

# Evaluation Framework for Text to Speech Modules: the RES Framework

Eduardo Luis Gaona<sup>1</sup>

<sup>1</sup>Jack Baskin School of Engineering

<sup>1</sup>UCSC

<sup>1</sup>Santa Cruz, USA

<sup>1</sup>eddielgaona@gmail.com

*Abstract*—This paper will be submitted to the 13th International Conference on Emerging Technologies.<sup>1</sup> The URL to the conference can be found here.<sup>2</sup>

First paper I am basing mine on.<sup>3</sup>

Many students suffer with learning disabilities and usually it holds them back in education. Many different technologies are being developed to help out these students because it is an issue within our society. Computer assistive technology is a powerful tool that will help out the community[6]. Text to Speech implementations is one way on closing the gap between students with disabilities and education. Even though devices are being created at a fast pace developers are having a trouble knowing if their devices are efficient. The way these devices are being tested is manually. Manually means that testing kits are sent out to the developers and they must return their data, which in return does not keep the test data 'secret'. Also, the connecting of different developers' modules can not be handled without an exchange of software to be integrated locally. No one wants to share their code. This presents a real issue between developers. That is why the framework Remote Evaluation System(RES) is being created. The RES is based on clientserver architecture. The purpose of the framework is to minimize the sharing of code between developers and to better evaluate their Text to Speech (TTS) systems. Also, a potential goal of this software is to introduce different ideas to developers and with those ideas develop a better device. Different developers worldwide are able to test their modules into the RES client and receive evaluations from the RES module servers. Also, the RES system is able to separate the different modules so developers can work on them one by one instead of all at once. A split into various modules has the advantage that the developers of an institution active, can concentrate its efforts on a single module, and test its

performance in a complete system using missing modules from the developers of other institutions. With that being said, developers can better develop their modules by their module being tested and evaluated. Also, this opens up new ideas within the developers leading to different approaches of TTS modules.

## I. INTRODUCTION

An evaluation framework has been established for speech synthesis technology with the ECESS consortium (European Center of Excellence in Speech Synthesis) in cooperation with the EU-funded project TC-STAR. The most important element is the separation of different modules such as, text processing, prosody generation, and acoustic synthesis modules. These modules are separated in order to be tested thoroughly. Developers have a better understanding on what is going on and on top of that create better technologies to be used in the real world within their institutions. The architecture consist of three components; the RES module server, the RES client, and RES server. The RES module server encapsulates the TTS modules. RES client which sends data to and receives data from the RES module servers. Finally, the RES server which connects the RES clients and RES module servers, and organizes the flow of information. Using this RES framework developers and people who study these modules can access the TTS modules via the internet. There is no need for face to face intervention and it is no longer required to download each module locally. Also, the test data nor the results have to be handled manually. Furthermore, developers can test their TTS modules with other

<sup>1</sup>13th International Conference on Emerging Technologies

<sup>2</sup><http://www.cust.edu.pk/icet2017/>

<sup>3</sup>Rojc, Matej, Höge, Harald, Kačič, Zdravko. Remote-based text-to-speech modules evaluation framework: the RES framework. *Language Resources and Evaluation*.44, 4 (2010), 371-386.

TTS modules by other developers in order to thoroughly test their performance. One the major flaw in the RES program is developers do not want to spend to much time integrating their modules into the RES servers. Thus, the *Unforma* RES tool has been developed. This RES tool is to help ease the embedding of a module into the RES module server, which allows for an easier way of constructing those data format parsers able to convert proprietary data formats into the RES system data format. Another tool was created that enabled the generation of numerous RES system's task configurations needed for evaluations or testing of various TTS modules and systems from different institutions called *ProtocolGen*. The RES architecture has the ability for evaluation of arbitrary software components. The end result of this new technology is to help all developers around the globe better understand their TTS modules. These TTS modules can definitely be used by students with disabilities. Many students suffer from learning disabilities and the TTS can help each of them become better students as well as get a proper education. The main purpose of the RES system is for different developers to work together in order to get more efficient TTS modules. Also, with the RES system it will open up more possibilities and different ideas for TTS modules. TTS synthesis is the process of converting human readable text input into audible speech output. It involves several steps including text and linguistic analyses, letter-to-sound conversion for pronouncing each word in context, and, finally, speech synthesis to create the speech waveform.[7]

