Extending Babelium Project with a revolutionary method for preparing Second Language Speaking exam using Web 2.0 tools in a collaborative way

Juan A. Pereira, Inko Perurena, Silvia Sanz, Julilán Gutiérrez
Dept. of Computer Languages and Systems
University of the Basque Country
Spain
juanan.pereira@ehu.es; inkoperurena@gmail.com; silvia.sanz@ehu.es; julian.gutierrez@ehu.es

Abstract: Second language oral exams are hard to pass. The students face a quite stressful situation as in real life contexts. Nowadays second language acquisition is a paramount skill in education, specially in non-English speaking countries. At the same time, students from all around the world use collaborative web 2.0 platforms creating huge social networks in their own languages. Our study lies in leveraging this collaborative tools and platforms in order to build an application that will enhance the speaking skills of their users. This paper explains how this necessary application has been designed.

Introduction

For many years, languages have been taught as a Foreign Language where the students learn the language in order to use it with any other target-language speaker, or as a Second Language where students used to live in a target-language community. (Harmer, 2007) points out that new idea that "it makes sense to blur the distinction and say, instead, that whatever situation we are in, we are teaching ESOL (English to Speakers of Other Languages)". Nowadays Official School of Languages' exams include some sort of speaking phase to measure the students' skills to communicate in their target-language .

The exam has a few simple stages: first off, the teacher, 20 minutes before the oral test distributes a couple of pages with photographs among the students. Each sheet refers to a specific theme, and the photographs are related to that topic. For instance, "Street gangs", "Home Schooling", "Social networks" could be some of the topics. During those 20 minutes, the student must prepare a "speech" of 5 minutes per subject. In order to help the students, apart from photos, the distributed paper-sheets include some questions that can serve as conversation catalysts. Once the scheduled time ends, the student goes to the evaluation room, where he or she should speak for 5 minutes on each subject. Actually the first topic is a monologue, and the second, a conversation with another student who is examined simultaneously.

The previous context describes a real stress situation. In spite of being subjected to this kind of test you could still state that it is going to be something simpler than the real life's challenges. In real life, there are only a few seconds for answering, not 20 minutes. If the student wants to learn how to communicate in a foreign language in real life, he or she must respond rapidly to stimuli (to a question from a third party, to a need of the student himself). However, it is clear that the pupil is actually learning the language, so it would be interesting to offer him some assistance while responding. In fact, in real life situations native speakers can - and often do - help the non-native to express.

The question is: can we simulate an environment similar to the described exam? And one similar to the real life? Perhaps a mixed environment it possible to train a student in that mixed environment to ensure that he or she will talk for 5 minutes at the right level of expertise? (A1, B1, B2, C1). Can the students train daily and self-assess their progress? Could the potential system tell when the students could take the exam with some warrant to pass?

The project that we have implemented attempts to resolve all the above questions and provide some measures and issues for reflection about future developments.

Babelium Project. An Overall Vision

Babelium Project (babeliumproject.com) is an open source application that tries to offer language students the possibility to improve their speaking abilities in a comfortable way, without setting up a meeting with any native or foreign person and with no way for fear about pronunciation and understanding problems. Even more, this system offers users the possibility of selecting their favourite topics and/or real life situations they need to practice or improve.

The project (Sanz, 2009) is based in the use of multimedia videos with real life conversations between two or more people. Users/students can watch videos in native language with subtitles (in various languages). The objective is to take the role of one of these people in the video and practice speaking when his turn comes. Users can record their conversations and publish them in order to be evaluated. The evaluation is carried out collaboratively by other users which native language is the one spoken in the video. These referees must assess the performance of the student. Also it is possible to send feedback to the student adding commentaries about the assessment. This is the basic loop for speaking learning. Users collaborate not only in the evaluation phase but also uploading new videos and creating the

subtitles for those videos, which in the case of non-native users is another way of learning that implies listening and writing.

Following there is a description, step by step, about how users interact with Babelium Project for improving speaking skills in a collaborative language learning environment:

1. Users upload interesting videos with short real life conversations between two or more people. These videos can come from YouTube or other sources, as well as from personal videos from users of Babelium Project.

