



e-ISSN: 2147-8228

www.dergipark.org.tr/ijamec

Volume 08  
Issue 01

March, 2020

**Research Article****A Mobile Application That Allows People Who Do Not Know Sign Language to Teach Hearing-Impaired People by Using Speech-to-Text Procedures****Emre Biçek<sup>a,\*</sup>** , **M. Nuri Almalı<sup>b</sup>** <sup>a</sup>Department of Informatics, Van Yüzüncü Yıl University, Van - 65080, Turkey<sup>b</sup>Department Electrical and Electronic Engineering, Van Yüzüncü Yıl University, Van – 65080, Turkey

## ARTICLE INFO

*Article history:*

Received 31 January 2020

Accepted 30 March 2020

*Keywords:*

Course Management System

Hearing Impaired

Mobile Application

Speech-to-Text

## ABSTRACT

Hearing-impaired people use sign language to communicate with each other. People who do not know sign language it is difficult to communicate with hearing-impaired people. In this study, this problem is tried to solve with speech-to-text infrastructure. A software project has been developed to enable people who do not know sign language to communicate with the hearing impaired people. The scope of the system which named "Subtitles Course Tracking System (SCTS)" is that with an application running on the Android operating system, user's speeches are converted into text (Speech-to-Text) and instantly transferred to remote a server. Speech texts which are recorded on the database of the remote server could be followed up in real-time on mobile phones and web pages by the software that has capable of asynchronous data exchange over the Internet (AJAX). Within the scope of this study, a web-based course management system has been developed and all courses in the system made accessible at any time.

This is an open access article under the CC BY-SA 4.0 license.  
(<https://creativecommons.org/licenses/by-sa/4.0/>)

**1. Introduction**

Hearing impaired people have a high level of hearing loss (pure sound loss above 71dB) therefore have little or no hearing [1]. 360 million people in the world are experiencing hearing problems. This corresponds to about 5% of the world's population. 32 million of the people with hearing problems are children under 15 years of age [2]. The most significant effect of hearing impairment on individuals is that individual communication skills with other people are not improved and language development is delayed compared to normal individuals. It is known that it is difficult for hearing impaired to follow a lesson in which the sign language is not used. Students are trained by lecturers who know the sign language in their special schools using various special teaching methods. These trainings are given either verbally or in written form, as well as in a customized language. It is very difficult for instructors who do not know the sign language to be able to teach in these schools and to communicate individually with the hearing impaired. This situation adversely affects the development of the students. In this study, it is tried to

provide a solution to this problem with developing software technologies.

A study was conducted on behalf of the hearing impaired adults to examine the communication methods and skills they use in different situations. A total of 100 hearing impaired individuals, 69 males, and 31 females, participated in this research. The following question has been asked "When you want to get information about something from someone who has never met before, which communication method do you understand better?" In response to this question, they stated that they understand better when they communicate with the sign language (60 people) and secondly with the text (54 people). According to this research, the importance of writing while communication with the hearing impaired is revealed [3]. In the recent past, many advances have been made for supporting hearing impairments with technological infrastructure. In addition to devices that heal hearing physically, many systems have been developed to support individuals in their social and educational life. One of the most important technologies among these is called

“Speech-to-Text” that enables to convert voice into text instantly.

