A Universal Design Infrastructure for Multimodal Presentation of Materials in STEM Programs

Leyla Zhuhadar
Dept. of Information Systems
Western Kentucky University
Leyla.Zhuhadar@wku.edu

Bryan Carson
Dept. of Library Public Services
Western Kentucky University
Bryan.Carson@wku.edu

Jerry Daday
Dept. of Sociology
Western Kentucky University
Jerry.Daday@wku.edu

Olfa Nasraoui
Knowledge Discovery & Web Mining Lab
Dept. of Computer Engineering
and Computer Science
University of Louisville
Olfa.nasraoui@louisville.edu

ABSTRACT
We describe a proposed universal design infrastructure that aims at promoting better opportunities for students with disabilities in STEM programs to understand multimedia teaching material. The Accessible Educational STEM Videos Project aims to transform learning and teaching for students with disabilities through integrating synchronized captioned educational videos into undergraduate and graduate STEM disciplines. This Universal Video Captioning (UVC) platform will serve as a repository for uploading videos and scripts. The proposed infrastructure is a web-based platform that uses the latest WebDAV technology (Web-based Distributed Authoring and Versioning) to identify resources, users, and content. It consists of three layers: (i) an administrative management system; (ii) a faculty/staff user interface; and (iii) a transcriber user interface. We anticipate that by enriching it with captions or transcripts, the multimodal presentation of materials promises to help students with disabilities in STEM programs master the subject better and increase retention.

Categories and Subject Descriptors
H. INFORMATION SYSTEMS.

General Terms
Algorithms, Design, Experimentation, Human Factors.

Keywords
Universal Design, Video Captioning, Accessibility, STEM.

1. INTRODUCTION
Over the past 30 years, more students with disabilities have taken advantage of the opportunities provided by higher education. The Americans with Disabilities Act [1] and the Rehabilitation Act of 1973 [2] require higher education institutions to provide these students with tools such as video captioning which are necessary for educational success.

Section 508 of the Rehabilitation Act [1] requires electronic and information technology communications to be made available in an accessible format by alternative means, perceptible by people with disabilities. These provisions apply to all entities—including colleges and universities—that receive federal money [2].

During the last eight years, faculty members at WKU have produced over 17,000 video lectures in the STEM disciplines. The Office of Distance Learning captions these videos, a process that is currently performed semi-automatically and is extremely tedious. As of the middle of 2013, only 43% of the videos were captioned. Meanwhile, it is projected that faculty will generate, on average, an additional 4,000 new video lectures each year. In 2007, the Office of Distance Learning developed the first open-source closed-captioning software [3]. This *Time Stamped Java program* takes text transcriptions (Word documents) as an input. The application then embeds time stamps within the script to generate a closed-captioning file (e.g., caption.txt) that synchronizes the text with speech. For more details, refer to section 3.

One problem with the current workflow is that it is extremely tedious and labor-intensive. Therefore, we propose using a Universal Video Captioning (UVC) platform to accommodate the unprecedented growth of video lectures. The proposed software platform will provide an automatic approach to embed captions into accessible STEM-related videos.

This project will focus on promoting better opportunities for students with disabilities in STEM programs to understand their materials and may eventually open new avenues for research on student learning styles. We anticipate that by adding captions or transcripts, the multimodal presentation of materials will help students with disabilities in STEM programs increase mastery of the subject matter and increase student retention. Later, by measuring the effectiveness of adding captioning to educational videos, we will be able to assess their impact on studying, grades, and persistence for individuals with disabilities. This would foster our understanding of how students with disabilities benefit from online learning and how they interact with their teachers.

The (UVC) platform, presented in this paper, is modular to provide interoperability, and will serve as a repository for uploading videos and scripts. All the building blocks of this project—literature, software, developed model, and instruments—will be made available to the public and other interested

WWW 2015 Companion, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05
http://dx.doi.org/10.1145/2740908.2744712

researchers online using the Creative Commons Open-source License (CC BY-NC 3.0—Creative Commons Attribution¬ Non-Commercial 3.0). Our contribution to open-source learning has the potential to provide a solution to thousands of individuals and universities who already have educational videos.

2. Literature Review

As higher education strives to serve students with disabilities, the STEM disciplines are no exception [4]. The Minority Science & Engineering Improvement Program [5] requires a special effort to recruit disabled students into these fields. Although often presented as an issue for distance education, the reality is that accessible learning is a barrier for students with disabilities in more traditional face-to-face classes as well. In fact, distance education solutions also hold some promise for helping these students. For example, lectures may be recorded and transcribed. This multimodal solution helps make hybrid classes and distance education available to all disabled students [6]. Actually, the recent emergence of distance education has caused educators to develop innovative ways to teach and deliver content in STEM programs in the online environment by using video capturing, reusable audio and video materials, and interactive simulated labs.