## II. FUNCTIONALITIES AND USES OF RES

### A. RES Components

The RES architecture consists of several RES clients, the RES server, and RES module servers encapsulating the RES modules. This setup can be seen in Fig. 1 below. The RES server is responsible for interconnecting the RES clients, and the RES module servers. All the different components of RES are connected with the internet via TCP/IP and UDP connections. Thus, all TTS modules are accessible through the TCP/IP network. First, developers must have to install their TTS module locally, embedded in a RES module server. The users of the program must download the RES clients locally. The client is is equipped optionally with an RTP player



Fig. 1. Functional architecture of the RES

for testing audio signals. The RES server must be installed by the 'administrating' institutions. The server talks to many RES clients at once, allowing the RES module servers to communicate with several RES clients at the same time. When performing evaluations, or testing developers simply select the desired TTS modules via RES clients and give corresponding input for the selected task. The input given by the user is then automatically transferred within the RES of the selected ECESS TTS modules, and their generated output is returned to the RES client.

RES users are able to use this technology to perform numerous testing and evaluation tasks. These tests can be tested using different RES components' architectures and their behavioral specifications. Each of these tasks to be performed starts by selecting those configurations of developers' TTS modules suitable for the desired task followed by the executions of certain tasks to be done by a TTS module. All the communication protocols used are additionally drawn in Fig. 2. It can be seen that the RES clients open RTSP (Real Time Streaming Protocol) sessions with the RES server, which are then closed after the requested task performed by the selected developer's RES module server is finished. The RTSP protocol is based on the TCP/IP, a secure, connection-oriented protocol.[2] Therefore, there is no need for the RES client or the RES server to implement any additional error-correction mechanisms. Within the RES system, RTSP defines packets content and packet exchange sequences between RES clients, and the RES server. These packets contain the MRCP content that is to be exchanged between the RES client and RES server depending on the MRCP protocol. The MRCP is used to control speech synthesizers and recognizes in order to provide speech recognition, and to stream audio from a common location to a user. The RTSP/MRCP protocols are only used between RES

client and RES server. After the client and server stabilize a connection, the RES server dedicates a special thread and the a connection is made with the different module servers. The special thread is important because this allows the RES server to perform many and different tasks at the same time. This thread is open until the job is finished and once it is done the results are sent back to the user that inputted the request. Also, the RES server acts as a mediator between any data exchange between and RES server and the RES client. Furthermore, all text data exchanged in the RES system are written in ECESS format that is later compatible with the TC-STAR data format.

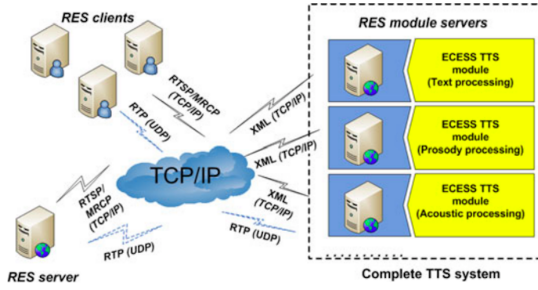


Fig. 2. RES system and related protocols

### B. Use and Installation of RES Components

Different configurations of the RES system can be seen in FIG. 3. It can be seen, the RES system needs three configurations to perform evaluation or test/development tasks. "Partner I" is dedicated to people that evaluate needing only RES clients locally. "Partner II" includes a RES client and a RES module server. This partner is made for testers and developers who wish to improve their TTS modules by comparing their results to results from other developers. Finally, "Partner III" only contains a RES module server. This partner is for developers who wish to be a part of the evaluation and who want to make their TTS modules available to other developers, but have no intention on testing or to run other developers TTS modules.

