benchmarking video presentations for CEFR usage in Cuba

Geisa Dávila Pérez¹, Frank van Splunder², Lut Baten³,
Jan Van Maele⁴, Yoennis Díaz Moreno¹

¹Dept. of English Language Studies, Universidad de Oriente,
Patricio Lumumba s/n Altos de Quintero, Santiago de Cuba, Cuba

²Linguapolis, University of Antwerp, Belgium

³Leuven Language Institute, KU Leuven (University of Leuven),
Belgium

⁴Faculty of Engineering Technology, KU Leuven (University of
Leuven), Belgium

This paper discusses language assessment by means of video recordings, particularly its use for benchmarking purposes regarding language proficiency in a Cuban academic context. It is based on videotaped oral presentation assignments of Cuban PhD students for peer and teacher assessment. In order to avoid bias and provide validity to the results, the PhD students' videotaped oral presentation assignments have been rated by language testing experts from three different Flemish universities, which are included in the Interuniversity Testing Consortium (IUTC). A selection of these assignments will be transferred to the university Moodle platform, and this compilation may be used to enable the start of a Cuban corpus of internationally rated presentations of academic English. Therefore, the results obtained will provide language teachers with a growing database of video recordings to facilitate benchmarking activities and promote standardized assessment in the Cuban academic context.

Keywords: language assessment, benchmarking, CEFR, Cuba, video recordings

Introduction

Cuba has only recently become involved in teaching English as an academic language at a country-wide level. The Cuban Ministry of Higher Education (*Ministerio de Educación Superior*, henceforth MES) officially declared, regarding its Language Teaching Policy, that all Cuban university students should reach and prove a B1+ level

of language proficiency by the time they obtain either a B.A. or BSc. degree² (Rivera, Torres & Estrada, 2017; MES, 2013). This policy, however, does not state “how” the Cuban educational staff should test this level in a valid and reliable way; it states the official strategy for teaching, but the only reference to testing is the language level expected from university graduate students according to the Common European Framework of Reference; henceforth, CEFR (Council of Europe, 2001). Moreover, the level of English for Academic Purposes (EAP) and Specific Purposes (ESP) tends to be generally low at Cuban universities (*English Proficiency Index*, 2017), and few if any standard setting and benchmarking practices are developed (Van Maele, Rodríguez, Díaz, van Splunder, & Baten; 2015). Consequently, the experience and expertise concerning English language learning, teaching and assessment in an academic context need to be developed. Even though the MES in Cuba allows for standardized testing (Maseda, 2015), and has made the CEFR the official standard to follow, there is a lack of overt and shared policy guidelines concerning English language assessment and testing.

Besides, language assessment and testing have not yet entered the university curricula as a subject in teacher training programmes for language instruction, and neither have they become a relevant research and work domain in their own right (Van Maele et al., 2015). Cuban language trainers design, administer and grade their own tests and tasks based on their experience as teachers rather than testers, and although previous efforts have been made in Cuba to align with the CEFR (i.e. international workshops, lectures and projects, especially in international cooperation), its implementation is new. Thus, while the new Language Policy in Cuba calls for testing in a standardized setting, standardization and benchmarking are not adequately developed; testing practices and test results can barely be considered valid, reliable and transparent due to the language teachers’ inadequate training in this field and the lack of practice as to nation-wide language testing.

Context within the VLIR UOS Project at Universidad de Oriente in Santiago de Cuba

Internationalization has become a crucial part of academic research and mobility in the Cuban context as a number of universities have had the opportunity to cooperate with universities in Europe, such as the cooperation of six Cuban universities with five Flemish universities in Belgium within the Flemish Interuniversity Council VLIR UOS (<http://www.vlir-uos.be>). On the one hand, there is the established dual degree for PhD students, which entails exchange of students and lecturers from both countries for research purposes. It is this international mobility that has also called for

²Infrastructure and staff professional development have not been achieved yet. That is why the MES have decided to gradually embrace this policy; currently, university graduate students should reach an A2 level. Retrieved from <http://www.mes.gob.cu/es/politicas-aprobadas>

assessment aligned with international standards. The first Cuban university opted for the CEFR already in 2004, and the need has strongly grown over the last thirteen years to assess Cuban PhD students', project leaders' and administrative personnel's language proficiency level. On the other hand, the MES decided to use the CEFR as its reference to start developing an appropriate framework for testing in 2015 nationwide. The basic requirement set for Cuban university students is that they should reach and prove a B1+ level of language proficiency by the time they graduate (Rivera, Torres & Rodríguez, 2017). Hence, bottom up and top down, the policies of the involved stakeholders meet, but the crucial question that arises is: who delivers the guidelines and the training regarding language testing standardization? It goes beyond saying that lecturers and teacher trainees should at least be trained into a good understanding and use of CEFR descriptors in an EAP context, if they are required to grant a certification.