The Accessible Educational STEM Videos Project permits us to combine the advantages of captioned/transcribed multimedia with the recognized potency of audio and video to reinforce learning (via the capability of re-listening or re-watching special segments on demand). Multimedia objects with captioning may also assist students whose native language is not English. In 2009, 32.7% of all students in the STEM disciplines were foreign nationals [7]. Use of subtitles is very common in foreign-language and English as a Second Language classes. Stewart & Pertusa [8] noted that native English speakers studying Spanish, who viewed films with the audio and subtitles in the same language, learned more vocabulary, showed a gain in comprehension, and had higher levels of engagement and satisfaction with the material. Similarly, Hayati & Mohmedi [9] reported a significant improvement in international students learning English, noting that multimodal delivery "[enhances] comprehension better than simply processing subtitle through silent reading" (p. 190). This line of research suggests that the use of captioned multimedia may also improve learning outcomes among STEM students whose native language is not English.

In 2008, higher education institutions enrolled 10.8 million students who identified as having a disability [10]. Figure 1 shows the relative distribution of students with disabilities. The project's significance will extend from the integration of multimodal delivery of audio and video in STEM courses to the transformation of curricular activities in higher education to accommodate students with disabilities. Disabled students vary in their learning styles based on the type of accommodation they require. While students with sensory or learning disabilities may be helped by assistive technology, the most important issue is proper pedagogical design [11].

As Thomas Klein [12] pointed out in a much-cited article, "[C]omputers are not a panacea; they cannot rescue a school from weak teachers, a weak curriculum, or the absence of sufficient funding." Universal design—making courses accessible for all students regardless of disabilities—goes beyond making courses usable and accessible. It involves "creating instructional goals, methods, materials, and assessments that work for everyone—not a single, one-size-fits-all solution but rather flexible approaches that can be customized and adjusted for individual needs" [13].

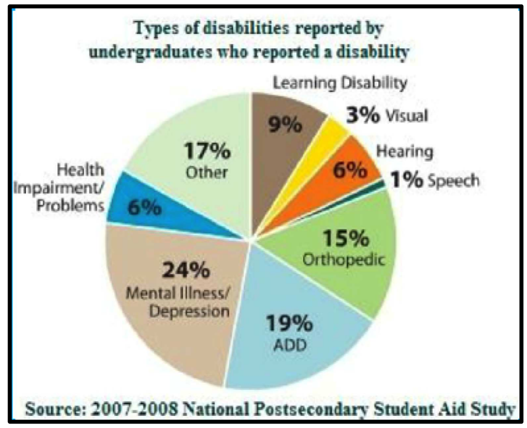

Figure 1

Proper course design from the very beginning, with attention to multimodal environments, will help all students, not just those who have identified as having a disability[14]. As van Hoorebeek [2] explained: Careful analysis and staged development of e-learning objects can lead to other benefits for the wider student populace, with other students benefiting from the new methods of delivery. . . . The increased use of alternative formats can significantly aid students that do not fulfill the disability criteria or have chosen . . . not to disclose their disability. . . . [These students] may also benefit from the increased focus on accessibility of e-learning provision, e.g. an unidentified dyslexic student may benefit from the audio and video materials from the lectures being available online or an animated equivalent of spoken text.

Marcyjanik [15, 16] noted that disabled students face greater challenges as learning or sensing difficulties can be amplified by improperly designed multimedia. Non-captioned videos, colored links, and animated images may all be difficult for disabled learners. Table 1 shows how selected disabilities interact with media, and how accommodations may be made. STEM students with disabilities are especially impacted by the challenges they face. Color, sound, and observation are all important in the study of scientific information.

For example, only one in four deaf students actually persist to graduation [10]. However, technology also provides some ways, in which some of these challenges can be overcome, particularly for students with sensory and learning difficulties. Nonetheless, technology also provides some ways to overcome these challenges, particularly for students with sensory and learning difficulties. But a one-size-fits-all approach will not work. Because the necessary accommodations vary by type of disability, it can seem overwhelming to faculty members. For example, a blind student might be able to hear online videos but will miss the visual details.