One of the international research groups, the Liberated Learning Consortium in Canada, has conducted studies on Speech Recognition infrastructural systems between the years 2010-2012. The company has developed IBM Speech Recognition technology. In Canada, they collaborated with many universities and research institutions on disability to develop hearing impaired practices. Hosted Transcription Service (HTS) as developed a text converter application to save the previously recorded audio file [4]. In 2014, a group of mobile app developers in the United States developed a mobile app for people who are deaf or hard of hearing. This application called Transcense enables people to follow up conversations via mobile phones and to make conversations with each other in an environment where many people come together [5]. The Nuance company, along with the Dragon brand, has an important place among speech recognition systems all over the world. The company offers many paid services to individual customers and also companies use this infrastructure to develop various applications in which voice is converted and voice is used [6]. Experiments have been conducted to automatically use speech recognition technology, to translate speech during the course, and to enable students to reach later and recreate their course notes [7], [8]. The researchers also worked on automatic captioning systems for the hearing impaired and hard of hearing individuals [9]. A multilingual system has been developed for telephone and video conferencing systems using the speech translation infrastructure. The system translates the speeches of the users into text synchronously and these texts are translated into other languages instantly and a multilingual communication is provided between the speakers [10]. In a similar way, many studies based on speech recognition have been conducted not only for the hearing impaired individuals but also for other disability groups. A voice-controlled wheelchair is designed for physically disabled patients. Voice controlled wheelchairs provide great convenience for many people who cannot use manual controls [11]. A web application that teaches programming to visually impaired people has been developed. The system using speech recognition technology that operates on the basis of voice commands is also an application based on semantic analysis, user interaction and braille engine [12]. Various studies have been carried out using Speech-To-Text infrastructure and quite successful results have been achieved on children with traumatic brain injury that causes memory loss, language skills weakness and rapid psychomotor disorders [13]. Previous research has focused on solutions for various disadvantaged groups using the speech recognition infrastructure as mentioned above. This technology has been used in many areas and successful results have been

achieved. This infrastructure was also used in the studies for the hearing impaired individuals. There are many studies for them in the literature. Automatic subtitle systems, call center solutions and assistive classroom applications are among the works in this area.

Unlike previous studies, this research introduces the idea of developing a mobile application and transforming the speeches into text via mobile phones and enabling students to follow these texts instantly from their mobile phones. In this way, the system can be used anywhere without the need for special classes prepared beforehand and without the necessary systems installed. Every environment in which a mobile phones exists becomes suitable for interactive lessons. In this study, a customized course management system and an interactive course environment are prepared for the hearing impaired. The system called “SCTS” consists of Android and web-based software. The speeches are converted to text synchronously with mobile software developed on the Android operating system. Speech conversions are instantly transferred to a remote server and stored in the database. As long as the lecturer continues to speak, the conversation texts, which are updated periodically and saved in the database, are printed on the screen with a web-based application that uses AJAX (Asynchronous JavaScript and XML) technology and these pages can be tracked instantaneously. With the software developed on the Android operating system, the speeches that are instantly transmitted to the web page are also reflected on the screen via the computer, projector, or smart board, and students are provided to follow the lesson in writing. In addition, a web-based open course management system has been developed to ensure that all lessons on the application can be accessed later. See also [14] reference in the context of the using Speech-to-Text infrastructure in lecture environment by using a mobile application.

The rest of the paper is organized as follows. Application Components and Software Development are presented in Section 2. In this section, detailed information about software technologies used in the developed system is given. Section 3 includes the units of the SCTS. This section contains the system usage and module descriptions and the paper is finalized with a conclusion section.

## 2. Application Components and Software Development

Many different software languages have been used to achieve the desired result for developing the “SCTS”. Two different applications both mobile and web-based have been developed. The Android application is developed using the Java programming language on the Android Studio mobile application development platform. The web application uses PHP, MySQL, SQL, AJAX, HTML and CSS.

### 2.1. System Architecture

The application briefly consists of the parts shown in Fig. 1. (The units of the application). The system uses a MySQL database on a Linux-based server. Both the mobile application and web application use a common database. The Android app was developed in the Java software language using the Android Studio mobile software development platform. The SQL class of the Java library was used for the mobile application's database operations. We used the Speech-to-Text service provided by Google for Android developers to convert speeches into the text. A separate class has been written in the Android application for connecting the Speech-to-Text server provided by Google and saving conversations converted to text instantaneously into the database. The speech data which is converting to text are saving to the database according to the flow chart in Fig. 3. (Speaking panel flow chart). Lessons can be accessed through the web-based course management system. On the live follow-up screen, the user's speeches are instantly writing on the screen, the conversation texts can be followed both from the web application and from the mobile application. Asynchronous JavaScript Technology and XML (AJAX) technology is used for instant tracking of the conversations.