This question became particularly adamant in 2013 when certification was required for oral proficiency of PhD students at Universidad de Oriente, Cuba. This university was chosen for the Inter-University Collaboration programme VLIR UOS, which is devoted to contribute to capacity building in higher education in Eastern Cuba (www.vlir.uo.edu.cu/en). This programme includes international mobility, research and recycling scholarships for graduate studies, such as MA and PhD courses at Flemish universities, and training and research programmes at the Cuban campus as well. Most of the applicants are teachers and researchers who are involved in any of the eight projects and who belong to different fields of science, i.e. Biology, Physics, Biophysics, Humanities, and Mechanical, Biomedical and Chemical Engineering.

In order to be accepted by the Flemish universities, these Cuban PhD students have to reach and prove a B1 level in oral proficiency and, in some institutions of higher education, even a B2 level is required to be allowed to defend a joint PhD. As a result, the B1 level for oral proficiency is the gateway to entering inter-university collaboration. Hence, certification of oral proficiency is essential.

In this context, the transversal project *Strengthening Foreign Language Skills for Intercultural and International Academic Purposes*, one of the projects within the VLIR UOS in Santiago de Cuba, sustains MA and PhD students who apply for scholarships abroad. The project management agreed with the decision in 2013 to implement the CEFR, to train the trainers accordingly and to sustain the development of local testing for EAP purposes. In order to raise professionalism in language learning, teaching and testing, recycling scholarships in Belgium were organised and gradually learner autonomy (Baten, Rodríguez Pozas & Van Maele, 2011) implemented while designing and piloting tailor made courses and tests for this target audience. As stated above, a core concern was oral proficiency, and along with it, the preparation of both the graduate students and their trainers for international oral proficiency in an academic context of English for Science and Technology, a domain entirely new to the local

language instructors of English, primarily trained as translators.

In the development of assessment skills, while learner autonomy and professional practices develop, mentoring plays an essential role (Baten, Beaven, Osborne, & Van Maele, 2013). Such detachment and collaboration indeed provide trainees with the opportunity to actively engage in further learning and reflect on how to improve professionally. In that sense, the foreign language skills project in Santiago de Cuba found inspiration in the outcomes of the WebCEF project; i.e. an online tool for assessment of oral proficiency in a foreign language. It used video recordings for peer and expert assessment, creating transparency as to the levels and descriptors used in CEFR for oral assessment for English for Specific Purposes. WebCEF (<http://www.webcef.eu>) succeeded in providing a platform for language teachers and learners to familiarize themselves with the CEFR and how to set up communities of practice (Wenger, 2006) involving peers to exchange their perception of each other's oral language proficiency.

Within the VLIR UOS cooperation, Flemish lecturers attempted to use WebCEF to give their impressions and comments on presentations of PhD proposals. Unfortunately, accessing the online tool was not an option in the Cuban context due to internet connectivity issues, but trainers adopted the concept and applied the approach to fit assessment training needs and learning challenges in the Cuban Eastern context. Thus, the need arose to collect PhD students' videotaped oral presentation assignments that would help trainers and teacher trainees overcome three main shortcomings. First, assessment is not part of the curriculum in teacher training programmes in Cuba; second, there is a lack of benchmarked Cuban videos contextualized for the purposes of these courses and, thirdly, Cuban trainers and students are not familiar with the autonomous learning approach.

Aims

This article focuses on the use of self-videotaped oral presentations of PhD students, which were rated by testing professionals, as a tool to foster autonomous learning among graduate students and to assist language trainers and teaching trainees in the standardized assessment of oral performance in the Cuban context. The video-recorded assignments would consequently become a necessary tool for both the Cuban language trainers and PhD students in order to upgrade professional levels of assessment and acquaintance with autonomous language learning formats with graduate students (peer and self assessment).