Table 1. Access Issues and Possible Modifications for Select disabilities (Roberts & Crittenden, 2009).

Disability Type	Access Issues	Possible Modifications
Hearing or Hard of Hearing	• Unable to hear audio materials. • Background noise may impede ability to hear.	• Closed or open media captioning. • Transcripts for audio files • Use of audio & video cues.
Blind or Low Vision	• May not be able to see or read small text or graphics. • Materials may not be accessible to assistive technology devices. • Screen readers or magnification software most often used.	• Provide alternative text write up. • Save PowerPoint as rich text. • Provide descriptive text for other visual information. • Order texts in Braille. • High contrast between colors in background and foreground. • Format for mouseless operations.
Mobility Impairments	• May have limited use of hands and arms. • May also have decreased eye-hand coordination. • May be unable to use mouse or make multiple keystrokes. • May require the use of screen reader.	• Provide alternative text write up. • Save PowerPoint as rich text. • Use descriptive text for visual information. • May need information in a format in which the student can control the pace of instruction. • Format for mouseless operations.
Learning Disabilities	• May not be able to process information if material moves too quickly. • Material presented with too much background can be distracting.	• Provide information in an alternate format. • May need information in a format in which the student can control the pace of instruction. • Provide written instructions.

Generating simultaneous captioning for the video that describes the action may be good for a deaf student. Evidently, a combined approach of describing the action and the dialogue will benefit both learners. Blind students can use a screen reader to listen to the captioning details, while deaf students will be able to read the transcription. Learning-disabled students will also benefit from the simultaneous multimodal presentation of information and materials.

3. Project Details & Design

3.1 Current system

In mid of 2013, only 43% of the 17,000 video lectures, produced in the STEM disciplines (Figure 2), are captioned. The Office of Distance Learning is capturing these videos semi-automatically.

The process involves five undergraduate students (*transcribers*) who perform the transcription process (Figure 3). The process consists of downloading the video to a desktop computer, listening to the audio, and typing the script.

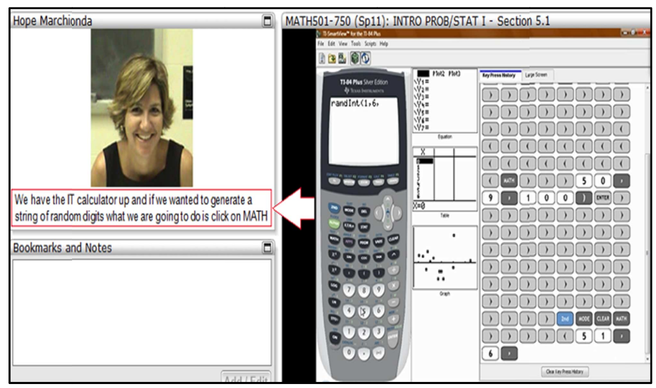

Figure 2

The transcribers use a "Wave Pedal" to control the audio speed (start, stop, move forward, move backward, etc.). The script is saved as a Microsoft Word document in a shared drive that can only be accessed locally.

Figure 3

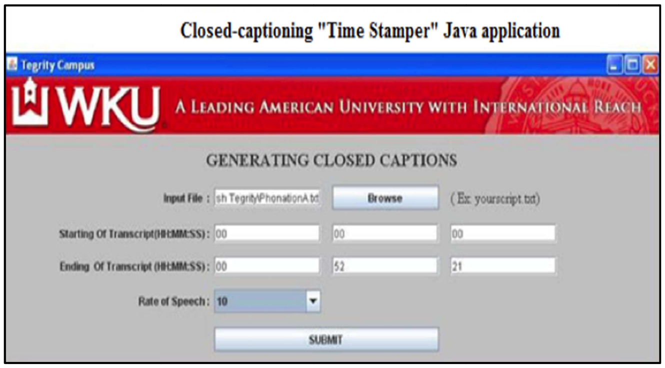

Figure 4

In 2007, the office of Distance Learning developed the first running closed-captioning software (Figure 4). It is a *Time Stamped Java program*. This program takes text transcriptions (Word documents) as an input. The application then embeds time stamps in the script to generate a closed-captioning file (e.g., caption.txt) that synchronizes the text with speech (Figure 5). This software is available in [3].

3.2 Workflow

Students enter the starting and ending times of the lecture in HH:MM:SS format. The software works with rate of speech (1-10) using a simple algorithm. The program divides the vocabulary size in the script over the length of the audio.