2. Native users (or those that master the language spoken in the video) add metadata and subtitles to the video. Metadata consist on: difficulty level, keywords, roles, and so on. Subtitles are added for each role (people who talk in the conversation) and for each moment (from second X to second Y) that role talks. When various subtitles are available for the same conversation moment a list ordered by most popular ones (more voted) is shown. This work is published for helping the rest of the community to learn. Besides, the user can vote (one time) for the subtitle he thinks fits better with the video conversation.

3. A user that wants to practice speaking a second language looks for videos using some keywords he is interested on. The system returns him a list with those videos that fulfil his search.

4. This user can watch the selected video including metadata and subtitles created for it by other users. The more voted subtitle is the one that the user watches at the same time with the video. If various languages are available for subtitles, they can be changed in real time.

5. The next step is to select a role (person) of the video conversation. This is the role that is going to be played by the user who wants to practice speaking.

6. Video starts reproducing. When the selected role turn comes, audio channel stops and video (image) channel continues reproducing. This is the point where user must start talking as if he were that person. The system will record his speaking. This process is carried out every time that role takes part in the conversation. If user wants, he has the support of subtitles (like a karaoke).

7. User can check his recorded-speaking and change it as many times as he wants.

8. Once he is satisfied with his video-exercise, it is time for publishing it for evaluation phase.

9. Evaluation phase is developed collaboratively by a number of referees, who are system users that are native or master the language spoken in the video. The evaluation task consist on reviewing a speaking exercise and score it following some items () in a 1-5 scale. Also, and maybe more important, it is possible to send feedback to the speaker about pronunciation, vocabulary or other help a referee considers for speaker to improve or just for congratulating him for the performance. Feedback can be sent in text mode, with a voice message or with a video message.

Participation in the system is controlled using a credit system. The operation is simple: users who collaborate (in any way) with the system receive credit and users that consume any resource of the system pay with credit. It is, as well, an ordinary manner of supporting collaborative environments avoiding user abuses. In Babelium Project, users receive credit by uploading videos, labelling them, editing subtitles and evaluating other users, and consume credit when they want to be evaluated. It is also possible to consult the evolution of credit along different periods of time (las day, last week, last month, last year) and another interesting statistics that show the evolution of users and their interaction inside the language community.

Preparing the Environment. Simulating the Real Life's Pressures

We constantly work for improving and expanding Babelium. In the last months, we are interested in simulating the environment of the real exam since its inception. So, which is the origin? In general, the issues addressed at the oral exam are chosen by the teacher and are commonly related to topics of interest or trendy news of recent months. The same as what happens in typical real life conversations: people prioritize the issues in the news. This does not mean that there are no deeper discussions of more complex issues but we want to focus on simulating the exam environment. That is why we'll need to have a list of possible topics for consideration.

What are the topics of interest that could be chosen for discussion in the exam or as a conversation in real life? People from Newsy.com have wondered the same thing, and to get the answer have used multiple sources (television, newspapers and online news sites). They analyze this items on a daily basis, searching for the most popular topics of conversation in the world (e.g., in January 2011, Wikileaks appears as a hot issue, as well as the ban on smoking in public places in Spain, the rise of electric cars, or electronic cigarettes).

In addition to the issues on the rise, it's also possible to find listings of common topics of discussion throughout the world, offered by different language teachers altruistically (The Internet TESL Journal and English Study and Learning Materials web pages are some examples).

Creating an Ad-Hoc Video Exercise

Once a topic of interest X is selected, another objective will be to help the student in his training of oral expression. With this objective, Babelium offers the possibility of creating a video from images related to the topic X. Once the images are selected, Babelium will propose the teacher i) a series of words related to X, ii) a set of catch phrases for that topic and iii) a set of conjunctive adverbs. These 3 components will be inserted into the final video (built from the images) every Y seconds.

After the student has read the topic, his aim will be to describe the images that will be shown in the video. To support the student in this effort, he/she will receive help in the way of superimposed words and catch phrases (same words and phrases that the teacher selected from Babelium suggestions in the previous step).
In the following sections we discuss in detail the points we present.