Literature Review

Teaching and learning have moved to learning facilitation, collaboration and autonomy (Holec, 1981; Beckman, 1990; Dooley, 2008), creating an environment where

teachers and students play different roles; i.e. the teacher becomes a “guide” for the work to be done and students actively participate in the building up of their knowledge. Language teachers also apply such principles and use classroom collaborative interaction, adult-oriented strategies and problem-solving techniques to encourage students theorize about what they have learnt in class (Preece & Griffin, 2002). Such dynamics are considered to be positive because students become critical thinkers and more independent, they retain the information longer, appear more satisfied with classes and take charge of their own learning (Holec, 1981; Beckman, 1990; Goodsell et al., 1992).

In the same line of thought, language assessment practices have also undergone major shifts to be consistent with current language teaching trends and to address the demanding emphasis on standardized assessment (Hamayan, 1995); i.e., aligning tests to the CEFR. As a result, alternative assessment practices have been suggested to facilitate teachers to tackle differences in learners, to address learning over a period of time, to include communicative performances in a variety of ways, and to have students provide reasons and clarifications, instead of marks, to support their assessment in self and peer assessment activities (Brown & Hudson, 1998; Dooley, 2008; Black, Harrison, Lee, Marshall & William, 2009).

These new alternative approaches provide a constructive environment in which learner autonomy is enhanced and mentoring can take place at different levels: between teacher and learner, between teacher trainer and trainee, but also among peers (Baten, Beaven, Osborne & Van Maele, 2013). A number of research publications (Hardford, & MacRuairc, 2008; Rich & Hannafin 2009; Newhouse, Lane, & Brown, 2007; Rosaen, Lundeberg, Cooper, Fritzen & Terpstra, 2008) have particularly suggested that self-recorded videos of oral performance provide an excellent opportunity to analyse and discuss oral abilities. This is mainly due to the fact that body language as well as oral production are captured, offering language teachers detailed and rich data on the communication process. Van Es & Sherin (2010) also claim that self-recorded videos of oral performance allow students to view his/her performance repeatedly and observe patterns of behavior as well as language that should be modified or changed.

Regarding language assessment standardization, the CEFR indeed provides a transparent measure in educational institutions worldwide. Nevertheless, it has proved to be a challenging mission to match communicative ability, language tasks and learning outcomes to the descriptors proposed (De Jong, 2010; Davies & Lishman, 2014). And many practitioners find it difficult to determine what the levels actually mean in spite of the wide distribution and careful formulation of the CEFR scales (North, 2008; Baten, Osborne, D’Silva, 2009; Fulcher, Davidson, & Kemp, 2011).

Consequently, several endeavors within the European context have been made to illustrate the CEFR scales in order to enable their application in the assessment practice of language proficiency. For example, the project “Multilingual Platform for the European Reference Levels: Interlanguage Exploration in Context” (www.merlin-platform.eu), a freely accessible online platform, which uses learner corpora to illustrate the CEFR levels for German, Italian and Czech and contribute to the validation of selected CEFR scales; the project “Modularising Multilingual and Multicultural Academic Communication Competence” (www.magicc.eu), a freely accessible tool, which provides pedagogical scenarios and assessment resources within a multilingual and multicultural academic setting; and as previously stated in this paper, the WebCEF project. These projects contribute to transparency, validity, reliability and benchmarking of assessment because the selection of language samples and scenarios have been consistently rated at a particular level by participating expert assessors (Van Maele, 2009; Wisniewski et al., 2013; Meima & Neuner-Anfindsen, 2016). At the same time, learners are directly involved in the assessment practice as they can play the role of assessors of their peers’ and their own language production (Black et al., 2003; Van Maele, 2009).

It should be noted, however, that connectivity-related issues may hamper the application of these online tools in a Cuban setting. Furthermore, it should also be highlighted that they have been developed in a European context by means of approaches which may be called innovative and challenging for the Cuban scenario. Rather than strictly implementing these specific practices, their core concepts should be gradually blended in the English language learning practices in Cuba in order to answer the current demands and achieve progress.

