Przemysław Krakowian

Katedra Języka Angielskiego i Językoznawstwa Stosowanego,

Uniwersytet Łódzki

ONLINE PEER ASSESSMENT OF SPEAKING WITH THE EPSS (ELECTRONIC PERFORMANCE SUPPORT SYSTEM)

Abstract

While developing and emphasizing oral skills in foreign language learning in line with the Common European Framework of Reference for Language (CEF) seems to be the general trend in the whole of Europe, the place and attention given to oral skills in the majority of schools across Europe is less than it deserves. This is so owing to a number of factors, one being that oral skills are slow to develop, require time and investment in terms of attention and resources and yet leave largely insubstantial evidence of effort, and finally teaching materials oftentimes focus on more manageable language elements and skills. The Web application described in the subsequent pages is a reasonable solution to this problem as it allows teaching staff to jointly evaluate the oral skills and proficiency of their pupils through a web-based environment. The main objectives of the rationale behind the web app are in bringing teaching staff in various institutions by providing them with a community where their assessments can be validated by their colleagues and to provide them and their pupils with tools to distribute and share annotated video and audio materials on the internet. This article outlines the workings of the web app, its major components, the logic behind it and the logistics in creating online assessments of oral proficiency in a collaborative assessment environment with a Community of Practice as a safeguard for the continual and sustained operation of the web-based evaluation environment for assessing oral skills.

Keywords: Assessment of Speaking, Assessment, Online Assessment, Computer Assisted Testing, Electronic Performance Support Systems, Knowledge Banks

Introduction

Computer Mediated Communication (CMC), as well as Electronically Mediated Communication (EMC), are becoming a staple feature of many working environments. Those working environments that rely on communicating notions for the purpose of evaluation of spoken performance in a learning and testing environment more and more commonly turn to the communal and collaborative arrangements for development, vetting and training solutions. One such arrangement may be rendered as a CoP (Community of Practice) where solutions are sought through a particular arrangement of duties, undertakings and responsibilities. Community of practice (CoP) is often defined as a network or a forum, both informal and with varying degrees of formal structuring and internal organisation, through which ideas are exchanged and solutions generated (Wenger, 1998). It implies the existence of a group of professionals, associated with one another through similarity of interests and expertise, and working on a common set of problems, in common pursuit of solutions, and themselves

constituting a store of knowledge (Wenger, 1998, Wenger, McDermott and Snyder 2002). The community of practice additionally entails the process of social learning that takes place when individuals who have common interests in some field or problem collaborate and share ideas, come up with solutions and otherwise interact with each other to work (Saint-Onge and Wallace, 2003; Hildreth and Kimble, 2004).

It constitutes a very attractive arrangement for many undertakings, as nothing binds people together faster than common interest and pursuit of solutions to common problems (Wenger, 1998; Wenger, McDermott and Snyder, 2002; Saint-Onge and Wallace, 2003; Hildreth and Kimble, 2004). One such scheme in which communities of practice, through the use of a web-based environment, collaboratively evaluate spoken performance is an Electronic Performance Support System. This web-based application, or webapp, is an offshoot of a project[1] sponsored by the EU and was carried out with the support of the Education, Audiovisual and Culture Executive Agency and the European Commission, Directorate-General for Education and Culture, under the Socrates - Minerva programme. The code behind this solution is freely available as Open Software. Currently, there are four instances of the application working in four different locations: KU Leuven in Belgium, University of Łódź, Universidad Central "Marta Abreu" de Las Villas in Santa Clara in Cuba, with the original copy maintained at the Open University in Milton Keynes.

1.   Electronic Performance Support System

The EPSS hosted on a server at Łódź University is a substantial re-write of the initial PHP code made available by the original consortium of project partners, with Łódź University vitally interested in making the code available to the general public under GNU General Public Licence, following the premise that the work on the project was EU financed, MySQL and PHP based and therefore open to inspection by any party interested (Meeker, 2008). The parts that have been re-written by the author to make the system usable for the purposes intended by the local examinations programme pertain to the portions of the system that relied on the use of proprietary software and which would require a commercial licence to be implemented locally. Two such areas were identified; the first involved a Flash®-based upload facility, which provided the server with a security contingency as well as a progress bar indicator to prevent the users from abandoning the upload process as a result of lack of progress update; the second a commercial video and audio encoder; both of which were replaced with a GPL alternative and integrated into the scheme. Portions of the system that now allow users to self-register have been written by the author from scratch.

