CineLibras: Automatic Generation and Distribution of Libras Tracks for Digital Cinema Platforms

Leonardo A. Domingues[1], Tiago M. U. de Araújo[2], Danilo A. N. S. Silva[3], Felipe L. S. Ferreira[4], Guido L. de S. Filho[5], Fernando T. Paulino[6]

Federal University of Paraíba, Paraíba, Brazil
{leonardo.araujo, tiagomaritan, danilo, lacet, guido}@lavid.ufpb.br and fernandotpaulino@gmail.com

Abstract. Deaf people face many problems to execute their daily activities. These difficulties include barriers for both access information as well as communicating with people without disabilities. In this context, the main goal of this article is to identify the main problems faced by deaf people to access information in movie theaters and to propose a solution to better address their requirements. In this context, a computational system was developed to automatically generate and distribute accessible video tracks in Brazilian Sign Language (Língua Brasileira de Sinais - Libras) in cinema rooms. This solution uses mobile devices as secondary screens, in a way that deaf people can have access to the content presented in their natural way of communication. Finally, experiments were performed with groups of Brazilian deaf in order to ensure the viability of the proposed solution and the data collected are analyzed and discussed.

Categories and Subject Descriptors: C.2.4 [**Distributed Systems**]: Client/server, Distributed applications; D.2.2 [**Design Tools and Techniques**]: User interfaces; H.5.1 [**Multimedia Information Systems**]: Animations, Evaluation/methodology, Video; K.4.2 [**Social Issues**]: Assistive technologies for persons with disabilities

Keywords: digital cinema, accessibility, hearing impairment, Brazilian Sign Language, multimedia applications

1. INTRODUCTION

Information and Communication Technologies (ICTs) are rarely developed taking into account the specific requirements and needs of deaf people [Haddon and Paul 2001]. They communicate naturally using sign languages, but the support for these languages is rarely addressed in the design of TV, Web and Cinema technologies. When supported, they are generally limited to a window with a sign language human interpreter (wipe) for few contents. In consequence, deaf people have serious difficulties to communicate and access information.

However, the digitization of TV and Cinema, besides the improvement of audiovisual content, is also making possible the development and introduction of new services and applications on these platforms. These applications and services can expand the system functions, allowing users to participate in innovative environments, to access new information, to interact with audiovisual content, among others.

In Cinema, an example is the "13th Street Interactive Movie Experience" [Elkins 2013]. In this experiment, viewers interact with the movie character using their smartphone, taking part in history. More specifically, during the film, a character calls a viewer in the room, and asks him for help to make some decisions, such as "Which way should I follow: left or right?". Then, based on the viewer's

We would like to acknowledge for the support of Rede Nacional de Ensino e Pesquisa (RNP) and Ministério do Planejamento, Orçamento e Gestão (MPOG) of Federal Government of Brazil in partnership with the Núcleo de Pesquisa e Extensão do Laboratório de Aplicações em Vídeo Digital (NPE/LAViD) that helped greatly in developing this work.

decision, the film takes a different path. Thus, the film can follow a different route in each new session. The interaction is done through software that makes phone calls to a randomly chosen viewer and allows a "dialogue" between them through speech recognition techniques.

Services and technologies developed to provide access for people with disability in movies are another important example. Through these services and technologies, we can generate accessible content to deaf users, for example, and distribute them to a secondary display (e.g., tablet, smartphone, among others). Thus, deaf users could watch the translation of a movie in sign languages, their natural language of communication. In scientific literature, there are some papers developed to improve access for deaf people in different environments [Tambascia et al. 2012; Ferreira et al. 2011; West et al. 2009; Sony 2012]. Tambascia et al. [2012], for example, propose a solution to include deaf students in the educational environment. Their solution captures the teacher audio and translates it to the sign language grammar. Then, a sign language video stream is generated from a dictionary of signs (represented by a 2D avatar) and transmitted along with the text to the deaf students devices.

Ferreira et al. [2011] and Lemos et al. [2011], proposed a solution that generates Brazilian Sign Language (Libras) windows into the Digital TV context. The Libras window is generated from the machine translation of the closed caption (CC) track broadcasted by the TV station. The proposal is compatible with DTV middlewares compatible with the International Telecommunication Union (ITU) J.202 specification, i.e., it uses just Application Programming Interface (API) and components defined in this specification.