The analysis of the academic literature in these areas also illustrates the swift development from paper to digital in language teaching and assessment. The internet has transformed the way in which foreign language courses are taught (Blake, 2008; Levy, 2009; Jarvis & Krashen, 2014; Richards, 2015), and learning management systems (LMSs), such as Moodle, provide suitable course and assessment formats for foreign language instruction (Berns, Gonzalez-Pardo, & Camacho, 2013). For a more thorough review on the use of Moodle for language teaching, see Brandl (2005) and Warth-Sondheimer (2011).

Methodology

Participants

Twelve PhD students’ videotaped oral presentation assignments were used. These were PhD students who enrolled in EAP courses taught by Cuban language trainers in the context of the transversal project *Strengthening Foreign Language Skills for*

Intercultural and International Academic Purposes, one of the projects within the VLIR UOS in Santiago de Cuba. Nevertheless, five out of the twelve videos were not rated as the recording quality did not meet the standards for rating purposes. The seven videos selected were rated by minimum two raters involved in the Interuniversity Test of Academic English for Students (i.e. ITACE for Students, <http://www.itace.be/students/about.html>; a language test which is used by all Flemish universities. It relates test taker's skills with the CEFR levels and has been validated by an international team of language experts). These raters were at least from two different universities, i.e. a total of six raters.

Tools

Bearing in mind, on the one hand, the importance of gaining mastery in the use of the CEFR descriptors for self, peer and teacher assessment of PhD students' oral proficiency and, on the other hand, the impossibility of using WebCEF in the Cuban context, local language instructors of English relied on the use of students' self-recorded videos to foster self and peer assessment, and eventually benchmarking activities of oral performance. It was conceived that the video-recorded assignments of PhD students needed to be included in an e-platform for open access. Moodle emerged as the most suitable option due to two main reasons: it is the e-platform widely available and accessible in the Cuban context, and it is user-friendly for non-technical people (Berns, Gonzalez-Pardo, & Camacho, 2013).

In order to avoid that self-recording became an issue due to technological availability, the PhD students were asked beforehand whether they had a means to record themselves. Most students involved in the study either had a Smartphone or a camera, while the remaining students could make use of a digital camera provided by the project.

Selection of a video corpus

For the purpose of academic mobility, Cuban PhD students attended a 10-week EAP course which concentrated on effective communication skills in an international academic environment, particularly in the context of a science project. The course included tasks to improve their oral presentation and interaction skills in English. To reach the required language level, the students had to master specific language functions, such as: expressing their opinions, giving clear, detailed descriptions and presentations on a wide range of subjects, expanding and supporting ideas with relevant information on the results of their research work as well as negotiating aspects related to both work and academic environments in face-to-face encounters in larger meetings.

One of the tasks the students had to complete in the courses was to video-record themselves while presenting their research to a simulated board of experts in their academic field (see Appendix A). Since this practice combined both oral presentation skills in English and research validation, the video recordings collected from this task were used for this study.

Once the students had video-recorded their presentations, they handed in the video in advance to their teacher and one peer assessor, selected from their class. The peer had to comment on the video considering the teacher's input concerning language and oral presentation techniques in an academic context. For that purpose, the peer assessor was given a template (see Appendix A) with the key elements to be considered for assessment. These elements referring to the use of language, organization of the presentation, nonverbal communication, and time management had been previously covered in class.

After the presentation of each video in front of the class, the designated peer assessor, who had reviewed the video in advance, commented on and asked questions about the presentation based on the template given and the learning outcomes from previous lessons. The rest of the class was also prompted to provide assessment and feedback on the same video. The role of the teacher was to act as a mediator in this activity and to intervene only when necessary. The twelve videotaped assignments selected included different academic domains (Biology, Physics, Humanities, Engineering, Biomedical Science, and others) and approaches, i.e. proposals for PhD studies and dissertations.

Rating of the self-videotaped oral performances of the PhD students

Since there is limited experience in Cuban higher education concerning the assessment of language proficiency aligned to the CEFR, the validity of the video assessment and benchmarking needed to be attended carefully. In order to avoid bias and to validate the claim that the videos displayed a B1 level of language proficiency, external expertise was needed. As the project is involved in an international programme comprising several Flemish Universities (i.e. Ghent University, University of Leuven, University of Antwerp, Vrije Universiteit Brussel, University of Hasselt), in which the Cuban students pursue their PhD, support from these universities was sought to validate the CEFR level of the videos.

