Generation of educational content for Open Digital TV and IPTV to assist the learning of Brazilian Sign Language (LIBRAS)

Hilda Carvalho de Oliveira, Celso Socorro Oliveira, Edson Benedito dos Santos Jr.

Abstract—Today there are 24.6 million of the people with special needs in Brazil, with 23% having some type of hearing loss and 2.9% of these deaf. The Accessibility Brazilian Law establishes that the accessibility of services for the deaf shall be provided by interpreters or people trained in Brazilian Sign Language (LIBRAS). Brazilian Laws provides that the teaching of LIBRAS must be part of the curriculum of all courses in Special Education, Speech Therapy and Magisterium at high school and college courses. Furthermore, all school systems in Brazil must provide a bilingual education (LIBRAS and Portuguese) as a right to deaf students. According to the National Federation on Education and Integration of the Deaf there is great need for qualified educators, geographically available in appropriate locations. The TV is an important means of distributing educational content since the 1950s and it is available in 98% of Brazilian homes. By 2016 the Open Digital TV (ODTV) in Brazil must cover the entire national territory. In this context, this paper presents the system SynchLibRAS, which facilitates the generation of educational content to assist the learning of LIBRAS. This system takes as input a video with audio and allows inserting subtitles and a LIBRAS window in a synchronized way. The window LIBRA is recorded by an LIBRAS interpreter in front of a webcam, with automatic synchronization among caption, audio and image. The resulting content is processed by a system called HiddenLIBRAS that uses the middleware Ginga-NCL – standard for Brazilian ODTV and IPTV. It allows the caption and the LIBRAS window are optional for viewing in both environments. The focus of this paper is to present the architecture and implementation of the SynchLibRAS, with emphasis on the synchronization process and the practical way of generating content with LIBRAS windows.

Keywords—Accessibility, deaf people, Digital TV, IPTV

I. INTRODUCTION

The evolution of technology in our days should include the People with Special Needs (PSN), facilitating their access to information and interaction with society [1]. A Person with Special Needs (PSN) is understood as an individual whose mobility or perception of the environment characteristics are reduced, limited or annulled [2]. These limitations related to sensory-motor organs can manifest themselves individually or together.

Fusco [3] and Baranauskas [4] comment that there are few consistent initiatives that support the development of technological alternatives to deaf people. It is understood as deaf person who has a hearing loss greater than 25 db but is still capable to hear some sound. When there is total lack of hearing is called anacusia, different of the deafness [5].

A large step was made by decree of Salamanca, which was signed in 1994 by UN in order to direct the principles, policies and practices in the area of special educational needs [6]. This Decree emphasizes the development of digital technologies that surpasses barriers for PSN.

Accessibility Brazilian Law (Decree-Law n° 5296 of 02/12/2004) regulates two important Laws: (1) Law n° 10048, of 08/11/2000, which deals with subjects about priority services to PSN; (2) Law n° 10098, of 19/12/2000, establishes ways to promote accessibility to PSN, including projects and constructions with public or collective purposes [7]. The World Wide Web (Web) and the TV are considered works of public and collective purposes. In relation to the deaf, Law n° 10098 establishes that services must be provided for interpreters or individuals trained in LIBRAS.

For communication between the deaf and society, most countries provide a national Sign Language. Unfortunately there is not a universal Sign Language. In Brazil, the French Sign Language influenced the emergence of two Signs Languages in Brazil: (1) Language Kapoor, currently restricted to a tribe of Indians with high deaf rate; (2) the Brazilian Sign Language (LIBRAS), regarded as the official Brazilian Sign Language, by Decree-Law n° 10436 [8]. This decree established that the teaching of LIBRAS must be part of the curriculum of all courses in Special Education, Speech Therapy and Magisterium at high school and college courses. In addition, all school systems in Brazil must provide a bilingual education (LIBRAS and Portuguese) as right to deaf students (LIBRAS should be the first language to be learned) [9].

According to the National Federation on Education and Integration of the Deaf [10], there is great need for qualified educators in LIBRAS and which are geographically available in appropriate locations to support this demand.