The bare system, with a minimum of two sample files and a test user, owing to the compactness of PHP and MySQL, amounts to a little over 90MB in a compressed archive, which can be downloaded from the author's home pages as an ISO image of the EPSS CD (see Bibliography). With minimum PHP savvy, a computer literate but non-specialist enthusiast should be able to set it up in a little more than ten minutes.

There are very few systems that offer electronic support and remote access to examiners involved in assessing oral performance and examinee performance in the form of annotated samples comprising examinee portfolios. Most systems listed below offer some online access to assessment tools, but few could be of use in assessing speaking and examiner training.

DIALANG (www.dialang.com) cannot be treated as such, as it was never meant as a support system for examiners, it should instead be treated as a low stake feedback facility for learners of various languages. Additionally, it has been out of commission at its original

---

[1]   cf.: [www 1]

location, with Lancaster University running the technology of DIALANG "pro bono" and for some time now working to try to find a sustainable future for the system.

LOLIPOP – the Language On Line Portfolio Project (www.lolipop-portfolio.eu), though capable of storing various types of student work, is not designed to be a database facility for developing expertise in oral assessment or for training examiners.

SACODEYL (www.um.es/sacodeyl) is a web-based system for open distribution of European teen talk in education, which makes speech samples available, but offers no facility for annotation or benchmarking.

CEFTRAIN (www.ceftrain.net) offers practical sessions in assessing language samples, which are meant as a benchmark to explicate the workings of CEF, hence the name, but samples are few and far between.

VERSANT (www.ordinate.com), which was formerly known as the Spoken English Tests (SET), offers automated tests of spoken production, using speech processing and recognition technology, incidentally the same technology that is part of the French Arianespace programme. Tests can be taken over a telephone or a computer equipped with a modem and a microphone, and are automatically scored by the system.

Finally, LIVEMOCHA (www.livemocha.com) offers courses in six languages and aims at building a global user community, encouraging interaction and improving motivation of its members. The assignments, however, are in no way connected to any established marking or assessment system, and no support for examiners is provided.

One facility offering a similar type of support for examiners is ESOL CAMBRIDGE FRONTER (fronter.com/cambridge), essentially a VLE through which training materials and benchmarking activities are offered, but its use is restricted to examiners involved with the Cambridge ESOL exams. It provides three general areas where oral examiners can familiarise themselves with the standard, practice marking and undergo a verification procedure. Since the Fronter system is essentially a VLE, and as such is not meant for public access, it only offers access to the examiners associated with ESOL and its use was never intended for the general public seeking access to speaking samples and striving to hone their assessment skills and expertise.

An interesting development can be seen in another EU sponsored project – SpeakingApps (www.speakingapps.eu – this one, however, is meant as a plug-in resource for the Moodle VLE environment, and while it offers a substantial experience upgrade to the otherwise mute VLE environment, it is not meant as an application for collecting, evaluating and semi-automated assessment and training procedures.

2. The EPSS Structure

Work with the EPSS follows a different line of thinking and a different structural organisation, allowing greater access and a greater variety to anyone who is interested in obtaining a user account. The system recognises four types of users: the system administrator, mentor of mentors, mentor and student. All four have different system privileges and levels of access. The hierarchical organisation of user accounts was dictated by the original arrangement of the project, which assumed that certain files, information and the overall structure of user networking should be restricted to users with specific access level as well as group membership. All of this was dictated by issues related to user and data privacy, as any on-line, web based and publicly available system, whatever the intentions of the designers may be, poses privacy and data protection issues. In essence, an administrator account allows the kind of access that permits account creation, editing and access and privilege augmentation and suspension, with deletion possible only manually and for logistical reasons outside the system as any user activity leaves a trace in the system in the form of interconnectivity between the user, task and group interdependencies, thus possibly resulting in a situation where

the integrity of the database and the effectiveness of MySQL queries may be at risk, and in extreme cases render the system unusable.