The Interuniversity Testing Consortium (IUTC), which comprises four Flemish universities (i.e. Ghent University, University of Leuven, University of Antwerp, Vrije Universiteit Brussel), developed the ITACE for Students. Bearing in mind its expertise, this group was asked to validate the videotaped presentations. Six expert assessors from IUTC rated and benchmarked the material to validate the required B1 level

according to the CEFR. These raters had participated in training sessions to familiarize themselves with the procedure.

Each candidate was assessed twice and independently by different raters, and an additional rater was asked in case of disagreement or doubt. The holistic rating scale which was used consisted of five major items (vocabulary: control and range; grammar: accuracy and range; fluency; pronunciation and intonation; coherence and cohesion), and these items are linked to the communicative language competence scale of the CEFR.

Results and discussion

Useful information was gathered from the video-recorded oral performances of PhD students. A number of specific guidelines should be formulated in terms of recording quality; for instance, trainers should verify that all candidates use the same video format, background noises are to be avoided, and all video recordings should be the same length (no less than 2 minutes and no more than 3 minutes). Addressing these issues in advance will contribute to the uniformity of the video corpus for similar undertakings in the future.

In terms of language proficiency, the candidates were within the broad B1 range in general (see Appendix B). Whereas three candidates were at the lower end of B1 (i.e. B1-), the rest was at the higher end (B1+). Overall the raters agreed on the level, except for one candidate who had been rated A2 by two raters, and two other candidates who were rated B2 by one rater (and B1+ or B1 by the other raters). In such specific cases of discrepancy, a third rater was asked to settle the issue. The typical mistakes found in the videos referred to grammar (tense use, non-use of *it*, use of *of* to vs. *for*, word order and sentence structure, articles, preposition combinations), pronunciation (word stress, rolling /r/, mispronunciation of some consonants and vowels; mainly /v/ vs /b/; /et/ vs /t/ ; /k/ vs /dzj/ and /ʌ/ vs /u:/), vocabulary (*professor* vs. *teacher*; *get better professional skills*), and coherence and cohesion (misuse of specific linking words and cohesive devices *on other hand*, *in contra*). However, it should also be highlighted that the candidates used appropriate collocations (*acceptable performance*, *obtain results*, *complex situation*, *run the examples on a computer*); they correctly organized the information by means of linking words and cohesive devices (*I would like to start with... and then..., as you can see, firstly, in any case, I'm going to show you*); even though there was some hesitation, in general the candidates' speech was fluent and natural.

The results are fairly consistent, and all videos (except the A2 candidate) can be uploaded to the university Moodle platform. Furthermore, an individual file with the

assigned level and some raters' comments on language features for each candidate's video needs to be included for teaching or other purposes (see Appendix C). The PhD students' video recordings thus can become a reference for teachers and students alike. As a result of this study, three major conclusions may be drawn. Firstly, PhD students have been helped to monitor their presentations skills and their language level transition in an academic setting by providing feedback to their oral output and keeping track of their oral performance. Hence, the students are aware of what has been learnt already, what needs to be improved and what still needs to be learnt. Secondly, awareness of autonomous language learning with professionals (peer and self assessment) can be raised. Thus, it encourages PhD students to actively participate in their English language training and become used to this newly embraced praxis in the Cuban setting. Moreover, it promotes independent action, critical reflection and decision-making strategies; capacities that are taken for granted with graduate students. And thirdly, these benchmarked video recordings of PhD students' oral performances will remain available for teacher practice and training as a CEFR-based model tool to rate oral performance in the Cuban academic context. The outcomes of this project may therefore help increase Cuban language teachers' understanding and familiarity with the use of the CEFR scales by means of benchmarked PhD students' videotaped oral performance, improve their professionalism as language assessors in a standardized setting, and increase transparency and validity regarding Cuban language testing practices.

Furthermore, although the study started from a small sample of twelve self-recorded videos, it is advisable to develop the stock as a growing source of material for learning and assessment training. For that reason, more self-recorded videos should be added to the university Moodle platform in order to enrich the available database and provide more examples. Additionally, the scope should be broadened to other levels of language proficiency, mainly B2 and C1, as there is an increase in international academic collaboration and these are the levels which are required in this particular context.