Like the native Language of a country has differences in pronunciations and vocabulary itself in the region, LIBRAS has structural differences according to Brazilian region. For example, there are different ways to sign representation to certain elements in the North and South.
of Brazil. Thus, where an educator or interpreter of LIBRAS will act is important for a successful work.

In this context, this work contributes to teach LIBRAS for professionals with educational purposes - which affects directly the learning and social inclusion of the deaf people. The idea is to provide a support instrument for learning LIBRAS similar to what is used for learning foreign Languages: videos with audio in the foreign Language with optional caption written in the Language of learner or of the audio.

In this direction line, the proposed solution has two components: a system that generates the desired video content (SynchrLIBRAS) and another that converts content to digital format, allowing visualization in IPTV and ODTV in Brazil (HiddenLIBRAS). In preview, the writing caption and the LIBRAS window can be hidden, similar to the process of closed caption. They can be enabled by the user itself. The LIBRAS window can also have their size and position adjusted by the user via the remote control or mouse/keyboard.

The main goal of this work is to present the system SynchrLIBRAS - a Web application which takes as input a video with audio and allows inserting subtitles and a window with an interpreter in LIBRAS (LIBRAS window) in a synchronized way. In the first version of the SynchrLIBRAS, the videos are downloaded from YouTube.

The generated content by SynchrLIBRAS can be executed by any system that implements the middleware Ginga-NCL, such as the Brazilian System of Digital TV (BSDTV), which is being introduced in Brazil, and IPTV. For this, HiddenLIBRAS allows all the videos stored in the repository of the SynchrLIBRAS be transmitted for the display on devices that implement the Ginga-NCL specifications. The preview display shows the user/telespectator a menu to select the desired option: subtitles, LIBRAS window or both, as illustrated in Fig. 1.

Thus, the section 2 presents some comments on LIBRAS in the context of the Web and ODTV in Brazil (BSDTV). Section 3 presents the proposed system in this work, SynchrLIBRAS, to synchronize video, audio and LIBRAS windows. The conclusions are in section 4.

II. LIBRAS ON THE WEB AND ON THE OPEN DIGITAL TV IN BRAZIL

On the Web, there is a great amount of documents and software that help the inclusion of accessibility in the pages, sites and portals [11]. Access to the computer should not be restricted to people without Special Needs (not PSN).

According to W3C [11], the major barriers faced by deaf on the Web are due to the lack of captions or transcripts of audio files and self-descriptive images. This complicates the understanding of deaf people who learned sign language as first language, because this requires the association between oral and written languages.

There are many research groups working on automatic translators between LIBRAS and Portuguese language. However, still faces different kinds of obstacles, such as: the evaluation of the most effective method to be adopted in automatic translation and the large number of cultural and linguistic variations in the Brazil’s large territory [12].

In the translation between sign and oral language, it is usually used an intermediate language (Interlingua) to facilitate the conversion between two distinct grammatical structures [13]. In this intermediate stage should be treated specificities relating the structures and problems of grammatical ambiguities. For an example, the word "manga" in Portuguese may be interpreted as "part of the shirt," as a "fruit" or even as a Japanese comics ("manga"). There is also the complexity of finding the correct representation in LIBRAS.

There are cases like the word "fermento" in Portuguese (Baking Powder) that interpreters use the ideological way of translation, signaling in LIBRAS something like "white flour that makes grow." In this situation, the interpreters use dactylology as a resource, spelling the word in Portuguese through LIBRAS. This feature, however, is only useful for people already alphabetised in Portuguese - which is difficult for the deaf from birth.

This approach of automatic translators is not used in the proposal presented in this paper. Instead, the system SynchrLIBRAS uses a practical way with the interpretation of an expert in LIBRAS which analyzes the gestural translation more appropriate to the context. This avoids certain problems pertaining to automatic generation. In addition, there are no restrictions on video selected for translation, which can be chosen through didactic and pedagogical criteria.