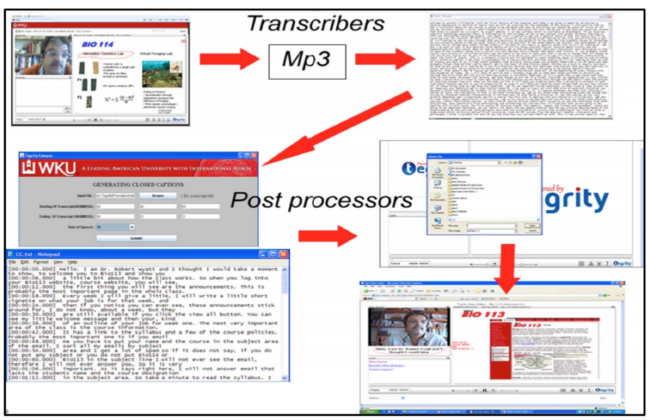

Figure 5

3.3 Current Workflow Limitations

The current workflow is extremely tedious and labor-intensive. Transcribers have to work in one location using specific desktop computer hardware equipment with appropriate software and wave pedals (Figure 6).

Technically, increasing the rate of speech increases the vocabulary size presented in the video. An example of a video lecture with embedded captions can be seen in Figure 2.

In addition, the communication process between the faculty members and transcribers is limited and tedious, being mostly based on emails.

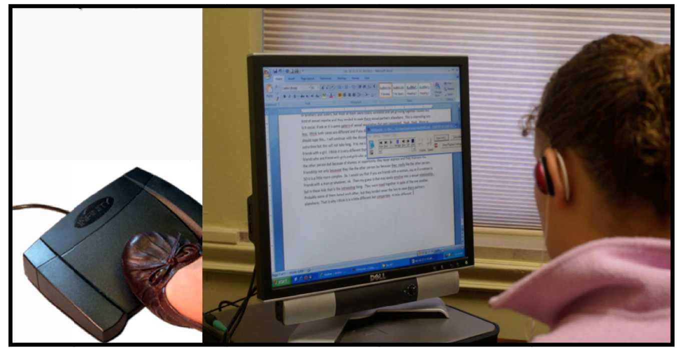

Figure 6

3.3.1 The workflow consists of the following procedures

- Communicating with faculty members to inform them that captions or transcripts are going to be added to their video lectures.
- Assigning tasks to students (which videos need to be transcribed).
- Follow-up with faculty members and transcribers.
- Archiving the process to generate semester-based reports.

The current workflow cannot handle the unprecedented growth of video lectures at Western Kentucky University. The emergence of the Web—along with new data-intensive computing platforms and technologies such as WebDAV, Hadoop, Solr,, and others—promise a solution to handling the large amounts of data in addition to enhancing the ability to search for specific resources.

Therefore, we will propose a solution to this challenge where we can manage resources in a timely matter. The solution is replicable, and could be used by other universities in assisting STEM students with disabilities.

4. Proposed Architecture for the Universal Video Captioning Platform

4.1 Proposed Infrastructure

The purpose of our research is to establish an automatic workflow that is accessible over the web to everyone. The proposed infrastructure is a web-based platform that uses the latest WebDAV technology (Web-based Distributed Authoring and Versioning) to identify resources, users, and content. The front end consists of a Solr Search Engine, Flex, and PHP, while the back end will consist of a MySQL Server and an Apache Tomcat Server.

The platform infrastructure consists of an administrative management system, a faculty/staff user interface, a transcriber user interface, and a synchronized captioning applet.

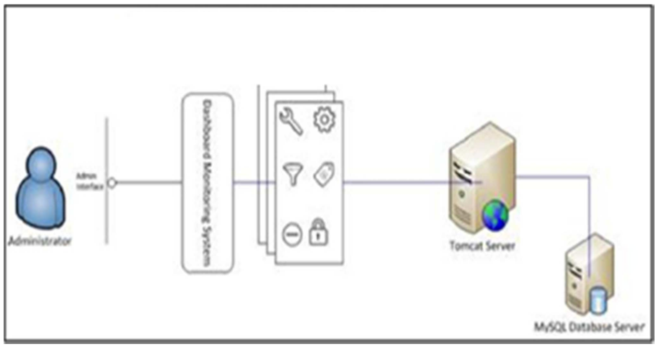

Figure 7

4.1.1 *Layer1: Administrative Management Interface*

The purpose of this layer (Figure 7) is to control user access, provide upload and download interfaces, maintain the database system, and troubleshooting problems. In addition, the administrator will be able to track the transcription quality by opening or closing any raised cases, request feedback from faculty and staff, evaluate transcriber performance, provide a monthly report to each transcriber, and generate statistical reports. Finally, the administrative module will provide a search engine mechanism to allow searching for a specific course, faculty, video, caption, or script.