Moreover, steps should be taken to set up a community of practice concerning standardized testing in Cuba, which may lead to the creation of a national group for exchange and cooperation to help develop a national assessment strategy³, and certify reliable and valid test results. By means of these actions, Cuban language teachers' professionalism may improve as language testers, and transparency regarding

³ These considerations materialized after the ILTA funded workshop: Assessment and Evaluation. Setting up a Cuban Assessment Network for Teachers. Havana, 10-14 July 2017. Retrieved from <https://www.iltaonline.com/?page=2017WorkshopAward>

language testing in Cuba may also be enhanced. Therefore, international visibility will be attained.

Conclusion

As observed earlier by Baten, Beaven, Osborne & Van Maele (2013), the use of students' self-recorded videos in an open-access platform may be regarded as a useful tool in mentoring practices because it supports the participants to critically ponder their actions and collaboratively build their knowledge. Such approach may also develop learner autonomy and professionalism for language students and teachers respectively. On the one hand, PhD students' autonomy is enhanced as they acquire awareness of their oral performance in an academic and standardized setting. On the other hand, language teachers' practices are further developed as Cuban language trainers gain experience in assessment practices within a common framework of reference for languages, i.e. the CEFR (North, 2008). On a more fundamental level, these videos can be used to professionalize language teaching and testing practices in Cuba, as a result of which it may be easier to gain access to the international market of higher education.

Acknowledgements

This work has been supported by the Belgian Development Cooperation through VLIR-UOS (Flemish Interuniversity Council-University Cooperation for Development) in the context of the Institutional University Cooperation programme with Universidad de Oriente.

The authors of this paper express their gratitude to the ITACE experts, who provided their expertise and valuable time to validate this work. Likewise, the authors would like to acknowledge the endeavours undertaken by Professor M.A. Humberto Miñoso from Universidad Marta Abreu de Las Villas, Cuba, to implement WebCEF in the Cuban context.

References

- Baten, L., Rodriguez Pozas, J.C. & Van Maele, J. (2011). *The ethical challenge of language assessment in a North/South collaboration programme*. Presentation at Ethics in Language Testing and Assessment, 8th Annual EALTA Conference, Siena, May 5-8, 2011. Retrieved from <http://www.ealta.eu.org/resources.htm>

- Baten, L., Beaven, A., Osborne, J., & Van Maele, J. (2013). WebCEF: An online collaboration tool for assessing foreign language proficiency. In Favaron E., Guerra L. (Eds.), *Handbook of Research on Didactic Strategies and Technologies for Education: Incorporating Advancements* (pp. 559-70). Hershey, PA, USA: IGI.
- Baten, L., Osborne, J., & D'Silva, Y. (2009). WebCEF: On-line Collaboration and Oral Assessment within the Common European Framework of Reference. *CercleS*, 25. (pp. 18-19).
- Beckman, M. (1990). Collaborative learning: Preparation for the workplace and democracy. *College Teaching*, 38(4), 128-133.
- Berns, A., Gonzalez-Pardo, A., & Camacho, D. (2013). Game-like language learning in 3-D environments. *Computers & Education*, 60 (1). (pp. 210-220).
- Black, P; Harrison, C; Lee, C; Marshall, B, & William, D. (2003). *Assessment for Learning. Putting it into practice*. Maidenhead: Open University Press.
- Blake, R. J. (2008). *Brave New Digital Classroom: Technology and Foreign Language Learning*. Washington, D.C, USA: Georgetown University Press.
- Brandl, K. (2005). Are you ready to "Moodle"? *Language Learning & Technology*, 9 (2). (pp. 16-23). Retrieved from <http://ilt.msu.edu/vol9num2/review1/>
- Brown, J., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly* 32 (4). (pp. 653-675).
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Davies, L & Lishman, J. (2014). *I'm lost! Mapping EAP descriptors to the CEFR*. Presentation at The CEFR and Language Testing and Assessment. Where are we now? 11th Conference of EALTA, University of Warwick, 29 May-1st June, 2014. Retrieved from <http://www.ealta.eu.org/conference/2014/presentations/Lucy%20Davies%20oster%20EALTA%202014.pdf>
- De Jong, J. (2010). *Getting the levels right: deriving item writer guidelines from CEFR descriptors*. Presentation at the 32nd Language Testing Research Colloquium, Cambridge, April 12-16, 2010. Revised at [https://cdn.ymaws.com/www.iltaonline.com/resource/resmgr/docs/LTRC_Past_Year_Programs/LTRC_2010_Program\(1\).pdf](https://cdn.ymaws.com/www.iltaonline.com/resource/resmgr/docs/LTRC_Past_Year_Programs/LTRC_2010_Program(1).pdf)
- Dooley, M. (2008). Constructing Knowledge Together. Extract from *Telecollaborative Language Learning. A guidebook to moderating intercultural collaboration online*. M. Dooly (ed.). (2008) Bern: Peter Lang (pp. 21-45). *Education First (n.d.)*. *English Proficiency Index*. Retrieved December 22 2017, from www.ef.com/epi
- Fulcher, G., Davidson, F. & Kemp, J. (2011): Effective rating scale development for speaking tests: Performance decision trees. *Language Testing* 28 (5) (pp. 5-29).
- Goodsell, A., Maher, M., Tinto, V, & Associates (Eds.). (1992). *Collaborative learning: A sourcebook for higher education*. Pennsylvania State University.

