Synchronizing Brazilian Sign Language (LIBRAS) on videos for Open Digital TV with educational purposes

CELSO S. OLIVEIRA[1], EDSON BENEDITO DOS SANTOS JR[2], HILDA C. DE OLIVEIRA[3]
[1]Departament of Computer Science, [2,3]Departament of Statistical, Applied Mathematics and Computing
Universidade Estadual Paulista (UNESP)
[1]Av. Eng. Luiz Edmundo Carrijo Coube, 14-01 - Bauru – SP, [2,3]Av. 24-A, 1515 - Rio Claro - SP
BRAZIL
[1]celso@fc.unesp.br, [2]e.jr@fc.unesp.br, [3]hildaz@rc.unesp.br

Abstract: - The TV is a potential midia of communication that affects all social classes and it is available in 98% of Brazilian homes. It has been used as a distributor of educational materials since the 1950s. By 2016 the Open Digital TV (ODTV) in Brazil might cover the entire national territory, replacing the TV with analog signals. Concerns about accessibility for People with Special Needs (PSN) in that media have become more intense since the 1990s worldwide. In Brazil, it is estimated 24.6 million of PSN, 23% of them having some type of hearing loss. Of these, 2.9% are reported as deafs. Brazilian Sign Language (LIBRAS) is considered the first literacy language for deaf people in Brazil. In this context, this paper presents a proposal to facilitate the generation of educational content for ODTV based on two components. One is called SynchrLIBRAS and allows subtitles synchronization in Portuguese and a LIBRAS translator window of videos downloaded from the Web. The second component allows the visualization of this content through the Brazilian System of Digital TV and IPTV - environments that implement the middleware Ginga-NCL. The main focus of this paper is the presentation of the first component: SynchrLIBRAS. This proposal has educational purposes, contributing to teach LIBRAS to people who may collaborate with social inclusion of the deaf people.

Key-Words: - Accessibility, deaf, LIBRAS, Digital TV, IPTV, closed caption

1 Introduction

The technology evolution in our days should include the People with Special Needs (PSN), facilitating their access to information and interaction with society [1]. PSN are understood as individuals 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. A Deaf person is understood as 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 [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 nº 5296 of 02/12/2004) regulates two important Laws: (1) Law nº 10048, of 08/11/2000, which deals with subjects about priority services to PSN; (2) Law nº 10098, of 19/12/2000, which 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 of public and collective purposes. In relation to deaf, Law nº 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 nº 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 a 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 and which are geographically available in appropriate locations to support this demand.

In this context, this study contributes to teach LIBRAS for professionals with educational purposes - which affects directly the learning and social inclusion of the deaf people.

The main goal is to present a Web application called SynchrLIBRAS, which results in the generation of videos with subtitles and a window with an interpreter in LIBRAS. In the first version of the work, the videos are downloaded from Youtube. The generated content 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, also it was developed a system called HiddenLIBRAS, which formats the content for transmission and viewing. The caption and the window can be hidden, similar to the process of closed caption. The window can also have their size and position adjusted by the user via the remote control or mouse/keyboard.

Thus, the section 2 presents some comments on LIBRAS in the context of the Web and Open Digital TV in Brazil. Section 3 presents the SynchrLIBRAS system architecture. The conclusions are in section 4.

2 LIBRAS on the Web and on the Open Digital TV in Brazil

On the Web, there is a great amount of documents and softwares 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 differents 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 gramatical 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 an agile mechanism of recording LIBRAS translation on a video which has already been inserted the subtitles.

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 [14] 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 Supplement Standard nº 01/2006 for BSDTV, approved on 27/06/2006, stipulates that all programming

presents a 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.