In relation to TV, nowadays there are still a few numbers of programs with subtitles or with simultaneous translation in LIBRAS for open analog TV, by broadcasting. The window with LIBRAS translation occupies much of the screen and generally people without Special Needs think it's a nuisance.
The Brazilian standard NBR 15290 [15] deals with accessibility in communication on television in general. The standard defines closed caption as being "the transcription, in Portuguese, of the dialogues, sound effects, ambient sounds and other information that could not be perceived or understood by people with hearing impairment". This standard defines "LIBRAS window" as a delimited space on the screen where the information is interpreted in LIBRAS.

The subtitle which optionally may appear on the TV screen according to activation of a decoder device (internal or peripheral) is called "CC (closed caption)". In digital systems, the closed caption is generated, transmitted and decoded using only logical filters (software) and its capturing is done by codecs for interpretation of CC signs. The digital transmission system, for IPTV or ODTV, permits these features are best exploited.

The Supplement Standard nº 01/2006 for BSDTV, approved on 27/06/2006, stipulates that all programming presents an optional window with interpreter of LIBRAS. The deadline was two years so that TV stations would promote the necessary adjustments and prepare a schedule of progressive amount of daily TV programs with accessibility. This period was considered too short by the TV stations. Thus, the deadline for TV stations transmit 24 hours per week of adapted programming was extended for ten years. The only requirement until 07/01/2011 is that TV stations introduce audio description on TV programmes (oral description of images) - which will benefit the visually impaired.

The System SynchrLIBRAS

SynchrLIBRAS has three basic functional modules.

The first module lets you generate captions written synchronously with the core of video/audio core.

The second module handles the recording of LIBRAS signs generated by the users who are LIBRAS experts (interpreters). Each stretch of recording must be synchronized with the caption written, audio and video. The infrastructure and the results of the recording must be in compliance with the Brazilian standard NBR 15290.

The third module uses the SQL (Structured Query Language) to extract the information stored by the other two modules and generating XML (Extended Markup Language) files. Then this XML content is processed to generate content in one of the following formats: (1) XHTML and SMIL (WEB environment); (2) NCL-Lua (BSDTV and IPTV environments), through XSLT (Extensible Stylesheet Language Transformations) [17] [18].

Markup languages help to synchronize different multimedia objects by the system [19]. The languages of management of temporal and spatial synchronization provide use of algorithms to monitor the synchronization between video streams [20].

The system uses the SMIL 3.0 (Synchronized Multimedia Integration Language) [14] to support the implementation of temporal synchronization mechanism in Web environments. SMIL (pronounced "smile") is recommended by the W3C to manage synchronization relationships (temporal and spatial) between multimedia objects [14]. This language is precursor of NCL (Nested Context Language), used by BSDTV.

The SynchrLIBRAS system follows the guidelines of the NBR 15290 for the treatment of subtitles. According to NBR 15290 closed captions can be produced during a live program ("live CC") or may be produced after the program finished and was recorded ("pre-recorded CC"). Live CC should be aligned to the left of the screen, while the pre-recorded CC can be aligned where best facilitates the telespectator view: left, right or central part of the screen. The SynchrLIBRAS system works with pre-recorded CC. The guidelines of NBR 15290 for LIBRAS window include topics that cover the production in studio, preview window and the conditions to the interpreter record the translations.

The recording process of LIBRAS translation uses a webcam and it is suggested that recording conditions follows the guidelines of the NBR 15290. The studio where the interpreter image is recorded must have enough space to avoid shadows, adequate lighting to enhance the image quality, fixed or supported video camera, as well ground marks for the adequate position of the interpreter.

The window with the interpreter must have sharp contrasts, covering all the movement and gestures made by the interpreter and avoid shadows/ blurring in the eyes of the interpreter. Consider that the capability displaying a small window on a video image is known as wipe.

When the image of the LIBRAS interpreter is on the wipe, this wipe must be placed in a position that is not clouded by the black stripe of the subtitle. Other images must not be included or overlapped in wipe. When there is displacement of the wipe on the screen, the window image must be continued. The window height must have at least half the height of the TV screen, with a width of at least a fourth part of the screen width. The costumes, hair and skin of the LIBRAS interpreter must be contrasting with each other and the background of the scenario.