- Hamayan, E. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics*, 15, 212-226.
- Hardford, J., & MacRuairc, G. (2008). Engaging student teachers in meaningful reflective practice. *Teacher and Teaching Education*, 24 (7) (pp. 1884-1892).
- Holec, H. (1981). *Autonomy and foreign language learning*. Oxford: Pergamon. (First published 1979, Strasbourg: Council of Europe).
- Jarvis, H., & Krashen, S. (2014). Is CALL Obsolete? Language Acquisition and Language Learning Revisited in a Digital Age. *Tesl-Ej.Org*, 17(4), 1–6. Retrieved from <http://www.tesl-ej.org/wordpress/issues/volume17/ej68/ej68a1/>
- Levy, M. (2009). Technologies in Use for Second Language Learning. *The Modern Language Journal*, 93(Focus Issue), 769–782.
- Maseda, B. (2015, julio 14). Hacer el TOEFL en Cuba: Yes We Can. *OnCuba*. Retrieved from <https://oncubamagazine.com/sociedad/hacer-el-toefl-en-cuba-yes-we-can/>
- Meima, E., & Neuner-Anfindsen, S. (2016). MAGICC: A Project of the EU Lifelong Learning programme: Modularising Multilingual and Multicultural Academic Communication Competence. *European Journal of Applied Linguistics* 4(2). doi.org/10.1515/eujal-2016-0008
- Ministerio de Educación Superior. (2013). Documento base sobre el perfeccionamiento de la enseñanza del idioma inglés en las universidades cubanas. Aprobado por el Consejo de Dirección del Ministerio de Educación Superior, diciembre.
- Newhouse, C., Lane, J., & Brown, C. (2007). Reflecting on Teaching Practices Using Digital Video Representation in Teacher Education. *Australian Journal of Teacher Education*, 32 (3), 5.
- North, B. (2008). The CEFR levels and descriptor scales. In L. Taylor, L. & C.J. Weir (Eds.), *Multilingualism and assessment. Studies in Language Testing* 27 (pp. 21-66). Cambridge: Cambridge University Press.
- Preece, J., & Giffen, C. (2002). Critical and Radical Pedagogies. *Theory and Practice of Teaching in the Learning society*, Jarvis, p. (ed). London: Kogan Page.
- Rich, P., & Hannafin, M (2009). Video annotation tools technologies to scaffold, structure, and transform teacher reflection. *Journal of Teacher Education*, 60 (1) (pp. 52-67).
- Richards, J. C. (2015). The Changing Face of Language Learning: Learning Beyond the Classroom. *RELC Journal*, 46(1), 5–22. <http://doi.org/10.1177/0033688214561621>
- Rivera, S., Torres, M., & Estrada, P. (2017). Instrumentación de la estrategia de perfeccionamiento del inglés en las universidades del Ministerio de Educación Superior (MES). *Congreso Universidad*, 6(6), 73-89. Retrieved from <http://www.congresouniversidad.cu/revista/index.php/rcu/article/view/976>