Other levels of access include the mentor of mentors, a simple mentor and a regular student account. At all levels of access, with the only exception of a student account, the user can create new user accounts and assign the user a role in the system, as well as create user groups and determine the group membership. The role associated with a user account determines which of the groups are visible, and potentially accessible to the user. A mentor group created by a mentor of mentors, the highest access user account, will be open only to other users with a mentor of mentors status. On the other hand, any group created by a regular student will be available by invitation to any user enjoying a higher access status as well as any other student user. A peculiarity in the system lies in the fact that a group creator is not automatically a member of the group which is started in the process. Instead, just as with any other group membership, the creator of the group needs to add his account to the group membership list if he wishes to have access to the samples and annotations made within the group. The assumption of the creators of the original EPPS here was perhaps, especially with higher access accounts, that the creator was not necessarily always interested in what became of the group and the assessments that ensued. The EPSS at Łódź University includes a facility, which makes the group membership optional when creating a group, with the option set to yes by default.

3.   The EPSS Workflow

Work with the system begins with creating a group with group user availability determined by the creator level of access, but amendable through a set of drop-down boxes. In order to make the group active and usable, the group creator needs to associate the group with one of the assessment schemas and provide future users with the description of the group and types or scope of assessments envisaged. The group may be created prior to is actual use and kept inactive, as well as it may eventually become suspended if interest in the assessments and group activity is discontinued. For reasons associated with database integrity, a group may only be deleted manually and outside the system by a direct MySQL query by the administrator; which is, however, not advisable as it may unravel the delicate tapestry of the group-task-user-assessment interconnectivity.

The next step in the workflow entails designing an assessment task, and involves providing metadata pertaining to various aspects of the future assessment instances – envisaged level in relation to CEF, status of the task in related curricula, if any, setting formality, support material, including any audio and visual files to be stored in the system as well as any meta information on the relation of the task to any existing publications of examination tasks, task sets and exam materials. A separate provision is made to store any information relating to the sample recording protocol and the procedures used in acquiring samples. In order for a task to become usable, it needs to be associated with a group, which in turn determines the type of assessment schema.

The assessment schema is added to the database manually and by the system administrator; at the moment there is no provision to do so at the individual user level, and while in the future such a functionality may be added by popular demand, the original WebCEF project [www 1] was centred around the CEF scales, and no need to add, clone and modify scales was perceived. The EPSS hosted on a server at Łódź University is currently implementing two separate assessment schemata on top of CEF, in no way related to CEF, and a process of refining and fine-tuning the scales is under way, with all of the changes to the performance descriptors dictated by the process and feedback information from exams. Tracking the various interim scales, managing scale integrity, and various consistency-ensuring tasks are, in

the absence of a separate facility to automate the task, managed by the author, who is the EPSS administrator.

With an assessment schema in place, work can begin on samples that first need to be uploaded, transcoded, made ready for use and eventually associated with the task. Two modes of operation are possible with the current EPSS code: the first involves encoding done on the server directly after the sample has been uploaded and before it has been made available for further processing, while the second assumes that the user will have already encoded the sample into the required format and performs a compatibility check on the sample. The choice as to which of the two is used is again made at the level of the code and by the system administrator. While the first of the modes is more user-friendly and makes work with the system more straightforward, it puts considerable strain on the server resources, both file allocation wise as well as in processing times.

To illustrate the point, the Open University in Milton Keynes, where the original WebCEF server [www 1] operates, uses a separate, dedicated graphics workstation sporting a four quad core processing unit to cope with the demands of transcoding as well as employs a distributed computing cloud-type storage facility to process incoming multimedia traffic. This results at least partly from the fact that the server accepts high fidelity and high definition video and limits upload to 2GB. Since the resulting video to be later used in sample evaluation, which was originally meant to be undemanding in terms of bandwidth and hardware, is a mere192x144 FLV flash movie, such liberality in reserving considerable resources was seen by the EPSS designers in Łódź as wasteful and overly extravagant. The added advantage of the arrangement employed in Łódź lies in the fact that the file is available for processing and evaluation the moment the upload is complete, provided of course that the file has been properly transcoded by the user beforehand. A separate workstation, on the other hand, requires scheduling transcoding and sometimes bottlenecking may transpire as a result of errors in encoding, when the system is unable to determine automatically if the encoding stalled or possibly failed and when it merely is still in progress.