The RES server is installed by only one institution, which is also responsible for administrating the RES system. The administrator is then in charge on making a list of all the available RES module servers made available by the developers. This list is then sent to the RES clients and the RES users

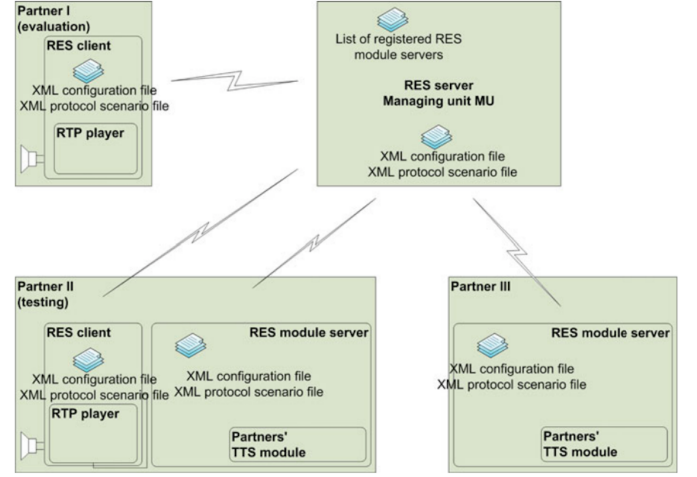


Fig. 3. Configuration to use the RES system

are free to select between the different modules. The users can download a RES client and server on any platform they wish because all components of RES are of pure Java application (Linux or Windows). RES client access to reach the rest of the RES components, regarding TCP and UDP traffic, is setup in XML configuration file. In this way RES users are free to choose between any RES module, running a specific TTS module. If the user wishes to run a module from another developer, they have to configure their IP/UDP port for RTP protocol in the RES client XML configuration file. Installation of the RES client is simple, since the RES users just have to copy the software package into some directory. After, they have to choose the desired RES task from the task list provided. The 'task list' identifies those sets of XML files describing the needed behavior of RES modules. They then input data as specified by the evaluation campaign. The input is then moved via the RES server to the specified RES module server, where it is stored in a predefined file. After, the RES module runs the TTS module or the script specified in the XML configuration file. The TTS module or script then outputs the data into another predefined file, also in XML, and the RES module server takes care of sending the data from RES server to the RES client.

### III. EMBEDDING A TTS MODULES INTO A RES MODULE SERVER

TTS modules developed by different developers will lead to different ways to input and output data.

These formats will then differ from the format required for the RES system. The translation of the different models to adapt into the RES system can require great effort and enormous amounts of time. Thus, a solution had to be created in order to speed up that process. If a solution had not been found many developers would have not been able to make their TTS modules available which would defeat the purpose of the RES system. The solution found is in order to translate the format into RES is RES should be able to make this fix internally. For each TTS module made by the developer to be executed by the RES module server, two Java parsers must be written by the administrator, one for the conversion and the other for the conversion of the RES system to the developers format. The Java framework that is used is JavaCC. Once this is completed the RES system is ready to be used. These two Java classes will automatically change the format for each RES client. Fig. 4, *Unforma*, shows how this implementation works. The Unforma tool was made to make the development of Java parsers as easy as possible. The Unforma tool is composed of more compilers. First, the developer specifies the parsers name, which is traditionally composed by the name of those data formats.

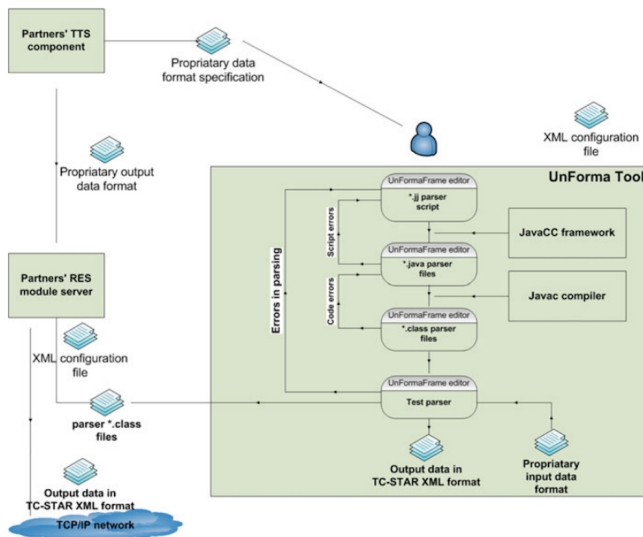


Fig. 4. *Unforma* tool - functional architecture

After, the administrator writes a parser script. The purpose of the JavaCC compiler is to create the required Java classes. For conversion in the opposite way, a new parser script has to be written and additional Java classes have to be generated. In this