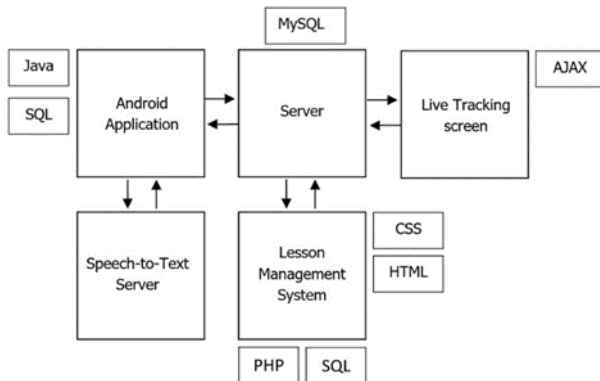


Figure 1. The units of the application.

### 2.2. Database Structure

SCTS is an application that consists of two different software as mobile and web-based. In practice, a common database maintained on a server with a Linux operating system and providing data exchange between this two software. A functional but minimal database has been created considering the application requirements. The database table attributes and relationship diagram of the table are shown in Fig. 2. (Table qualification and relationship scheme). The application has two different user types as lecturer and follower. User privileges vary according to the type of entry. The lecturer can select a course that already registered in the system through the Android application, or create a new one also lecturer has the authority to add, edit and delete courses from the

course management system. Followers are authorized to follow lessons using the Android app. On the course management system, users have the authority to view and download all available courses on the system. There are two tables in the system's database. User information is kept in the "tblusers" table and course information are kept in the "tblcourses" table.

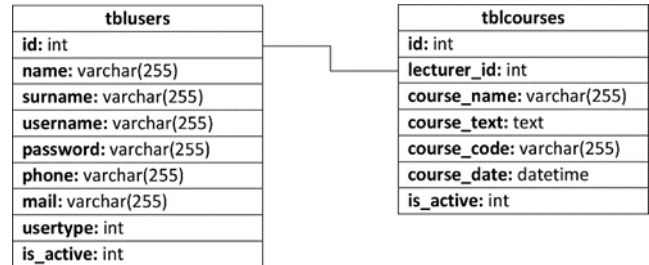


Figure 2. Table qualifications and relationship scheme.

The "tblusers" table holds the "usertype" parameter. The value of the field can be "1" or "2". If the value is "1" user is a lecturer. If the value is "2" user is a follower. The "is\_active" attribute indicates whether the user is authorized to use the system. If the value is "0", it means that the user cannot use the system, if the value is "1" the user is authorized to use the system. The "lecturer\_id" parameter in the "tblcourses" table associated with the "id" field in the "tblusers" table and thus, the course-teacher relationship is provided. The "course\_code" parameter includes the live tracking code. This code is a unique two-digit code that is automatically generated by the system which must be entered for tracking the relevant lecture. The code is automatically generated by the mobile application when the lecturer creates a new lesson which is not on the system. A user with a lecturer authority over the course management system has the authority to change this code. The "is\_active" attribute keeps the value that whether the course is active or not. If the value is "0" the course is not shown to the users, if the value is "1" the users can show the course.

### 2.3. Speech-to-Text Integration

Speech Recognition is a multi-layer pattern recognition function. Acoustic signals are processed in a certain hierarchy to form words, expressions, and cues. There may be some temporary restrictions on each of these layers such as pronunciation problems or incorrect grammar. Such mistakes can be statistically resolved in low-level layers in the hierarchy. In the upper layers, voice recognition processes performed using algorithms based on estimation. We can examine the voice recognition system in several main sections: Feature extraction, acoustic model database, dictionary, language model and voice recognition algorithms that have been tested and verified previously [15].

The Application uses Speech-to-Text service, which is offered to Android application developers by Google. A

number of procedures have been applied for using software libraries provided by Google. A number of additional functions and classes have been written for sending converted voice to the remote server instantly. The texts that transferred to the remote server can be processed in the course management system. The speeches on the mobile application continue as long as the instructor continues to speak without interruption and is instantly saved in the database as soon as the conversation is completed. The flowchart of the speech panel is shown in Fig. 3. (Speaking panel flow chart).