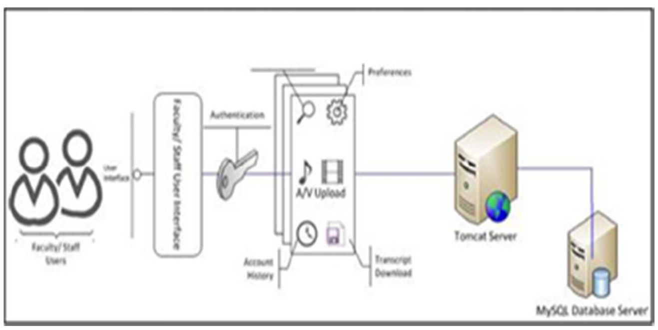

Figure 8

4.1.2 *Layer2: Faculty/Staff User Interface*

The Faculty/Staff User Interface module (Figure 8) will consist of the registration and login access. The uploading dashboard will allow users to upload video or audio material or references to video or audio by using WebDAV technology. A universal identifier will be associated with each resource. If the multimedia item is already located on the web, only a reference to that resource will be added to the database. The downloading dashboard will list the current status of each submitted request (download captioned video, download script, and status ("in process" or "completed")).

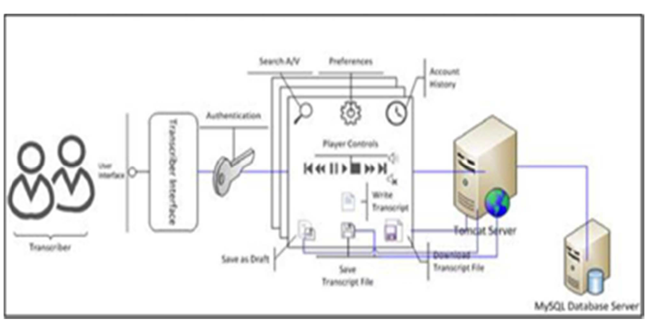

Figure 9

4.1.3 *Layer3: Transcriber User Interface*

The transcriber user interface (Figure 9) will provide registration and login access, as well as task management. Four dashboards are available for these users. The transcribers will receive notices of requested audio or video files that need transcription. When a transcriber receives a notice, this task will be moved to his/her personal interface. The task will be removed from the queue, and the status of the resource will become "in process." A separate dashboard will list all the queued resources that the transcriber is working on. Each transcriber will see a different (adaptive) interface listing the individual specific tasks.

Once the project has been assigned, the transcriber will listen to the audio and transcribe on the platform in real time. Scripts in progress can be saved as drafts. When the individual transcriber logs into the system again, all assigned and uncompleted work will show up. Once the work has been completed, the captioned video and script will be moved from the transcriber interface to the live faculty/staff interface and the status of this specific task will become "completed." Instead of using a wave pedal as a hardware device, this functionality is provided through the use of arrow keys on the keyboard.

An enhanced Web-based captioning Applet allows transcribers to generate automatic synchronized captions with time stamps. Furthermore, transcripts with content description will be provided in plain text as additional content.

4.2 Synchronized Captioning Applet

As a starting point, we plan on developing an enhanced web-based version of our software. We noticed over the last five years that the software has three drawbacks: 1) it is not web-based and cannot be accessed in real-time over the web; 2) the accuracy of the synchronization between the audio and the script is affected negatively when there are long pauses; 3) the application only works with specific video formats.

We therefore propose using a JavaScript Applet, embedded inside the platform, which can be accessed easily by transcribers. To improve accuracy, we propose using a segmentation process to the audio file before we add the script. For instance, the short-time energy ($Short\_time\ En$) can be used to detect long pauses in the audio. This measure can differentiate silence from speech and is calculated as follows,

$$\text{short\_time\_energy} = \sum_{-\infty}^{\infty}[x(m)w(n-m)^2] = \sum_{m=n-N+1}^{m=n} x(m)^2 \qquad (1)$$

Where $x(m)$ is the audio signal, $w(m)$ is a rectangular window of length N, and n is the index of the $Short\_time\ En$. When $Short\_time\ En$ drops below a certain threshold, we consider this frame as a pause. After such a pause has been detected, we save its location. Therefore, there will be no captioning during long pauses. This process is fully automated since we are not going to make any physical changes to the original audio. The speech that is converted to equivalent text in Dashboard 2 (transcriber interface) will be selected, long pauses will be identified, text vocabularies will be divided over the length of the audio after extracting long pauses, and vocabularies will be embedded in the correct time frame with no pauses. Because manual captioning is labor intensive, this solution will have a significant impact on the system performance.