- Rosaen, C., Lundeberg, M., Cooper, M., Fritzen, A., & Terpstra, M. (2008). Noticing noticing: How does investigation of video records change how teachers reflect on their experiences? *Journal of Teacher Education*, 59 (4). (pp. 347-360).
- Van Es, E. & Sherin, M. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher and Education*, 13 (2). (pp. 155-176).
- Van Maele, J. (2009). WebCEF: a conducive environment for staging useful language assessment? In: Bijnens, H., *WebCEF-Collaborative evaluation of oral language skills through the web* (pp. 21-26). Leuven: AVNet, KU Leuven.
- Van Maele, J., Rodríguez, R., Díaz, Y., van Splunder, F., & Baten, L. (2015). Weighing the consequences of introducing an international large-scale English language test in Cuba. LTRC 2015: From language testing to language assessment. Toronto, 16-21 March 2015. Retrieved from https://www.academia.edu/11603344/Weighing_the_consequences_of_introducing_an_international_large-scale_English_language_test_in_Cuba
- Warth-Sondheimer, C. (2011). Using Moodle for Language Teaching. Retrieved from [http://www.wideminds.eu/moodle/file.php/1/tutorials/Using Moodle for Language Teaching v1.2.pdf](http://www.wideminds.eu/moodle/file.php/1/tutorials/Using_Moodle_for_Language_Teaching_v1.2.pdf)
- Wenger, E. (2006). Communities of practice: A brief introduction. Retrieved from <http://www.ewenger.com/theory/>
- Wisniewski, K., Schöne, K., Nicolas, L., Vettori, C., Boyd, A., Meurers, D., Hana, J., Abel, A. (2013). MERLIN: An online trilingual learner corpus empirically grounding the CEFR Reference Levels in authentic data. *ICT for Language Learning, Conference Proceedings*. 2013. <https://conference.pixel-online.net/conferences/ICT4LL2013/conferenceproceedings.php>

Appendices

A. Sample task for students

Based on the previous class discussion about the elements for good presentations:

1. Prepare a 3-5 minute presentation on your own research results.
2. Video-record your presentation and exchange it with a partner.
3. Evaluate your partner's presentation and provide feedback taking into account the key elements for a good oral presentation covered in class:
 - Use of language:
 - Vocabulary
 - Pronunciation
 - Grammar
 - Register
 - Fluency
 - Nonverbal Communication
 - Eye-contact with the audience
 - Voice and intonation
 - Posture
 - Organization of the presentation
 - Structure (introduction, body conclusion)
 - Support (relevance and quality)
 - Time management
4. Time allotment for the task: 3-5 minutes per student
5. The rest of the class also discusses and provides feedback on the presentation being assessed.

B. Overall rating

PhD students	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	OVERALL
Candidate 1	-	B1-	-	B1	B1-	B1	B1
Candidate 2	-	A2	-	A2	B1-	B1-	A2
Candidate 3	-	-	-	B1	B1	B1	B1
Candidate 4	-	B2	-	B1	B1	B1+	B1
Candidate 5	-	B1	B1	-	-	B1	B1
Candidate 6	B1	B1	B1	-	-	B1	B1
Candidate 7	B1+	B2	B1	-	-	B1+	B1

C. Example of the raters' individual feedback for PhD student

Name	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Rater 6	Overall rating
Candidate 4	<p>fluent speech, clear pronunciation, natural intonation use of longer sentences, little hesitation, not too many mistakes</p> <p>good signposting, manages to get across rather complex message, stresses important issues</p> <p>Vocabulary: good collocations, does not have to search for words</p>			<p>coherence and grammar compensate for pronunciation</p> <p>Fluency: natural pace, a few hesitations but not problematic</p> <p>Pronunciation: strong foreign accent can sometimes hamper intelligibility</p> <p>Grammar: uses wider range with relative clauses and modal verbs (<i>principles that are difficult to understand, you need to show students that...</i>)</p> <p>Coherence: excellent use of linking words and cohesive devices: <i>from general to specific, I'd like to start with... and then...</i></p> <p>Vocabulary: good range for B1 with a few academic collocations (<i>put them in a context of, express concepts, present applications, run the examples on a computer</i>)</p>		<p>fluent, natural speech;</p> <p>well structured,</p> <p>uses quite complex sentence structures and linking words</p>	B1