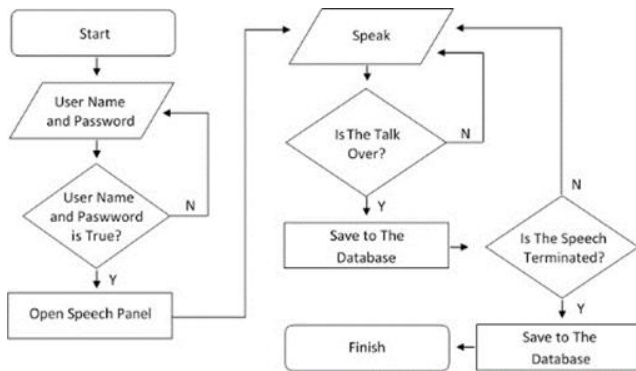


Figure 3. Speaking panel flow chart.

2.4. Ajax Integration

Web pages traditionally consist of a series of HTML pages and the parts that make up those pages. They must be reloaded by pressing a button to update a section or all of the contents of these pages. It is not possible to update a section of the page in the classic web applications; the page is completely updated in every update process. On the contrary, with Web 2.0 based on Asynchronous JavaScript Technology and XML (AJAX), it is possible to instantly update only the desired sections without needing to click any buttons, instead of updating the entire page. This not only significantly reduces the bandwidth used by the page, but also offers a wide range of possibilities for the software [16].

In the classic method, when the user sends a request to the server, the server must wait until the request will be answered. The user has no chance to send another request to the server without waiting response from the server. In the Ajax model, when the page uploads, the ready Ajax engine interacts with the user instantly. When the user sends more than one request, these requests handle by the Ajax engine and the necessary operations are sending to the server asynchronously. It may be thought that adding an extra Ajax engine to the page can make the work more complex, but the opposite is true [17]. The comparison of classic web applications with the Ajax model shown in Fig. 4. [17]. (The comparison with classic web applications and AJAX model).

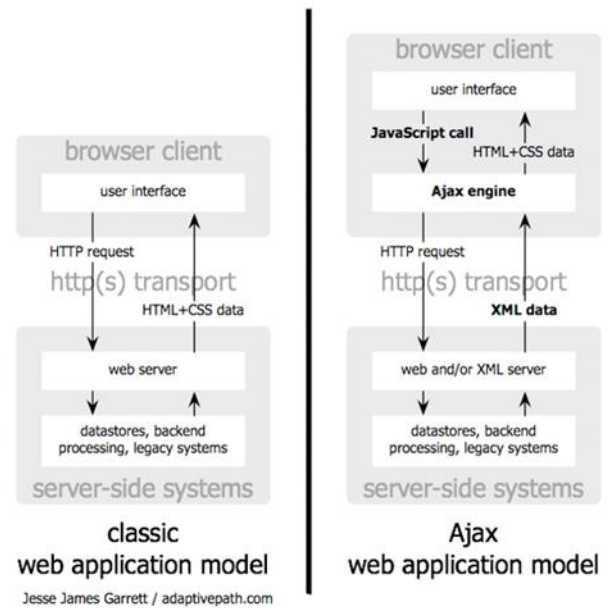


Figure 4. The comparison with classic web applications and AJAX model.

In the classic "request-wait-response" structure, the page must be reloaded for each user activity. In the AJAX framework, multiple requests are accepted asynchronously in the background and use to pre-process and update a specific section of the page when necessary. This approach provides more interactive and faster web experiences for users. Before the AJAX concept has emerged, this structure has been strongly felt in applications such as Google Suggest, Google Maps, or Gmail, Flickr, and Digg [18]. With the mobile application developed in the context of this application, the conversations are converted and instantly transferred to a remote server. These texts on the remote server are instantly transmitting to a computer or mobile application. At this point, through AJAX user can follow the speeches instantly without having to refresh the page.

3. Subtitled Course Tracking System

The main aim of the system is to enable people who do not know the sign language to be able to communicate the hearing impaired. With the developed software, the lecturer's speeches were converted into text with an Android-based mobile software and it is provided that students could follow these texts synchronously through mobile phones or a web page. It is possible to provide access to the contents of the courses by the course management system. It is provided to add, edit or delete a course by using the system. The application is created by using more than one integrated software. Teacher-student interaction of the SCTS is shown in the Fig. 5. (Teacher-student interaction in classroom).