West et al. [2009] and Sony [2012] propose solutions addressing the inclusion of deaf people in movie theaters. To perform this task, they developed special glasses with appropriated microdisplays to display movie subtitles. The microdisplay uses special lenses that make subtitles appear to be floated. Thus, users can see both images (movie and subtitles) smoothly. However, according to Wauters [2005] and Menezes and Cavalcante [2008], deaf users have difficulties to read and write texts in spoken languages.

In order to reduce these problems and improve access for deaf people to Cinema, in this article, we propose a practical and effective solution called CineLibras, which aims to automatically generate and distribute Brazilian Sign Language (Libras) video tracks in movie theaters. The Libras video is automatically generated from the machine translation of movie subtitles and is distributed for deaf users in secondary screen devices (e.g., smartphones, tablet, glasses, among others).

To synchronize the Libras video track and the film, the movie player is responsible for initiating the process of translation and distribution of accessible contents. To build a proof of concept of this proposal, we used the Fogo Player [Aquino Júnior et al. 2013], one player based on software developed to display content with 4K resolution, as our movie player. It was chosen because it is open source and allows synchronization with new services.

The rest of article is organized as follows. Section 2 presents the related works. Section 3 presents some concepts related to the Fogo Player. The solution proposed in this article, called CineLibras, is presented in Section 4. Section 5 discusses some tests performed with deaf users to evaluate the proposal. Final remarks are presented in Section 6.

2. RELATED WORKS

As mentioned in Section 1, some works have been developed to address the communication needs of the deaf people [Tambascia et al. 2012; Sony 2012; West et al. 2009; Ferreira et al. 2011]. These works offer technological solutions for daily activities that make possible for deaf people to watch and understand television, movies, to improve their inclusion in educational environment, among others.

Other works are related to machine translation of spoken language contents to sign language contents [Haddon and Paul 2001; Veale et al. 1998; Morrissey 2008; Zhao et al. 2000; Fotinea et al. 2008;

Huenerfauth 2008; Huenerfauth et al. 2007; Anuja et al. 2009; San-Segundo et al. 2012; San-Segundo et al. 2008; San-Segundo et al. 2008]. Veale et al. [1998], for example, proposed a multilingual translation system for translating English texts into Japanese Sign Language (JSL), American Sign Language (ASL) and Irish Sign Language (ISL). The work applies and extends some Artificial Intelligence (AI) concepts to sign languages (SL), such as, knowledge representation, metaphorical reasoning, among others, but there is no testing or experimentation to evaluate the solution. Then, it is not possible to draw conclusions about its feasibility and translation speed and quality.

Zhao et al. [2000] developed an interlanguage-based approach for translating English text into American Sign Language (ASL). It analyses the input data to generate an intermediate representation (IR) using their syntactic and morphological information. Then, a sign synthesizer uses the IR information to generate the signs. However, as well as in [Veale et al. 1998], the solution lacks experimental evaluation. Morrissey [2008] proposed an example-based machine translation (EBMT) system for translating text into ISL. However, the data set was developed from a set of "children's stories", which restricts the translation for this particular domain.

Fotinea et al. [2008] developed a system for translating Greek texts into Greek Sign Language (GSL). This work uses a transfer-based approach for generate the sentences in GSL, but it main focus is the strategy of animation that explores the parallel structures of sign languages (e.g., the ability to present a hand movement with a facial expression simultaneously). To perform this task, a 3D avatar was developed to explore the parallel structure of sign languages. However, no testing or experimentation was conducted to evaluate it translation speed and quality.

Huenerfauth [2008] and Huenerfauth et al. [2007] proposed modeling classifiers predicate1 in a English to American Sign Language (ASL) translation system. Some tests performed with deaf users showed that contents exploring the use of classifier predicates (generated by the Huenerfauth solution) were significantly more natural, grammatically correct and understandable than the contents based on direct translation. The translation speed, however, was not evaluated by author.

Anuja et al. [2009] proposed a system for translating English speech into Indian Sign Language (ISL) focused on helping deaf people to interact in public places, such as banks and railroads. The system also uses a transfer-based approach for translating speech entries into ISL animations. This solution is restricted to a specific domain and according to authors it takes a long (and unacceptable) time to generate the translation (the time values were not described in the work).