With a sample uploaded, associated with a task, a user group and an assessment schema, the user may proceed to make assessments. This is achieved by navigating through a number of screens where individual scales are displayed and where the user may review the sample, pause it where necessary, and rewind it using the controls in order to finally make selections as he deems fit. Once the selection is made and saved, the user proceeds to subsequent scales and finally publishes the assessment. At any time, the assessment may be stopped and work on it can be suspended and the state of assessment is then automatically saved by the system with a corresponding time stamp for further editing. Any number of assessments of a given sample may be made, and any number may be left unfinished. Samples that have not been published, and therefore not made available to other group members, may be deleted, while the ones that have already been completed and published may only be deleted manually and by the administrator, with the reservations mentioned previously.

Apart from the scales that are employed in any of the available marking and assessment schemas, the system makes provisions for manual annotation, notes and comments from the assessors. Those come in two separate forms, as comments applicable to the entire sample or time specific comments, which have the added advantage that the relevant part of the sample can be played at the click of the mouse. The annotations can be used as comments, explications as well as illustrations of the finer points of the scale that may be of importance to the future user of the system as well as any CoP (Community of Practice) working on a particular task.

The system additionally allows the user to recycle samples through a functionality that makes a provision for any sample visible to the user, whether this is the sample originally uploaded by the user or any of the samples assigned to the user group the user is currently

a member of, to be cloned. Cloning produces an identical instance of the sample, but stripped of any meta information or existing evaluations that may have been performed. In essence, the sample must undergo an identical procedure as any other newly uploaded sample before it can be used for evaluation. This functionality is especially useful in situations where the oral performance of more than one subject is evaluated. The cloning saves the time otherwise used on uploading and limits the space allocated to the file, as the system stores reference to the original file and not the duplicate.

The file cloning procedure, serves an additional purpose of allowing the cloning party to strip an interesting file of any associated information. The idea behind a community of practice (CoP) is that people collaborating on a task and sharing their opinions and points of view, through the differences of their professional opinions create a multidimensional account of the phenomenon accounting for its richness and complexity. With convergence being paramount in evaluating oral performance, there is a natural temptation to consult the opinions of others before committing oneself to a judgment, especially that within the same user group all evaluations are open to user inspection once they have been published. The more evaluation traffic any sample receives, the greater the risk that the opinions of the subsequent users may be biased and tainted by unsavoury curiosity how others view, perceive and evaluate it. On the other hand, a sample stripped of any history of evaluations, of any contextual information as to its origin, provenience and purpose, and for that matter also of its original name and timestamp, would require considerable effort to be traced within the system, and if additionally belonging to a different user group, would in fact prove untraceable and for all intents and purposes be seen by the intended audiences as an entirely new file.

Sample cloning may additionally be used to separate the uses and the intended applications of the different instances of evaluation. The annotation functionality does not limit the database allocated storage in practical terms, or does so for only some of the samples that would perhaps be previously eliminated by the upload limitations of the server. The annotation, therefore, does not need to be restricted to short, incidental observations made on top of the evaluation to highlight a point that in the opinion of the rater is instrumental in justifying or emphasising a particular choice made or a band chosen on the scale. Instead, the entirety of the sample can be transcribed and then used as a starting point for further analysis. Since the text is relatively lightweight and undemanding in terms of the database storage, there is essentially no limit to the degree to which the transcription can be augmented with various kinds of information, starting with accuracy, range, appropriateness, discourse markers and linking words to prosodic features and their effective use for message enhancement.

Given the fact that various features of the EPSS and different intended and unintended applications and uses, as well as the emergence of a variety of interest groups of users are bound to generate sample and assessment traffic, an issue arises as to how to manage what is being done by different groups and whether perhaps different groups are not redundant in their efforts. The system does not provide a listings or a directory functionality, neither is there an internal discussion forum or a system of internal mail or point to point communication or chat-room at the disposal of the users. Obviously, the users can establish and maintain communication outside the system, but the prerequisite for this is that they know about one another. A fragmentary glimpse into the complete extent of activities in the EPSS is offered through a metadata export functionality. Evidently, even though referring to whole of the samples and their assessments, the metadata provides only the sort of information that has been included by the user uploading or cloning the sample, hence, in order to avoid incompleteness of the insight into the assembled oral performances and their evaluations, it is recommended that the metadata goes beyond what is required by the system.