way the administrator checks if the JavaCC parsers compile with no error. Then the administrator needs to check if the conversion was made in the correct format. If it is not then the administrator needs to fix the parser and recompile the entire Java class and do the entire process over again. When the entire program compiles with no errors and the parsers output the correct information, the generated parser can be included into the developers' RES module server. Only additional entries are needed in the XML configuration file. In this way, it is unnecessary to deploy a new version of the RES module server, only the new generated parser classes should be given to the developers.

#### IV. IMPLEMENTATION OF NEW TASKS BY THE RES SYSTEM

An issue the RES system has is that it the RES components need to be able to cover all the different scenarios developers and testers put into the system. With this being said, it is impossible or inefficient to hard code scenarios into the system. This can all lead to nightmares and confusion all across the system. The RES framework must have a way to take in all the different information. Thus, all the RES components have been implemented with finite-state engines using the UniMod framework. Each module performs actions to reach a goal, meaning there is a sequence of steps the RES system takes. Each task can be described by a set of states, as shown in Fig. 5. Transition between states are triggered by events. Additionally, 'guard' functions can be used for control if all the conditions are met before the transition can occur. If some conditions are not met the RES system turns into an error returning back the start state. Such graphs can be drawn offline, then rewritten in the XML language in proprietary data format, and added as a new XML protocol scenario files to the RES modules. The RES system is able to learn new scenarios. Developers and testers are able to draw graphs in XML languages and implemented into their RES system. The system then takes the new scenario into consideration in the next run.

Using this approach can ensure flexible and fast configurations of all the RES modules and even remote specifications. Also, no task or behavior is hard coded into the RES system. Instead of the hard coding, the RES system is made up by human





and sending data into the RES server, the RES server automatically sends data to the selected text processing RES module first. The results are then sent to the selected prosody processing RES module server. They are then finally sent to the acoustic processing RES module server. Generate audio data are transferred at the end, via the RES server to the RES client, using RTP protocol. By using this RES system configuration, developers are able to test their module or algorithm, by selecting a RES client, their module encapsulated in the RES module server and use other required modules to compose the complete TTS system from other developers.

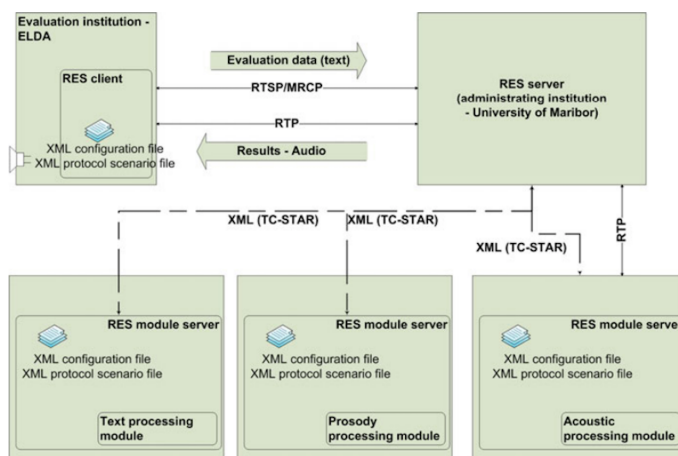


Fig. 7. RES system configuration for remote ECESS evaluation/testing of a complete TTS system

Furthermore, the checking of selected IP settings should be performed, before they are being used and saved by the RES system to ensure that each partner interested files work with no issues within the system. If one or more modules did not work the RES system would not stop running. The RES system would disconnect that module and no output would be returned to the user. The overall purpose of the RES system is for people to collaborate in development, testing, and evaluating of different TTS modules.