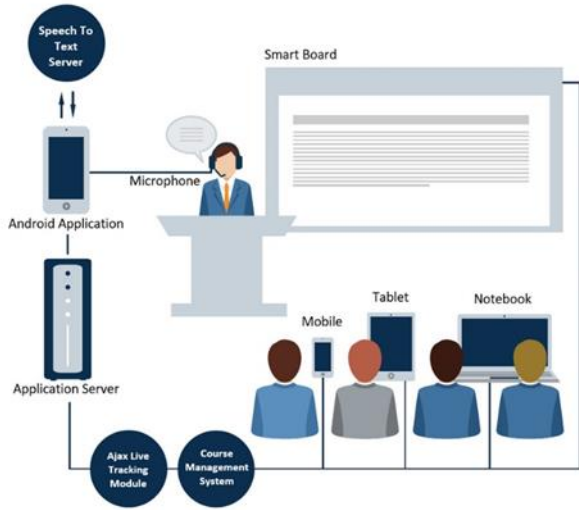


Figure 5. Teacher-student interaction in classroom.

Two different applications have been developed for the project, both mobile and web-based. Both applications are linked to a common database. The system basically consists of three parts. These are; Lectures, Course Tracking and Course Management Units.

### 3.1. Course Lecturer Unit

When the users open the application, they encountered two options in the form of "Course Lecturer" and "Course Follow-up". When users select the "Course Lecture" option, they can log in to the system by user name and password which are previously created on the course management system. After the user logs on the system can start a lecture that already exists in the system. When the users want to start a new course, they need to write the course name and the system automatically creates a course code and the users reach the lecturing panel. Course code is a two-digit unique code that is required for users to enter the system to follow the generated lesson. Fig. 6. (Course Lecturer unit and interfaces) shows the lecturing unit and interfaces.

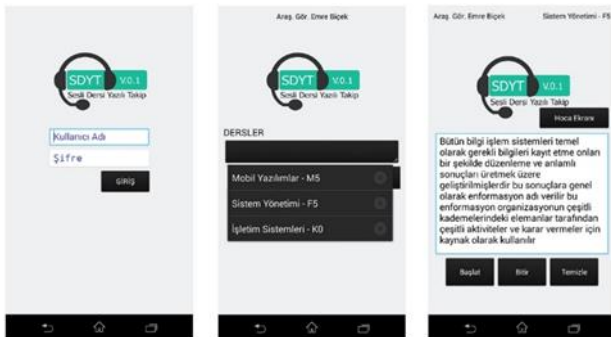


Figure 6. Course Lecturer unit and interfaces.

When the start button in the lecturing panel is pressed, the conversation panel opens and the user starts talking. The user's conversations are instantly recorded in the database instantly as long as the conversation continues.

When the user presses the end button, the conversation panel closes. When the user presses the clear button, all conversation records in the database is deleting.

Also, the user is able to edit the whole conversation texts on the panel.

### 3.2. Course Follow-up Unit

When the live track button is pressed, the course code entry panel opens. The user can join the course by entering the live tracking code of the course. Also, users can reach a lesson that has been registered previously in the system. A course tracking panel opens when the user entering the course code. In the upper part of the panel, the name and surname of the instructor and the name of the course are displayed. In the upper right corner, the red button allows the speech flow to stop. When the green colored play button is pressed, the speech flow continues again. It is possible to enlarge and reduce the text further. There is a text flow area in the middle of the screen where the user can follow the lesson. This field automatically scrolls upwards as the speeches continue and the current texts are displayed on this screen. Fig. 7. (Course follow-up part interfaces) shows the lesson follow-up part interfaces.

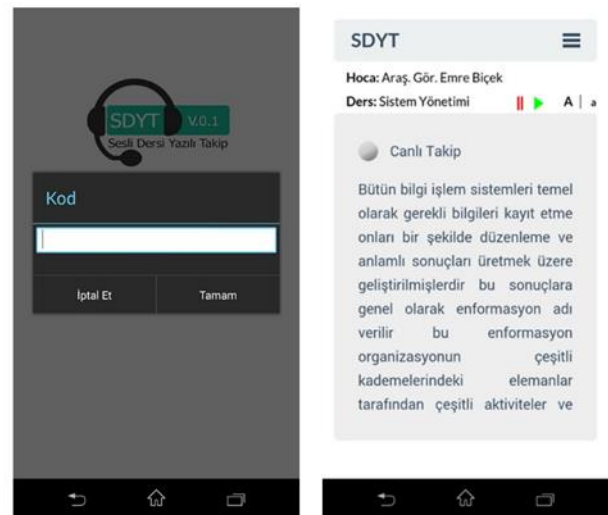


Figure 7. Course follow-up part interfaces.