Fig. 2 shows an example of synchronization of video and legend to help understand the temporal view of the system SynchrLIBRAS [16]. Note that two distinct regions are generated: one for the allocation of video objects (rgVideo1) and another for the objects of legend (rgLegenda). The beginning of the process of subtitles (markers 2, 4 and 6) is controlled by the links of sync "onBegin1StartN".

The end of the caption text is controlled by "onEnd1StopN". All links represent the start time and end time defined in the coding. Each link corresponds to a file with the contents of a part of the subtitle ("legenda01.html", "legenda02.html" and "legenda03.html") [16]. Markers numbered 1 through 7 represent the periods of the video playing. The markers 1, 3, 5 and 7 represent the periods of video playing without the presence of subtiles. The markers 2, 4 and 6 represent the periods of subtitle displaying. Serg [16] presents an NCL code that shows how links of synchronization implement the connection between an
object and its display region (rgLegenda) on the screen. The user interface allows it to type the URL (Universal Resource Locator) that references the selected video of the site YouTube, as illustrated in Fig. 3. Thus, SynchrLIBRAS searches the video on own YouTube site, through the Jscript API (free and open source). In future, releases the input of video can be extended to URLs of other sites and new forms of video input. The videos are recorded and processed by the system with the FLV file format.

In the environment of caption written of the first module, the user has possibility of pausing the video to insert the caption at any time. This frees the text field below the main video for the user type the text. Completing the text, it clicks on the button "add caption" to insert the caption into the system. The system adds the phrase typed in JScript Grid as well as temporal information of beginning and end of the subtitle display - taken from directly from the main video timeline.

So it is possible adding each sentence, composing a list sorted by insertion in JScript Grid, displayed in a window on the right side of the screen. It is possible to add, delete or edit (double-click the mouse on the line) in the captions JScript Grid a way manually. A text line can be "broken" with <enter> to better distribute the text caption on the video page. This contributes to a flexible user interface.