## VI. USER EXPERIENCE

Partners' user experience feedback has been overall positive. Initially, there were some issues with proper configurations of the necessary communication infrastructure at partners' sites. Overall, the RES system enables fast and easy integration of the developers modules, regardless of their operating

system, and the possibility of automatic input/output data format conversion seemed to be very useful. There were issues for the user to work around firewalls, where some extra knowledge about setting the a proper IP address and incoming and outgoing ports were needed. Manual configuration of these IP setting for different users is needed. In order to ease the pain of these IP a GUI will be developed to automatically set all the necessary configuration data. Overall many of the users saw this work as having great potential and it could become something great. The easiness of the user interface is powerful and new updates are being developed to further improve the RES system.

## VII. WHY IS TTS IMPORTANT?

One of the major reasons this system exists is because there are many people out there who unfortunately are visually impaired. The TTS synthesis technology helps people that are visually impaired because it enable the device to read text to them. TTS will help the student to better understand how to read [4]. TTS is a feature most technologies possess today, but they are yet to be improved. With the RES system many developers worldwide will be able to help each other by testing each others modules. Each module is important because it could either inspire or complete a developers work. This way of tackling the issue will lead to better software for TTS synthesis and will improve the community. Also, TTS is important because it will also help many students with their education. These students are at a disadvantage and they need to be aided with the on growing technology today. Many developers wants to help their society and with the RES system it's going to make that possible.

Computer technology has been recognized as a good way to teach students with disabilities[10]. Many students are impaired and lack some skills that others might find simple. Computer technology offers varies ways to keep the student engaged and learning. The implementation of TTS into these technologies will only make computer technology that much better to use for these students. Some students might be visually impaired and require the use of TTS technology to know what the computer is saying. This gap needs to be close because people should be given the opportunity to an education in order to reach their goals. Many smart phones

have implemented this technology into their systems already. The RES system will help further progress the technology that exists already. It is important to make this technology so people with disabilities feel a sense of independence from other humans. Upcoming technology will help them reach that goal.

### VIII. TEXT-TO-SPEECH IN LEARNING SCENARIOS

In today's society students feel the need or obligation to attend their classes or lectures. The teacher or professor was seen as the one who "knew it all" and the only way to master a subject was to attend class and listen to what this person has to say. Also, often these teachers or lecturers teach the subject more than once a year or sometimes continuously over a period of time. With emerging technologies and the significant drop in laptops, computers, and access to the internet have been questioning the role of traditional universities in modern society. The purpose is not to replace the instructors, but to better aid the students in the classroom. With the text to speech modules it can be implemented to the lecture slides as it would in the lecture. Instructors can aid their slides in order to better assist their classes. To often times lectures rely on the expertise of the instructor which would lead to issues. The instructor might forget to mention something important that might end up on an exam or homework assignment. The purpose of this implementation is to ease the work of the instructor to better serve the classroom.[8]

### IX. CASE STUDIES OF TEXT TO SPEECH MODULES

One of the important aspects to consider is choosing which classroom to experiment on. The classes that were chosen were picked strategically in order to get one class from each of the departments. Jonassen(1991) argues that if the only socially acceptable learning principle is based on objectivism where there is a defined set of objectives to attain, then why not elaborate on ways to enhance the learning and evaluation processes based on this approach.[5]

The modules or classes that were chosen can be seen below:

- History of Mauritius: From the social sciences department involving mainly declarative knowledge acquisition.
- Introduction to Management: A module from the Law and Management department being offered on distance education module to more than 500 students across faculties.
- Matrices: A unit in a distance education Mathematics module involving procedural and conceptual knowledge as well as problem-solving skills.
- English literature: A module from the Humanities department involving a blend of knowledge needed and which is delivered face-to-face.
- Computer programming: A module delivered face-to-face to large cohorts. The experiment limits to classroom lectures and exclude the practical sessions.
- Waves and Optics: A module from the Physics department extensively depending on materials freely available from the web using sources like Wikipedia.