### 3.3. Course Management Unit

There are two different types of entry in the web-based course management system. The first one is developed for the lecturer and the other for the followers. Both entrance types have different authorities. At the entrance of the lecturer, users can view their own courses in the system, download, edit, add new courses and delete existing lessons. Also, personal information can be updated. A follower management panel has been created for the users who want to follow and download the lessons previously made through the application. The main menu in this panel is the same as the lecturer management panel. The list on the follower's dashboard main page contains the course name, lecturer name, course date, and operations columns.

The operations column has only a download button, unlike the lecturer management panel. Followers can only download courses. They are not authorized to take any action regarding the courses. The Course management system lecturer unit interface is shown in Fig. 8. (Course management system lecturer interface)

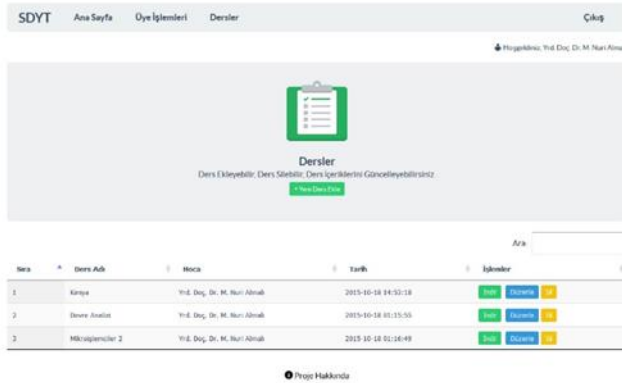


Figure 8. Course management system lecturer unit interface.

#### 4. Conclusion

Rapidly evolving technologies offer solutions that make life easier for everybody. Mobile application which are becoming very popular nowadays, allow people with hearing impaired to do things they cannot do a few years ago. SMS, e-mail and instant messaging applications making it much easier for people with hearing impairments to communicate with hearing impaired. With today's technology, the steps taken on voice and image processing are using in applications developing for hearing impaired, thereby enabling solutions to change their lives. Considering the number of individuals using devices called smartphones, it turns out that these technologies are commonly achievable.

The application developed within the scope of the study offers a solution in the field of hearing-impaired education. As known, hearing impaired children are educated in educational institutions which are given special education Teachers who are experts in their specific areas and who have mastery in sign language work in these educational institutions, but it is very difficult for an instructor who does not know sign language to be able to share the information with the students in these schools. It aims that teachers who do not know sign language can share their knowledge with hearing impairments people by using this application.

This application is an alternative for applications that have been made by using speech recognition infrastructure and have high costs. Within this project, an integrated system has been developed for instructors to communicate hearing impaired students from every area and a large number of hearing-impaired individuals have been reached by sharing these lessons over the internet. The application is not only available for the education of hearing-impaired

individuals, but also for organizations such as conferences and meetings. The right to education is the most important factor for an individual's development. It is necessary for our society to act with this consciousness and take steps towards the individuals who have difficulties in education due to various reasons. This project was initiated with this awareness and it was tried to contribute to the educational life of the hearing-impaired individuals. This system can be used in the hearing-impaired individual's schools all over the world.

#### Acknowledgment

This research was supported by Van Yüzüncü Yıl University Scientific Research Projects Coordination Unit as a Master Thesis project numbered 2015-FBE-YL059.