4.   Working with Samples

While there may be numerous applications of the EPSS, two seem to be the most prominent, with the former more forthrightly appealing to an average user and the latter more attractive to examiners or examiner trainers. Primarily, the EPSS may be used in the process of providing feedback on the quality of oral production, and in this it may be instrumental in targeting different interest groups, ranging from learners, peer-learning arrangements, teachers, learning and teaching supervisors and standards authorities, to finally cater to parents wanting to learn about their offspring. Secondly, but no less importantly, the EPSS may be useful to any parties interested in refining their understanding of rating scales; whether those are any of the aforementioned or those involved in assessment, assessment vetting and training of assessors. The EPSS, apart from its main function of being a repository of peer/mentor assessed speaking samples, additionally has the potential for storing and retrieving any text based information entered by the assessors/raters in the process of pinpointing the merits and shortcomings of oral performance, and may therefore be used as a corpus of learner language providing various insights of interest to a language researcher.

When attempting to assess a sample, a brief summary of the task that the sample was based on is provided on one of a number of preliminary screens. Any materials that were used for the task such as instructions, text, pictures, audio/video recordings will be listed in the box of "associated documents". Under most typical circumstances, file names in the right-hand column will allow to open and view the documents, or in some cases, if the document is on a website, a URL will be provided instead of a direct link to the file. While browsing the files and familiarity with the task behind the sample is not a prerequisite for accessing the sample assessment screen, it is advisable and highly recommended, especially so for scales that refer performance to task completion. Once the preliminary screens containing sample summary and task details have been displayed, the system opens the main assessment screen, which has been reproduced below. The main assessment window consists of four major zones which are responsible for displaying and controlling the playback of the sample, annotations, displaying the rating scales and providing summary information on the assessments of the sample on various subscales. The screen seems initially somewhat overpopulated and difficult to navigate through, but in essence it has a very straightforward arrangement and is easy to learn once the logic of making the assessment is understood.

Zone 1: The video player - this has the usual buttons for playing, pausing, rewinding, plus a row of buttons underneath such as "clear", "go", and "+" which are used for adding time-stamped annotations to the sample.

Zone 2: Annotations - this zone is used to add comments and explanations to the assessment. In a collaborative assessment, this can provide useful information on why a sample has been assessed at a given level as it explains how the assessor arrived at the decision. In training assessors, those comments can be used to help the assessors/raters better understand the scale descriptors as well as hone on particular levels.
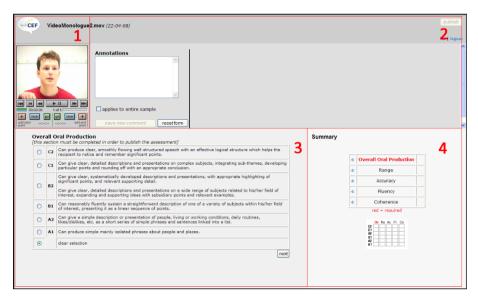
Figure 1. The main EPSS assessment window with zones

Zone 3: Assessment scales - this zone displays the descriptors for the scales (CEF in the illustration above) used in evaluating oral production, so that the level can be selected (here A1 to C2) which corresponds most closely to the sample being assessed.

Zone 4: Summary table – once the assessor/rater has arrived at the assessment of the level corresponding to the sample on a given scale, this choice will appear in this table, so it can be seen at a glance which levels have been chosen for each of the scales; in this example scales relating to Overall Performance, Language Range, Language Accuracy and other.

Since there may be two types of samples to be evaluated on the assessment tool: oral production in the form of monologue of sorts by one speaker and spoken interaction performed by two or more speakers, when evaluating interaction, there is need to indicate which of the two or more speakers is to be assessed - this information may be displayed in a caption on the video, or in the task description.

The assessment scales area (zone 3) shows the descriptors for all levels on individual scales. In the case of the example above, these are the standard CEF descriptors published by the Council of Europe. Note that the language the descriptors are displayed in is set by the person who created the group if the scales are available in different languages. To indicate the level to be assigned to the sample, radio buttons in the left-hand column are used. Confirming the selection is achieved through clicking on the "next" button at the bottom of the annotation form.