The material or technologies used in the experiment are:

- Microsoft PowerPoint
- Next Up Technologies TextAloud™ Software (trial version)
- ATT Natural Voice™ (Crystal and Mike)
- Bytescout™ PPT to Video Scout Software (trial version)
- Available lecture notes, textual explanations and learning activities

#### Step 1: Presentation Preparation

The first step of the process is to obtain the knowledge of the professor in the class they are going to teach. The team will help the instructor get down all the information as the instructor would do in the classroom. One of the pros of this approach is that there will be a preset lecture where everything is carefully and well planned.

#### Step 2: The TTS Approach

The following step consists of using the TTS software to generate the complete lectures prepared in sound format. The sound format (Fig. 8) gives a sample Microsoft Power-Point slide, the text written (a paragraph) for self study mode, and the corresponding format of the text to be converted by the TTS software. The course designer has the ability to use this technology as many times as they desire.

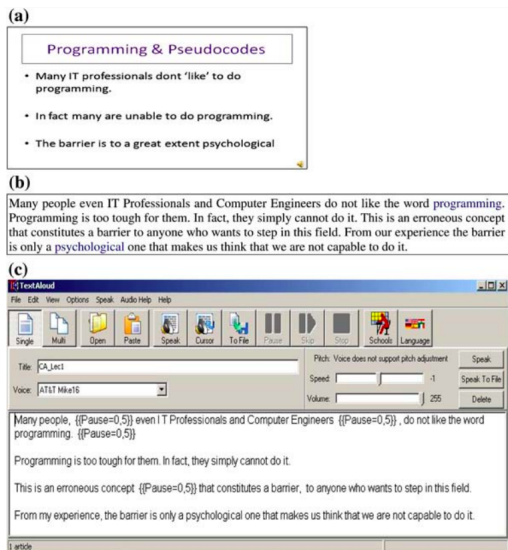


Fig. 8. Illustration of the TTS approach. (a) PPT slide, (b) Accompanying Notes and (c) TTS conversion.

They are able to play and replay the sound as many times as they wish until they are satisfied with the final version. The "To-File" option allows the user to implement the sound to specific power point slides.

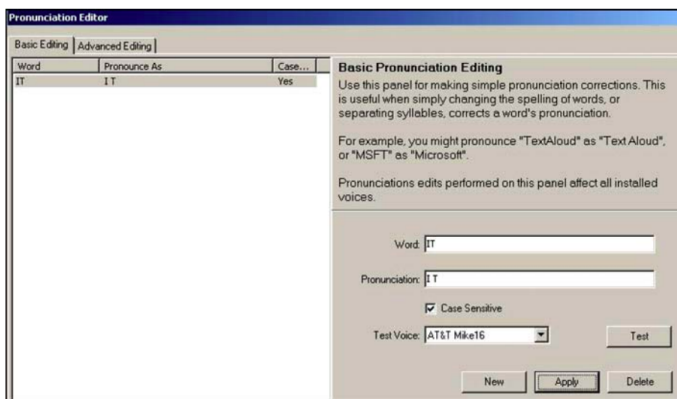


Fig. 9. The pronunciation editor

There is an issue with the pronunciation of certain words within the system. Figure 9 demonstrates the implementation of a feature in the system which helps the user pronounce words better. Using the pronunciation editor, words like 1958 for instance, need no longer be written as 19 58 in the software or can be pronounced as information technology depending on the wish of the user. Other examples are Year 1958 written as 19 58. There is an advanced feature within this feature that gives full control to the user on how each word is pronounced. To enter a pronunciation, sound it out using the strings in

the Phoneme Table incorporated in the functionality. Phoneme strings in the pronunciation must be separated by spaces. For example, the word hello is pronounced as h eh l ow. Other examples where the pronunciation editor is of use are words like words separated by / (slash), organizational (written organ I zational), IT (written I T), VCILT (written V C I L T). After the user is satisfied with their pronunciation and sound they are able to implement the files into their power point. They can implement them in the certain slide they desire and create a presentation how they would do it if they were in the classroom.