The states of player of the SynchrLIBRAS system are monitored by the function "onpyplayerStateChange(newState), as shown in Fig. 5. After inserting the URL and capture the video by the command "executar()", the player enters the paused state (newState === 1) and disables input field of subtitles, waiting for the user to click "play".

```
... function onpyplayerStateChange(newState) {
  if(newState === 2) {
    $('newcaption').removeAttr('disabled').val('').focus();
    $('#addcaption').removeAttr('disabled');
  } else if(newState === -1 || newState === 0 || newState === 5) {
    $('newcaption').attr('disabled','disabled');
    $('#addcaption').attr('disabled','disabled');
  } else {
    $('newcaption').attr('disabled','disabled').val('Press pause to add caption');
    $('#addcaption').attr('disabled','disabled');
  }
...```

Fig. 4 Activity diagram of the SynchrLIBRAS System.
When the user thinks that video playback reached an ideal point for inserting a part of the legend, the user pauses the video ("pause()”). Then the state of player changes (newState === 2).

A video playback is interrupted and the input field of subtitles is activated. The user enters subtitle text ("inserir()”) and the state of player changes, disabling the input field of subtitles and returning to play the video. This process is repeated until the user finalizes the subtitling process.

To start the recording environment of the signals of LIBRAS generated by the interpreter the user can use a click the button “Record LIBRAS”, as shown in Fig. 6.

This environment is developed with Flex / Flash that provides conditions for a request for permission to access the computer’s webcam be done - standard procedure for providing private information user. After the approval of the application, the system indicates the beginning and end of recordings, via audible and visual warnings (the system uses the temporal information captured along with the captions written stored).

Then the user’s screen (interpreter of LIBRAS) displays the main video on the upper left, as illustrated in Fig. 6. On the right side of the screen is shown the Jscript Grid. A window with the image of the interpreter that is in front of the webcam is displayed at the bottom of the screen. The system informs the user about the timing of the recordings through this window.

After recording a stretch of video, a menu is offered to the user with the following options: (1) remake of the video, in case that the recording quality or time used were not satisfactory; (2) save the video if it is in conformity with its expectations and the technical constraints of Brazilian law; (3) download the video to the computer itself; (4) leaving the environment.

The base to start recording of LIBRAS signs is a timeline resulting from the previous process of subtitling: the timeline marks correspond to the marks entered by the user to pause the system. Synchronization between video and subtitle text is done by the marks.

The time interval for recording LIBRAS signs must be identical to range used for displaying subtitle text of a section of the video. However, the user can change that interval if it is short for the gesticulation.

The typed text is inserted into an XML file, as shown in Fig. 5. Sentences are synchronized according to start and end time of the timeline of the player when a section of the video is displayed (see Fig. 7).

A webcam connected to the computer is used to establish synchronization between the subtitle and the LIBRAS signs recorded.

Information about start and end of sentences are captured directly from the video when user presses "Pause" (function "getPlayerTime()") (Fig. 8). The variable "time" is then set. The values of the variable "time" are captured by the function "getSec()" and then analyzed to check if they are according to the values of time possible (it is considered display hours, minutes and seconds).

...
If they are not in standard format, the function reformats these values (Fig. 8).

The system also allows adding or removing sentences via the button "New" and "Remove" (Fig. 3). Thus, it is allowed generate intermediate sentences after the end of subtitling or delete sentences. This prevents the user having to restart the whole process of subtitling if it forgets any sentence or resolves add extra information at the end (Fig. 9).

Consider that the process of subtitling has already been completed and all the values of start and end time of the timeline have already been defined. When the user presses the button "Add Captions" (Fig. 3) the system searches and executes the function "addCaption ()" (Fig. 10), which stores this information in XML file (Fig. 7) and displays the text on the Grid (right side of the video screen) (Fig. 3).

Thus, the user can see the previous sentences along with their textual information when adding new sentences. This allows the user to have the notion of the sequence of sentences of captions, easing the process of recording of LIBRAS signs via webcam.

After synchronizing the caption and the LIBRAS signs recorded, the system offers the option of generating independent multimedia objects, with possibility to export them to the data repository (see Fig. 11).

This way, the data repository maintains the multimedia content generated in the steps of subtitling and video recording of LIBRAS interpreter (original video, LIBRAS signals recorded and XML files with metadata descriptors and temporal information for synchronization between objects). These data are available for access via Web or any digital media.

Each piece (an object) of caption written is related to a corresponding video object. This relationship is defined by the length of recording time and the interval between the beginning and end of the display of the caption written in the video.

The caption objects, the video objects and the synchronization information are stored in a MySQL database system.

The third module of the system SynchrLIBRAS is responsible for capturing the information from the database and convert (export) the content into formats that enable the visualization by local software applications in the computer or Web environment or BSDTV environment.