The assessment will then appear in the summary table (zone 4), and the next set of descriptors will appear. At any time before the assessment is published and made available to other assessors/raters, it may be modified or deleted. The scales are displayed in the order shown in the summary tables above. However, it is possible to move back and forth through the scales and do the assessment in a different order if preferred. As the assessments for each of the scales are selected, the choices will appear in the summary table. In all cases, the Overall rating, whether it is Overall oral production or Overall spoken interaction, is a required field and the assessment is not completed without an Overall rating.

4.1 Adding annotations

Annotations, which essentially are comments, explanations and examples which are aimed at illustrating the assessments – can be added to any of the assessment pages, so separate comments can be added for each subscale. The comments simply need to be typed into the text box in zone 2, and they can be in any language and of whatever length that is desired for the purpose. There are two types of annotation: overall comments, which apply to the whole of a sample within a scale/sub-scale, and time-specific comments, which apply to a particular time segment within the sample, again within a scale/sub-scale.

4.2 Overall comments

An overall comment can be applied for each section of the scales used for evaluating the task behind the sample, and in order to add such a general comment, the text needs to be entered into the annotation box and "applies to entire sample" box needs to be ticked off. The "save new comments" button will then be activated and when clicked on, the text of the annotation will be transferred from the text box to the right-hand part of the annotation zone, accompanied by two icons that can be used to edit the comment, or to delete it altogether. There is no limit to how many comments or annotations can be added to the assessment of the sample, or to what their length may be other than what is set up in the database; obviously depending on the envisaged applications, the database can be altered and augmented accordingly by the administrator.

4.3 Time-specific comments

If a remark concerning a specific section of the recording such as any comment on a particular section that provides a good illustration of the speakers' performance such as range, fluency, discourse management etc. the annotation can be linked to a time-stamp which is placed in the recording and stored in the database behind the EPSS. Owing to this facility, when looking at the assessment the comment can be referred to the exact place in the recording and the recording will automatically start playing at the beginning of the section which has been time-stamped and selected for inspection. As many time-stamped annotations as there are needed may be added to the recording; the different annotations may be overlapping or repeated. The size of the annotation is not limited otherwise than by the setting in the database, which may at any time be changed by the administrator if a need arises.

4.4 Publishing the assessment

Before the speaking performance is rated on the compulsory overall scale, and before the entire set of ratings is published with the use of the "publish" button (zone 2), only the author can see the assessment; and at any time there may be any number of such assessments not visible to the general public browsing samples within the group. To make the assessment(s) available to other members of the group, the author needs to publish it. At any time when the assessment is considered complete, whether all of the scales have been used, or merely the overall scale has been selected, and whether any annotations have been made or not, the sample may be published by clicking on the "publish" button in the top right-hand corner of the screen. Once the assessment is published, it cannot be modified or deleted by the author, and while the administrator of the EPSS may intervene at the level of the database, such course of action is not recommended for reasons outlined earlier in connection with the database integrity and sample-assessment interconnectivity. Any number of subsequent assessments, however; may be published when the author decides that any previous instances lack precision, clarity or require additional comments and explications in the form of annotations.

5.   Concluding remarks

Oral skills in foreign language learning are of paramount importance, and while this is something that everyone will agree on, they are slow to acquire and their development is resource intensive. The place and attention given to oral skills in the majority of schools across Europe is less than it deserves and this is so owing to a number of factors, one being that oral skills are slow to develop, require time and investment in terms of attention and resources. The EPSS application described in the preceding sections is a reasonable solution to this problem as it allows teaching staff to jointly evaluate the oral skills and proficiency of their pupils through a web-based environment. This is an invaluable source of insight into learner competencies, learner language and progress and effectiveness of teaching.

Apart from the original applications of the EPSS to constitute a meeting ground for teachers, assessors, and students/learners and any other interested parties in the form of a community of practice (CoP), the system has the potential to offer numerous other opportunities including amongst others a training platform for assessors, a system for creating learner portfolios, a vetting system for rating scale construction or a multimedia annotated corpus of spoken learner language. Obviously, any information stored in the system that refers to instances of spoken performance may be used for research purposes. Finally, but not any less importantly, peer student-to-student assessment should encourage self-reflection and could be helpful in self-evaluation; a skill which is difficult to develop among students.