### Step 3: PPT to video (optional)

Once the sound file has been implemented into the power point, the resource can be used in that format. However, this step of conversion of the PPT to video in the process even extends the usefulness of this technique. This is a very simple step any computer user with basic knowledge or skills can accomplish. How long this process takes can vary depending on the frames per second or how much resolution the user wants in their video.

### Step 4: University Intranet Interactive TV Channel Through Video Streaming

Having different Wi-Fi points around a university campus will enhance the performance of e-courses across campus. Also, having a TV channel would be a useful component. Lectures would be able to be downloaded by various students across campus. It will be seen across multiple students, on the bus, walking to class, or having lunch with their MP4 players broadcasting the lectures. On the other hand, using a video streaming server not to stream video recording of chalkboard lectures as proposed by Canessa [3] with all the inherent complexities of such processes; but to broadcast the lectures through an interactive web based environment with interactivity features that will favor a new paradigm in the teaching and learning landscape at the University. Red5 is an open source flash server written in Java that supports streaming audio/visual. This new way of approaching the learning experience of the student. The student has more freedom to learn at their pace and repeat a lecture as many times as they need too in order to write down better notes.



## X. THE BENEFITS AND THE FUTURE OF TEACHING AND LEARNING LANDSCAPE

Persons with a visual preference tend to show a greater ability to analyze and integrate visual information, mentally convert non-visual information into visual, and show superior retention of mental images [1]. The benefits of this new approach on teaching styles are many and include:

- teaching can be carried out exactly as the lecturer wants it with minimal efforts required from the part of the lecturer in remembering stuffs, what to say, and timing of the whole lecture;
- deriving the synthetic speech recordings from written texts imply that the content of the latter can be easily be modified and recordings updated;
- replay of any slide at any time for discussion purposes;
- the presentations can be replayed any number of times;
- the technique caters for students repeating a module, especially where the failure rate is low and the module is no longer run;
- it follows from (3) that revision is easy (selected slides or whole lecture can be reviewed);
- students do not have to adapt to different lectures as far as understanding their languages is concerned;
- the classes can be better managed, if lectures are carried out in situ;
- the material prepared can be placed in the knowledge repository accessible to students and/or staff;
- the provision of portability of lectures in MP3/MP4 formats to suit mobility of students thus minimizing idle study times;
- the simplicity of the lecture preparation means that entire module lectures can be quickly prepared for the purpose without having recourse to sophisticated and expensive resources (such as video lectures);
- the online monitoring of individual students activities is easily done as the system tracks their virtual attendance automatically. By incorporating interactive activities such as journal writing, lecture summary postings and critical appraisals, students learning and progress become more manageable than ever;
- the possibility to cater for learners with varying styles and preferences to learn as well as for auditory and visually impaired learners through either audio podcasts or video podcast with text support. In some jurisdictions, this can be a useful technique to cater for related law provisions for students with disabilities and impairments.

It is to admit that the process to get a lecture online can be tedious and time consuming it will better server the classroom and the instructor. After the process on having the lectures online is completed the professor has more time to concentrate on research they are a part of as well as help out students on an online forum. This will be beneficial to both parties because the students will learn at their own pace and the instructor can make time for other activities they are a part of. Also, the student and instructor can come together when the student has a question. With the introduction to this new form of enhanced delivery of online learning, the teaching and learning landscape will no doubt change worldwide. These can be summarized as follows:

- The involvement of teamwork in the preparation of the presentations imply that, in general, a high quality of delivery will be ensured since all the lectures will be delivered soundly as per the wish of the team with no rooms for excuse. This also means that, as far as quality assurance is concerned, peer-reviewing of lectures can be carried out offline in its entirety rather than sampled and ad hoc. This also lends itself to a much better working environment for the academics.
- There will be a substantial decrease in the face-to-face component of the teaching especially when academics are called to address different cohorts of students with the same material. Academics will therefore have more time to devote to enhancements of their teaching, more time for research and consultancy and other duties. This will be also accompanied by lesser time tabling problems since many lecture theatres and classrooms will be freed. On the other hand, this can also imply increasing access to students with existing staff.
- Lesser face-to-face contact hours calls for special arrangements with students with the main

objectives to guide them, to discuss and review assignments, to discuss examination issues and other issues, and to clear any misunderstandings as far as the lectures are concerned. There are various ways that these could be entertained through regular meetings, possibly on a fortnight basis. Just like the peer-reviewing system discussed above, feedback from these meetings with students are essential to clear any bottlenecks, to address calls for upgrading lectures, to set up an FAQ system for each module, to sense the level of the intake of students, etc.