Selecting images

After the teacher fixes the topic to be discussed, Babelium will search for images related to that topic. To do this, the system will use Google Image Search (images.google.com). Among the images that it finds using the script *goosaver*:

```
#!/usr/bin/perl
# goosaver.pl
# Downloads images from a Google Image search for a screensaver.

use strict;
use WWW::Google::Images;

# Take the query from the command line.
my $query = join(' ',@ARGV) or die "Usage: perl goosaver.pl <query>\\n";

# Create a new WWW::Google::Images instance.
my $agent = WWW::Google::Images->new(
        server => 'images.google.com');

# Query Google Images.
my $result = $agent->search($query ,
                    limit => 25,
                    iregex => 'jpg'
);

# Save each image in the result locally, with
# the format [query][count].[extension].
my $count;
while (my $image = $result->next( )) {
  $count++;
  print $image->content_url( ) . "\\n";
  print $image->save_content(base => $query . $count) . "\\n\\n";}
```

the system will only select those that are really photographs (thus, images that are not drawings) because usually the former are of a higher quality than the latter. To being able to sort the images in these two groups (pictures vs. illustrations) Babelium uses the *imagemagick* tool with a patch (Banerjee, 2008). After several tests we have obtained a heuristic that indicates that the -overall- kurtosis and skewness of all of the image's channels, when both negative, reflect that it is a drawing and not a picture.

Selection of phrases and useful constructions for oral expression

Once the images that comprise the video are selected, Babelium will present the teacher with useful words and phrases aiming to help the student to explain what he/she sees in the video. Specifically, the teacher must select from among a group of conjunctive adverbs, sentence connectors (moreover, thus, hence ...) and helpful expressions to begin or end a sentence.

Even though the teacher selects a great number of verbal constructions, Babelium will not show the whole set to the student. Instead, the system applies a filter based on different parameters: the length of the video exercise, the student's past-records, language skills and a spaced repetition algorithm (Elmes, 2010).

Spaced repetition algorithms try to determine the number of repetitions that the student must successfully complete before changing the phrases or words of support.

Specifically, if the student has already successfully used several times in the past some of the phrases and words that our system proposes, he or she will be tested with other constructions not seen before.

Simulating the pressure of time

Once the images and the texts are selected, our system will automatically build a video using a simple Bash script that makes use of the multimedia framework (MLT):

```
$ melt ./.all.jpg out=2000 ttl=100 -filter luma:%luma01.pgm
luma.softness=0.2 -repeat 2
```

In the final video, each image is shown for Y seconds. Helping texts that the teacher selected and Babelium filtered will be shown as closed captions (depending on the student's level and his will, subtitles will be shown or not). In order to maintain the student's concentration the maximum video length has been fixed to 5 minutes.

Video-Exercise Recording

The student should speak for no longer than 5 minutes explaining what he/she sees in the proposed video. If possible, the student should not be silent and have little or no hesitation. All student's speeches are recorded in our system's server. Recordings can be made using a microphone and / or a webcam. As it is stated in (Frey et al., 2010), the use of a webcam can help improve the flow and minimize the number of errors in second language speaking (as opposed to a simple audio recording).

Measurement and Analysis of the Recording

After recording has stopped, the student can publish it in our system so that it can be evaluated by others in a collaborative way. We currently use the open source Babelium application described in section 2. Collaborative evaluation allows evaluators to issue a verdict on the students' performance, both in terms of a numerical score based on parameters like intonation and stress, pronunciation, rhythm, fluency and spontaneity, and as a set of text and video comments (the evaluator can leave a small video comment with his/her opinion or advice for the future).

We have added some new tools to the Babelium project that help students to get almost immediate feedback about their performance and to know some points of suggested improvements on their work. To determine these potential points of improvement, we have added to Babelium the following measurements: a) detection of silences in the audio of the video-exercise (the student is quiet when he/she is supposed to talk), b) the number of words and richness of language (this new tool compares the words used by native speakers vs. words from a non-native speaker), c) number of phrases / words suggested by the video that the user has used. Next, we describe in more detail the

aforementioned points:

Detection of silence

Our tool uses Sphinx 3, a speech recognition system (CMU Sphinx, 2010). We decide that there is a silence lapse in the X frame of the audio if it meets the following criteria: (1 frame = 10 ms).