5. Conclusion and Future Work

We presented a modular universal video-capturing (UVC) infrastructure to serve as a repository for uploading videos and scripts. We plan to make the developed platform available to the public and other interested researchers online using the Creative Commons Open-source License (CC BY-NC 3.0—Creative Commons Attribution¬ Non-Commercial 3.0).

The UVC platform will automatically embed captions inside educational videos using an enhanced and extended web-based version of our current software, which will be developed with rigorous algorithms to accommodate a variety of video formats. STEM teachers can use the online UVC platform to upload audio or video teaching materials, and will receive a notification to download both the original multimedia with synchronized captions and the transcript.

The fusion of video capture with captions and/or transcripts promises to immerse students with disabilities in STEM programs in multimodal presentations of information and learning materials that they can interpret or reinterpret in a better way, helping pave the way to a better and broader understanding. We anticipate the synchronized multimodal presentation of materials to help disabled learners increase their mastery of the subject and decrease their attrition. Our assumption is based on multiple studies that have found significant improvement in grade point average and reduced attrition among disabled learners who have used assistive technology. The proposed innovative platform can be extended by applying statistically valid instruments to help in understanding the learning process of the students, as well as to collect, analyze, and share data that can reveal patterns in the students' learning behaviors.

6. REFERENCES

[1] Kerr, S. Application of the Americans with Disabilities Act to the Termination of the Parental Rights of Individuals with Mental Disabilities, The. *J. Contemp. Health L. & Pol'y*, 161999), 387.

[2] van Hoorebeek, M., Walker, S. and Dermo, J. Analysing the parameters of the duty of reasonable adjustment to facilitate access to e-learning resources. *Multicultural Education & Technology Journal*, 3, 3 2009), 227-241.

[3] Zhuhadar, L. Open-Source Closed-captioning Software: Time Stamped Java program: http://ada-wku.pbworks.com/w/page/1367266/FrontPage2007-2013).

[4] Jerome, M. K., Neuber, K., Stegall, B., Emenova, A. and Behrmann, M. Technology for integration of students with disabilities in higher education. *The Engineering Handbook of Smart Technology for Aging, Disability, and Independence*2008), 101-116.

[5] August, D. and Hakuta, K. *Improving schooling for language-minority children: A research agenda*. National Academies Press, 2000.

[6] Keengwe, J. and Kidd, T. T. Towards best practices in online learning and teaching in higher education. *MERLOT Journal of Online Learning and Teaching*, 6, 2 2010), 533-541.

[7] Wasem, R. E. *Immigration of foreign nationals with Science, Technology, Engineering, and Mathematics (STEM) degrees*. Congressional Research Service, Library of Congress, City, 2012.

[8] Stewart, M. A. and Pertusa, I. Gains to Language Learners from Viewing Target Language Closed□Captioned Films. *Foreign language annals*, 37, 3 2004), 438-442.

[9] Hayati, A. and Mohmedi, F. The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 42, 1 2011), 181-192.

[10] Marschark, M., Pelz, J. B., Convertino, C., Sapere, P., Arndt, M. E. and Seewagen, R. Classroom interpreting and visual information processing in mainstream education for deaf students: Live or Memorex®? *American Educational Research Journal*, 42, 4 2005), 727-761.

[11] Moallem, M. Accommodating Individual Differences in the Design of Online Learning Environments: A Comparative Study. *Journal of Research on Technology in Education*, 40, 2 2007).

[12] Klein, T. Electronic Revolution at the Educational Crossroads: Foot-Dragging on Campus. *College Teaching*, 43, 4 1995), 151-155.

[13] CAST http://cast.org/udl/index.html. City, 2012.

[14] Moreno, R. and Mayer, R. Interactive multimodal learning environments. *Educational Psychology Review*, 19, 3 2007), 309-326.

[15] Marcyjanik, D. *An Analysis of Web Accessibility in Online Nursing Education for Persons with Disability*, 2013.

[16] Marcyjanik, D. and Zorn, C. R. Accessibility in Online Nursing Education for Persons With Disability. *Nurse educator*, 36, 6 2011), 241-245.