- This is also an opportunity to study new ways for students to interact amongst themselves for group work.
- Universities will be able to run programs with smaller critical number of students as the costs will be reduced. Furthermore, a real flexible and lifelong learning culture can be brought in the educational landscape as students can easily enroll on self-study modules at any time of the year without having to wait the start of semesters and new batches.
- There will, no doubt, be new opportunities for national and international universities to work together. Visits from overseas lectures can be significantly decreased if not eliminated altogether.

While research, especially from a survey of students profile at the University of Mauritius shows the oral component is very important to suit different styles of students, e-learning and distance courses often omit this component as the emphasis is laid on self-explanatory written text. [9]

## XI. CONCLUSION

RES is a web based distributed framework for the evaluation and development of TTS modules to better serve the disabled community. It is a client/server architecture composed of several components, all running at finite states. This technology is maintainable and easy to use. The user interface is made in a way that the user can easily ask RES to complete a job and return the results as well as interchange information, RES acting as a mediator. One of the most powerful tools that RES has is that it is able to run many different tasks at the same time. In this way developers and researchers are able to

concentrate their energy one module at a time. Each module then is tested thoroughly which leads to better algorithms. RES is starting to be recognized worldwide because it is easy to handle. TTS technology is something our community needs because people have disabilities. Whether it is by birth or an accident TTS is important. This will allow people to communicate verbally through software. It is difficult to communicate certain things via hand signals. This technology will help close that gap and in the near future RES will be the architecture developers use in order to create better and more efficient modules/algorithms for TTS technology. The integration of text-to-speech seems to be the approach that can help achieve most of the decisions taken without compromising the educational system but instead providing a better environment for every stakeholder. At this stage, however, the process and the technique are described but relevant policies are forthcoming. It is easy to notice the Test to Speech technology will play an important role in the education system in the future. The first steps are to test the new technology in different scenarios and soon this new form of teaching will forever change the way of learning for many students around the world.

## REFERENCES

- [1] AYERSMAN, D. J., . M. A. Individual differences, computers and instruction. *Computers in Human Behavior* 11, 3-4 (1995), 371-390.
- [2] BURKE, D. Speech processing for ip networks/media resource control protocol (mrsp).
- [3] CANESSA, E., F. C. . Z. M. Webcasting of traditional chalkboard lectures: The eya system. *European Journal of Distance and Online Learning* 2, 1 (2007), 123-125.
- [4] HALL, TRACEY, . H. C. . F. M. Computer assisted instruction in reading for students with learning disabilities: A research synthesis. *EDUCATION AND TREATMENT OF CHILDREN* 23, 2 (2000), 173-193.
- [5] JONASSEN, D. J. Evaluating constructivist learning. *Educational Technology* 38, 28-33.
- [6] KUMAR, DAVID., W. C. Computer technology, science education, and students with learning disabilities. *Journal of Science Education and Technology* 6, 2 (1997), 155-160.
- [7] LI, S. Z., AND JAIN, A. Text-to-speech (tts) synthesis. *Encyclopedia of Biometrics* 3, 3 (2009), 1332-1332.
- [8] RUGHOOPTH, S. D. D. V., AND SANTALLY, M. I. Integrating text-to-speech software into pedagogically sound teaching and learning scenarios. *Educational Technology Research and Development* 57, 1 (2009), 131-145.
- [9] SANTALLY, M. Individual instruction and distance learning: Application of learning and cognitive styles. *Malaysian Journal of Distance Education* 5, 2 (2003), 15-26.
- [10] SIGAFOOS, J. Introduction to the special issue: Evaluating assistive technology in the education of persons with severe disabilities. 1-3.