If the user chooses to store content in Web format, the content will be stored in a repository in the form of XHTML structures combined with languages based approaches for controlling the timing: XML and SMIL. The stored objects can be viewed through streaming multimedia generated by scripts "SMIL". Considering the relationship between NCL/LuaScript and SMIL, the stored objects and the associated synchronization information also allow direct export to the BSDTV.
IV. CONCLUSION

This paper presented an educational proposal that aims to contribute to the implementation of Decree-Law nº 10436 [8] in Brazil, benefiting the inclusion of the deaf people in the society. The idea is to provide a support instrument for learning LIBRAS similar to what is used for learning foreign Languages: videos with audio in the foreign Language with optional caption written in the Language of learner or of the audio. The proposed solution has two components: a system that generates the educational video content (SynchrLIBRAS) and another that converts content to digital format, allowing visualization in IPTV and OD TV in Brazil (HiddenLIBRAS). In preview, the writing caption and the LIBRAS window can be hidden, similar to the process of closed caption. They can be enabled by the user itself. The LIBRAS window can also have their size and position adjusted by the user via the remote control or mouse/keyboard.

The main goal of this work is to present the system SynchrLIBRAS - a Web application which takes as input a video with audio and allows inserting subtitles and a window with an interpreter in LIBRAS (LIBRAS window) in a synchronized way. In the first version of the SynchrLIBRAS, the videos are downloaded from YouTube. In the next version the videos may be from any source.

The development of SynchrLIBRAS used the MySQL database system and technologies associated with the languages: XML, XHTML, SMIL, XSLT, NCL and Lua. The educational content is constructed in a simple and practical from an existing video selected by the teacher or other specialist. It is not necessary specialized professional equipment - just a computer with a webcam. The environment for insertion of the caption written was development for non-experts in Informatics. The LIBRAS specialist does not need to be an expert in Informatics too. It can see the recording image through a window on screen and it is alerted with audible and visual signals to start and end a recording of caption. The proposed solution can be used throughout the Brazilian territory, allowing that the recording of LIBRAS embraces regional differences in accordance with the target audience for educational content. The solution can be extended to other countries, because the used languages (to audio, caption written and sign language to deaf) do not interfere in running the proposed software systems. In general, the work presented contributes to learning and dissemination of LIBRAS and another sign languages aiming to at social inclusion of deaf people.

ACKNOWLEDGMENT

The authors thank the contribution given to the work by Daiane Schiavon, LIBRAS expert, who contributed to the test recordings of the LIBRAS windows.

REFERENCES


Hilda Carvalho de Oliveira received her PhD in Electrical Engineering (Digital Systems) from the Polytechnic School of the University of São Paulo (USP) about information integration in open systems. She received her MSc in Computer Science from the Campinas University (UNICAMP) about databases unconventional. She is BSc in Mathematics from the São Paulo State University (UNESP), with research works in deterministic and non-deterministic expert systems. She is professor of Computer Science at UNESP since 1989, acting in undergraduate and graduate program. She is head of the research group "Software Engineering and Information and Communication Technologies (LesTIC)". She is member of the research group "Software Engineering and Quality Assurance". Currently, her interest areas are directed to digital convergence, e-Learning, t-Learning, software project management and domain-oriented software development. In addition, develops projects of university extension on digital and social inclusion as well as participates in projects administrative support.

Celso Socorro Oliveira received his PhD in Special Education from Center of Human Studies of the Federal University of São Carlos (UFSCar) about teaching mentally retarded deaf with Equivalent Stimulus Paradigm using educative software based on Graph Theory. He received his MSc in Electrical Engineering (Systems Engineering) from the Electrical Engineering College of the University of Campinas (UNICAMP) about use of Graph Theory modeling to solve logistic distribution of sugar and sugar plants. He is a Chemical Engineer from Polytechnic School of the Federal University of Bahia (UFBA), with research works on mathematical modeling of unit operations of refineries and distilleries. He is professor of Computer Science Department at UNESP, since 1993, acting in undergraduate and graduate program. He is the head of the research group "Software Engineering and Quality Assurance" and the head of the "Laboratory for Informatized Learning (LEIA)". He is member of the "Software Engineering and Information and Communication Technologies (LesTIC)". Currently, his interest areas is digital inclusion, e-Learning, Behavior Analysis and Human-Computer Interface. In addition, develops extension projects on "Teaching using Cartoons", "Brazilian Sign Language Teaching using Computers and Digital TV" and "Teachers capacititation using positive reinforcement and informatized tools".

Edson Benedito dos Santos Jr has a degree in Computer Science with Business Management in the Faculty of Technology of São Paulo (FATEC). Currently a Master's in Computer Science from São Paulo State University “Julio de Mesquita Filho” (UNESP) in the area of Computer Systems and line of research in software engineering and database. He is member of the research groups: "Software Engineering and Information and Communication Technologies (LesTIC)" and "Software Engineering and Quality Assurance". He is the head of the "Laboratory for Informatized Learning (LEIA)". Currently, his interest area is digital inclusion, e-Learning, and research in the area subject to Accessibility LIBRAS environments such as Web and open DTV.