3 The SynchrLIBRAS System

The proposed solution has two components. One is the system developed for Web environment called SynchrLIBRAS, which takes as input a video with audio and allows inserting subtitles and a LIBRAS window in a synchronized way. The other system, called HiddenLIBRAS, allows all the videos stored in the repository of the SynchrLIBRAS be transmitted for the display on devices that implement the specifications of middleware Ginga-NCL. Thus, the content may be available for IPTV or BSDTV. The display allows the user/ telespectator select the desired option: subtitles, LIBRAS window or both, as illustrated in Fig. 1.
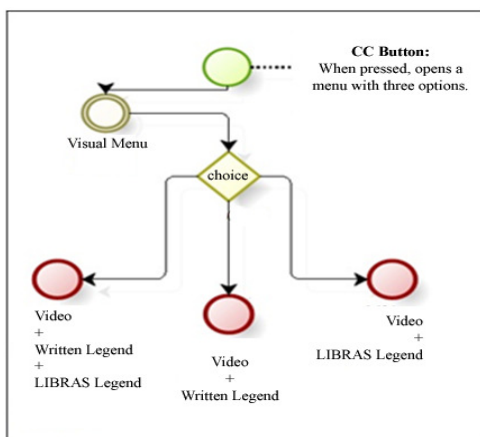

Fig. 1 Options to select subtitles, LIBRAS window or both.

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 was finished and 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 TV screen height, 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 [15]. Note that two distinct regions are generated: one for the allocation of video objects (rgVídeo1) and another for the objects of legend (rgLegenda). The beginning of the process of subtitling (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") [15]. 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 subtitles. The markers 2, 4 and 6 represent the periods of subtitle displaying. Serg [15] presents an NCL code that shows how links of synchronization implement the connection between an object and its display region (rgLegenda) on the screen.

In this first version, the user interface SynchrLIBRAS system allows URL (Universal Resource Locator) of a video on the site YouTube as input, as shown in Fig. 3. Thus, the system searches the YouTube video, via the API JScript. After the video download, it is expected a user command to start playing the video and the subtitle inserting process. In future releases the video input can be extended to URLs of other sites and new forms of video input.

The states of player of the SynchrLIBRAS system are monitored by the function "onytplayerStateChange (newState), as shown in Fig. 4. After inserting the URL and capturing the video by the "execute ()" command, the player enters the paused state (newState === 1) and disables input field of subtitles, waiting for the user to click "play". 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. The next steps are: export the XML code relating to subtitles or start the recording of LIBRAS signs ("inserirlibras ()").

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. 4. Sentences are synchronized according to start time and end time of the timeline of the player when a section of the video is displayed (see Fig. 5). 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. 6). 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. 6).

The system also allows adding or removing sentences via "New" and "Remove" button (Fig. 3). Thus, it is allowed to generate intermediate sentences after the end of subtitling or delete sentences. This prevents the user to restart the whole process of subtitling, if it forgets any sentence or resolves add extra information at the end (Fig. 7).

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. 8), which stores this information in XML file (Fig. 5) and displays the text on the Grid (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 (Fig. 9), the system offers the option of generating independent multimedia objects, with possibility to export them to the data repository (see Fig. 2). 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.

4 Conclusion

This paper presented a proposal that aims to contribute to the implementation of Decree-Law nº 10436 [8], benefiting the inclusion of the deaf in the society. For that, the proposal contemplates a Web system (SynchrLIBRAS) which enables the insertion of subtitles and windows with LIBRAS interpreter on a video. The mainly idea is that people can learn the way speeches and actions can be translated into LIBRAS. SynchrLIBRAS system was developed with XML and NCL programming languages. This facilitates the communication with the system HiddenLIBRAS, which treats the content to be displayed in environments that implement middleware Ginga-NCL, as the Brazilian Digital TV System (BSDTV) and IPTV. The system SynchrLIBRAS was the main focus of this paper.

Since the introduction of BSDTV in 2007 in Brazil, it is expected the use of interactivity (return channel via the Internet), scheduled for 2013. Many benefits must be introduced into the direction of the teaching-learning by Digital TV (t-learning). The merger of Digital TV and the Internet is already being introduced in Brazil and also

facilitate the convergence of digital educational applications. Whereas BSDTV is also transmitted to mobile devices, the proposed solution may be used for m-Learning, considering certain technical restrictions of the device´s display.
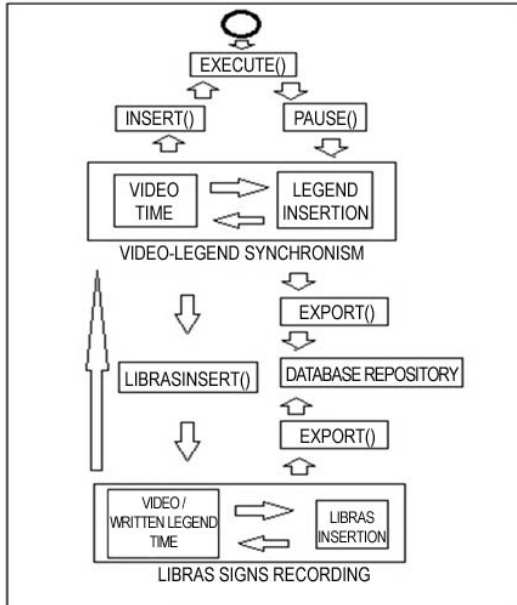

Fig. 2 Activity diagram of the SynchrLIBRAS System.