An advantage that is not to be forgotten is the fact that the system is available to anyone interested in hosting the system under the Open Source Licensing, meaning that not only may it be used for free, but any modifications to augment the system for particular needs of the user are both allowed and keenly encouraged.

REFERENCES

Bachman, Lyle and Adrian Palmer. *Language Assessment in Practice*. Oxford: Oxford University, 2010.

Brindley, Geoff. "Language assessment and teacher development." Paper presented at the International Conference on Teacher Education, City Polytechnic of Hong Kong, April 17–19, 1991 http://www.mq.edu.au/about_us/faculties_and_departments/faculty
_of_human_sciences/linguistics/linguistics_staff/honorary_staff/associate_professor_geoff_brindley/papers (access date November 5, 2015).

Brooks, Lindsay. "Interacting in pairs in a test of oral proficiency: Co-constructing a better performance." *Language Testing*, 26(3) (2009): 341–366.

Hildreth, Paul M. and Chris Kimble. *Knowledge Networks: Innovation Through Communities of Practice*. Hershey: Idea Group Inc. (IGI), 2004.

Meeker, Heather J. *The Open Source Alternative: Understanding Risks and Leveraging Opportunities*. Hoboken, New Jersey: John Wiley and Sons, 2008.

Saint-Onge, Hubert and Debra Wallace. *Leveraging communities of practice for strategic advantage*. Burlington: Butterworth-Heinemann, 2003.

Stiggins, Richard. "Assessment crisis: The absence of assessment for learning." *Phi Delta Kappan* 83(10), (2002): 758–765 https://www.pdkintl.org/kappan/k0206sti.htm (access date November 5, 2015)

Wenger, Etienne, Richard Arnold McDermott and William Snyder. *Cultivating communities of practice: a guide to managing knowledge*. Boston: Harvard Business Press, 2002.

Wenger, Etienne. *Communities of Practice: Learning, Meaning, and Identity*. Cambridge: CUP, 1998.

[www 1] http://webcef.eu

[www 2] http://filolog.uni.lodz.pl/krakowian/course_files/missingcd.iso

WZAJEMNE GRUPOWE OCENIANIE JĘZYKA MÓWIONEGO W SIECI ZA POMOCĄ ELEKTRONICZNYCH SYSTEMÓW WSPIERANIA PODEJMOWANIA DECYZJI

Streszczenie

Chociaż nacisk na rozwijanie umiejętności mówienia w języku obcym według Europejskiego Systemu Opisu Kształcenia Językowego (ESOKJ) wydaje się być powszechnym trendem w całej Europie, miejsce i uwaga jaką poświęca się umiejętnościom mówienia w znakomitej większości szkół europejskich jest mniejsza niż na to zasługują. Składa się na to szereg rozmaitych czynników, z których jeden dotyczy faktu, że umiejętności językowe nabywane są powoli, wymagają czasu i wysiłku, a jednocześnie brak wyraźnych efektów wydatkowanych nakładów pracy, oraz wreszcie nie sprzyja temu taka konstrukcja materiałów do nauczania, gdzie zwraca się uwagę na takie składniki i komponenty nauczania, które są bardziej namacalne. Aplikacja internetowa przedstawiona poniżej stanowi ciekawy sposób na uporanie się z tym problemem, pozwalając nauczycielom na wspólne, grupowe i wzajemne dokonywanie ocen umiejętności mówienia w środowisku internetowym. Głównym założeniem tej aplikacji jest networking między nauczycielami z rozmaitych instytucji edukacyjnych w celu stworzenia wspólnoty eksperckiej, która pozwoli na walidację ocen języka mówionego i stanowić będzie dla nich narzędzie za pomocą, którego będą dzielić się materiałami audio i wideo zawierającymi wyczerpujące oceny wypowiedzi ustnych w języku obcym. Ten artykuł ma na celu zaprezentowanie tej aplikacji, jej wewnętrznej logiki oraz podstaw procedury tworzenia adnotowanych próbek języka mówionego w środowisku internetowym przy udziale społeczności eksperckiej, której wysiłki stanowić mają gwarancję trwałości funkcjonowania takiego systemu oceniania języka mówionego.

Słowa kluczowe: ocenianie mówienia, ocenianie przez sieć, testowanie wspomagane komputerowo, elektroniczne systemy wspierania podejmowania decyzji, banki wiedzy