1. The average power of the interval [x-7 , x+7] is 8 dB lower than the average power of the interval [x-200, x +200].

2. The power range of the interval [x-7, x +7] is less than 10 dB.

These criteria are the same used by the creators of Sphinx in some of theirs experiments (Placeway et al., 2005).

Number of words and richness of language

We want to compare the number of words and the richness of vocabulary used between a learner-user and a native-user in a certain language. The result will be composed of two word-clouds (a cloud to the native user and one for the apprentice) like this:

Figure 1: Example word-cloud automatically generated by the Wordle.net service

In order to generate word-clouds, we must first extract the audio recorded by the user. We have tried different speech-to-text application alternatives: the open source Sphinx system - does not work properly in free transcriptions-, the Spinvox.com web application (it's API has severe limitations on the number of words per month) and finally, the automatic transcription system Google Transcript, available through YouTube's website. This system allows us to upload any English video and get an automatically generated transcription (using their "Machine Translation" option). It has two drawbacks, the first is that, as mentioned, they only accept English videos. The second is that the result is not immediately available - you have to wait 48 hours.

Number of suggested words and phrases vs. those that have been used

Using the transcript made by the Google's service, we contrast the number of words and phrases that were suggested to the user against those which he/she actually used. This comparison will serve us for building a personalized spaced repetition schedule, as we've mentioned previously (we want to know if the user masters the words and constructions suggested or if it's necessary to insist on them again in the future).

Conclusions and Future Work

Using web 2.0 tools and collaborative work we have been able to create an open system to allow second language oral practice. This system builds on the code developed under the Babelium application.

Based on this application we've proposed the introduction of a module that's specially targeted to prepare oral examinations that are similar to the ones the official schools of languages use in their assessments. To meet this end, we have shown the process that a teacher would follow in order to build a video-exercise that includes images related to the topic of interest as well as words and phrases that will help the student.

The assessment of the video-exercises carried out by other users are done in two ways: through collaborative work (other system's users assess the learner) and through heuristics (obtaining the video transcription and performing various measurements and analysis to provide a quick approximation to the recording quality of the exercise).

In the future we want to delve into an aspect of the assessment that has not been addressed so far: self-evaluation. We think that a motivated learner would be able to self-evaluate his or her response, but we do not know the degree of objectivity / subjectivity of such assessments. Does the student self-assessment correspond with third party evaluations? Are students more critical with themselves than external evaluators? Or rather are they more lenient? We might also take more measures in order to find out a potential correlation between the number of minutes or hours spent on the application and the degree of confidence in the users' self-evaluations (are the users who spend more hours better self-evaluated?). Finally, we want to know whether spending time using the application really helps in improving the received evaluations. For example, most active users, in time, would receive better evaluations? Do they really improve their oral skills? Does our system have any influence in this improvement?

References

Banerjee, Amartyo (2008). *Patch to calculate kurtosis and skewness of image channels using ImageMagick*. http://goo.gl/Ps8ef

CMU Sphinx. http://cmusphinx.sourceforge.net/

Elmes, Damien. *The Anki Spaced Repetition Algorithm.* http://goo.gl/YetUW

English Study and Learning Materials. http://www.eslgold.com/speaking/talking_topics.html

Harmer, Jeremy (2007). *The Practice of English Language Teaching*. Harlow: Pearson Education Limited.

P. Placeway, S. Chen et al (2005). The 1996 HUB-4 SPHINX-3 System.

Sanz, S., Pereira J. eta al. (2010). T*aking Advantage of Web 2.0 and Video Resources for Developing a Social Service: Babelium Project.* 10th IEEE International Conference on Advanced Learning Technologies.

Sanz, S., Pereira J. et al. (2009). *Practicing L2 Speaking in a Collaborative Video-Conversaion Environment.* ED-MEDIA 2009 -World Conference on Educational Multimedia, Hypermedia & Telecommunications

The Internet TESL Journal (2011). Conversation Questions for the ESL/EFL Classroom . http://iteslj.org/questions/

Timothy J. Frey, Abby L. Houlton & Elizabeth D. Gruis (2010). Learning Connections, Volume: 38, Issue: 3.