Fig. 3 Video from YouTube as input to start subtitling.

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

Fig. 4 Script to subtitle insertion in XML file.

```
...
<?xml version="1.0" encoding="UTF-8"?>
 <tt xmlns="http://www.w3.org/2006/04/ttaf1"
xmlns:tts="http://www.w3.org/2006/04/ttaf1#styling" xml:lang="pt">
 <head>
  <styling>
            ...
  </styling>
 </head>
 <body id="thebody" style="defaultCaption">
  <div xml:lang="pt">
      <p begin="00:02:00.00" end="00:04:20.00">Phrase 1...</p>
      <p begin="00:05:00.00" end="00:07:40.00">Phrase 2...</p>
      <p begin="00:07:50.00" end="00:08:55.00">Phrase 3...</p>
      <p begin="00:10:00.00" end="00:11:30.00">Phrase 4...</p>
  </div>
 </body>
 </tt>
...
```

Fig. 5 XML script containing subtitle and temporal information.

```
...function getPlayerTime(){
            var time = ytplayer.getCurrentTime();
            if (time<0){
                  return "00:00:00.00"; }
   var timestamp = parseInt(time/60);
            timestamp = timestamp < 10 ?'0'+timestamp :
timestamp;
   var secs = parseInt(time%60);
   timestamp += ':';
   timestamp += secs < 10 ?'0'+secs : secs;
            timestamp = "00:"+timestamp+".00";
            return timestamp; }
      function getSec(time){
            if(!time.match(/^\d\d:\d\d:\d\d.\d\d$/)){
                  return -1; }
            else{
                  return      sec =
parseInt(time.substr(0,2))*3600 + parseInt(time.substr(3,2))*60 +
parseInt(time.substr(6,2)); } }...
```
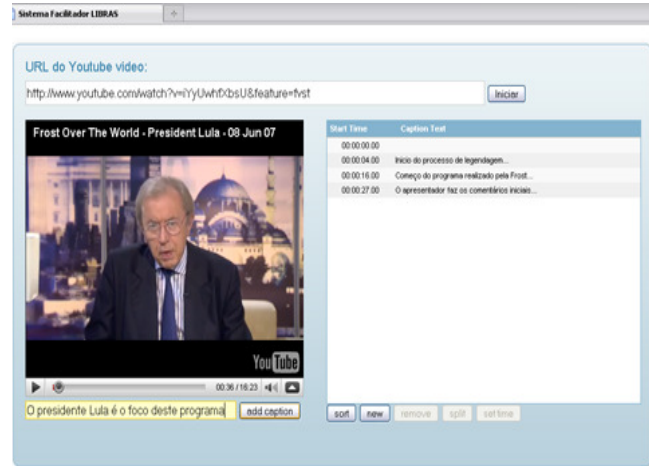
Fig. 6 Script to capture temporal information.

```
...
   function addLine(){
            var nextrow = getMaxId()+1;
            $("#list").addRowData( nextrow,
{id:nextrow,start:"00:00:00.00",caption:""}, "first" );
            return false;
      }
      function removeLine(){
            var id = $("#list").getGridParam('selrow');
            if( id != null ){
                  $("#list").delRowData(id);
            }
            else{
                  alert("Please Select Row to delete!");
            }
            return false;
      }
..
```