#### References

- [1] A. Ataş, A. Genç and E. Belgin, Odyoloji'de Kullanılan Temel Kavramlar, Pediatrik Kulak Burun Boğaz Hastalıkları, U. Akyol, Ed. Ankara, Güneş Kitapevi, 2003, pp. 35-50.
- [2] World Health Organization, Deafness and Hearing Loss. <http://www.who.int/mediacentre/factsheets/fs300/en/>, Accessed on: Oct. 10, 2016
- [3] Ç. Gürboğa and T. Kargin, "İşitme engelli yetişkinlerin farklı ortamlarda kullandıkları iletişim yöntemlerinin/becerilerinin incelenmesi," *JFES.*, vol. 36, no. 1, pp. 51-64, May. 2003. [https://doi.org/10.1501/Egifak\\_0000000074](https://doi.org/10.1501/Egifak_0000000074)
- [4] The Liberated Learning Consortium, Increasing Access to Speech Recognition. <http://www.transcribeyourclass.ca/projectdescription.html/>, Accessed on: Dec.07, 2015
- [5] Ava, Group Conversations Made Accessible. <https://www.indiegogo.com/projects/ava-group-conversations-made-accessible/>, Accessed on: Oct. 12, 2014
- [6] Nuance Communications, Inc. Speech recognition. <http://research.nuance.com/category/speech-recognition/>, Accessed on: Nov. 21, 2015
- [7] K. Ryba, T. McIvor, M. Shakir and D. Paez, "Liberated Learning: Analysis of University Students Perceptions and Experiences with Continuous Automated Speech Recognition," *E-Journal of Instructional Science and Technology.*, vol. 9, no. 1, Mar. 2006.
- [8] M. Wald, "Synote: Accessible and Assistive Technology Enhancing Learning for All Students," in *International Conference on Computers for Handicapped Persons*, Berlin, Heidelberg, 2010, pp. 177-184. [https://doi.org/10.1007/978-3-642-14100-3\\_27](https://doi.org/10.1007/978-3-642-14100-3_27)
- [9] S. Kafle and M. Huenerfauth, "Evaluating the Usability of Automatically Generated Captions for People who are Deaf or Hard of Hearing," in *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility*, Baltimore, Maryland, USA, 2017, pp. 165-174. <https://doi.org/10.1145/3132525.3132542>
- [10] A. A. Mirza, C. Mirza and D. Rhodes, "Conference And Call Center Speech To Text Machine Translation Engine," Patent US20190121860A1, Apr. 25, 2019. <https://patents.google.com/patent/US20190121860A1/en>
- [11] S. U. UPase, "Speech recognition based robotic system of wheelchair for disable people," in *International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, 2016, pp. 1-5. <https://doi.org/10.1109/CESYS.2016.7889851>
- [12] K. Lunuwilaga, S. Abeysekara, L. Witharama, S. Mendis and S. Thelijagoda, "Web based programming tool with speech recognition for visually impaired users," in *11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, Malabe, 2017, pp. 1-6. <https://doi.org/10.1109/SKIMA.2017.8294132>

- [13] M. A. Noakes, A. J. Schmitt, E. McCallum and K. Schutte, "Speech-to-text assistive technology for the written expression of students with traumatic brain injuries: A single case experimental study," *School Psychology.*, vol. 34, no. 6, pp. 656-664, Nov. 2019. <https://doi.org/10.1037/spq0000316>
- [14] E. Biçek, "Real-time inscriptive follow up system of audible lecture with an Android based and web based application," M.S. thesis, Dept. Electrical and Electronics Engineering., Van Yüzüncü Yıl Univ., Van, Turkey, 2016.
- [15] J. Tebelskis, "Speech recognition using neural networks," Ph.D. dissertation, School of Computer Science., Carnegie Mellon University., Pittsburgh, Pennsylvania, 1995. <https://www.examinations-hub.com/download-file/104644743220180811135756.pdf>
- [16] A. Mesbah, A. van Deursen, "Component- and push-based architectural style for ajax applications," *The Journal of Systems & Software.*, vol. 81, no. 12, pp. 2194-2209, Apr. 2008. <https://doi.org/10.1016/j.jss.2008.04.005>
- [17] J. Garrett, "Ajax: a new approach to web applications, Adaptive path," <http://adaptivepath.org/ideas/ajax-new-approach-web-applications/>, Accessed on: Oct. 14, 2015
- [18] Yang J., Liao Z., Liu F., "The impact of ajax on network performance," *The Journal of China Universities of Posts and Telecommunications*, vol 14, no. 1, pp. 32-34, 2008. [https://doi.org/10.1016/S1005-8885\(08\)60007-2](https://doi.org/10.1016/S1005-8885(08)60007-2)