Fig. 7 Script for inserting subtitle via extra Button.

```
...
   function addCaption(){
     var timestamp = getPlayerTime();
                   var caption = $("#newcaption").val();
                   var nextrow = getMaxId()+1;
                   $("#list").addRowData( nextrow,
{id:nextrow,start:timestamp,caption:caption}, "last" );
                   sortGrid();
           $("#newcaption").attr("disabled","disabled").val("");
               $("#addcaption").attr("disabled","disabled");
                   var curSec = ytplayer.getCurrentTime();
                   var reSec = 1;
                   var gotoSec = curSec>reSec?curSec-reSec:0;
                   $("#videoplayer").focus();
                   ytplayer.seekTo( gotoSec, true);
                   ytplayer.playVideo();
     return false;
   }
```
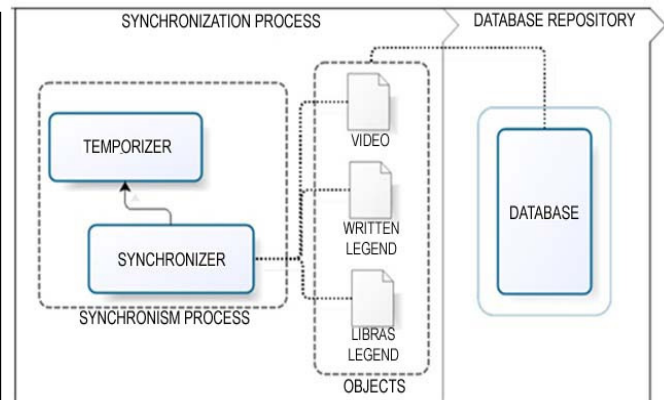
Fig. 8 Script for inserting subtitle.

Fig. 9 Synchronization and storage of multimedia objects.

References:

[1] G. A. Lira, *Deaf Education, Language and Digital inclusion*. Dissertation (Master Degree), Estacio de Sa University, Rio de Janeiro, Brazil, 2003.

[2] Brazilian Technical Standards Association (ABNT), *NBR 14.021: Transportation – Accessibility in urban or metropolitan train system,* Rio de Janeiro, 2005a, 39p.

[3] E. Fusco, X-Libras: an informative environment to the Brazilian Sign Language, *National Meeting of Research in Information Science, Cultural Diversity and Information Policies*, USP, Sao Paulo, Brazil, 2008.

[4] M. C. C. Baranauskas, M. T. E. Mantoan, Accessibility in Educational environments: Beyond the guidelines*, Rev. Online da Bibl. Prof. Joel Martins*, Vol.2, nº2, 2001, pp.13-23.

[5] R. Quadros, *Deaf Education: the Language acquisition,* Medical Art Ed., Porto Alegre, Brazil, 1999.

[6] United Nations Organization (UN), *Standard rules for opportunities equalization to disabled people – Salamanca Declaration,* A/RES/48/96, UN General Assembly. 1994

[7] Brazil, Decree nº 5296: *Regulates the Laws numbers 10048 and 10.098: general standards and criteria for the promotion of accessibility*, Union Official Diary, Brasilia, Brazil 2004.

[8] Brazil, L*aw 10.436:Brazilian Sign Language – Libras*, 2002, http:// www.planalto.gov.br/ccivil/leis/2002/L10436.htm.

[9] State Secretariat of Education (SEE), *Inclusion: Special Education Journal*, Vol.1, nº1, 54p.

[10] National Federation on Education and Integration of the Deaf (FENEIS), 2005, http://www.feneis.com.br/.

[11] World Wide Web Consortium (W3C), *Web Acessibility Initiative: WAI: Strategies, guidelines, resources to make the Web accessible to people with disabilities,* 2005, http://www.w3.org/WAI/ guidtech.html.

[12] C. R. Ramos, *Libras: The Brazilian Deafs sign language,* Arara Azul Ed., Petrópolis, Brazil, 2009, http://www.editora-arara-azul.com.br/Artigos.php.

[13] C. Alfaro, M. C. P. Dias, *Automatic Translation: a tool to aid the translator,* Cadernos de Traduçã*o* n°3, Centre of Communication and Expression, Federal University of Santa Catarina, Brazil, 1998.

[14] Brazilian Technical Standards Association (ABNT), *NBR 15.290: Acessibility in communication at television ,*Rio de Janeiro, Brazil, 10p. 2005, http://www.crears.org.br/crea/downloads/acessibilidade/NBR15290.pdf.

[15] Semiotic Engineering Research Group (SERG), *Manual of Audio-visual interactive Software development using authoring tools*, PUC, Rio de Janeiro, Brazil, 2006.