San-Segundo et al. [2012], San-Segundo et al. [2008] and San-Segundo et al. [2008] proposed an architecture for translating speech into Spanish Sign Language (LSE) focused on helping deaf people when they want to renew their identity card or driver's license. This translation system consists of three modules: a speech recognizer, a natural language translator and an animation module. However, as well as in the work of Anuja et al. [2009], this solution is also restricted to a particular (or specific) domain and the time needed for translating speech into LSE (speech recognition, translation and signing) is around 8 s per sentence, which makes the solution unfeasible for real time domains (e.g., television).

These works can be separated in two classes: one class of works that translate speech in the source spoken language to the target sign language (i.e., they use speech recognition) [Anuja et al. 2009; San-Segundo et al. 2012; San-Segundo et al. 2008; San-Segundo et al. 2008], and other class that translate written texts to the target sign language (i.e., they do not use speech recognition) [Fotinea et al. 2008; Huenerfauth 2008; Huenerfauth et al. 2007; Morrissey 2008; Veale et al. 1998; Zhao et al. 2000]. However, all these works have some limitations. The class of works that use speech recognition [Anuja et al. 2009; San-Segundo et al. 2012; San-Segundo et al. 2008; San-Segundo et al. 2008], for example, are just applied to specific domains and are not efficient considering signing and translation speed. However, these solutions do not consider specific requirements for specific environments and are not the best approach for places with shared experience for both deaf and non-deaf people, such as

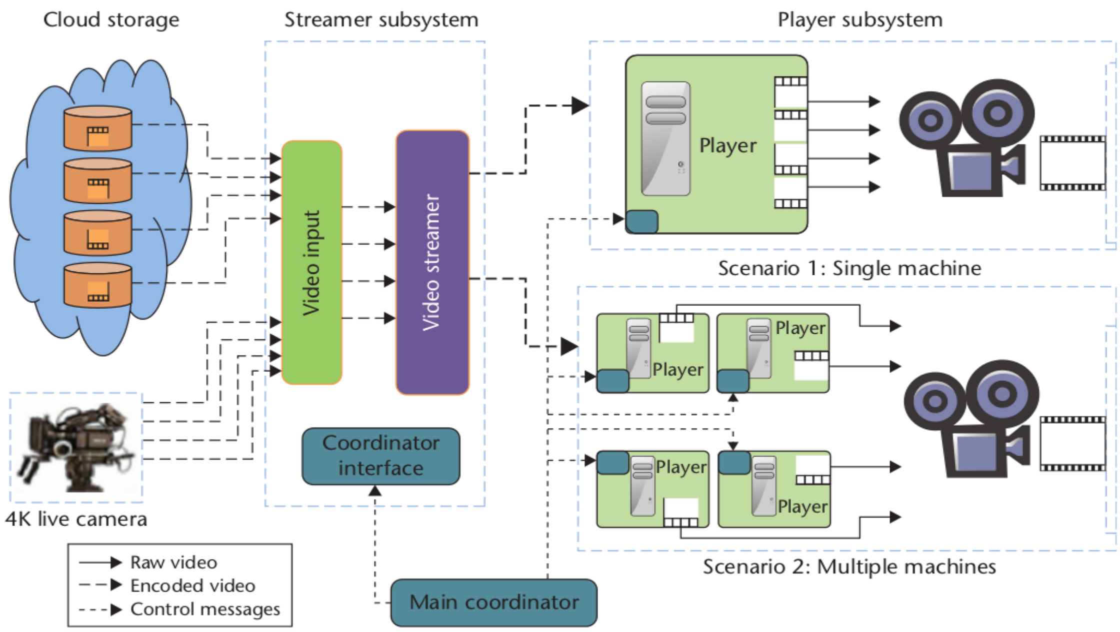

Fig. 1. Fogo Player distribution flow (taken from: [Aquino Júnior et al. 2013]).

Cinemas. In this scenario, the exhibition of sign language track over the movie content could disrupt non-deaf people. These limitations led to the development of our proposal, which will be presented in detail in the Section 4.

3. FOGO PLAYER

Fogo Player [Aquino Júnior et al. 2013] is a computational system that uses a scalable, flexible and low cost architecture to allow 4K-video (4096 x 2160 pixels) transmission and playback. Its architecture is shown on Figure 1. The system can use multiple content sources, which makes its use feasible in distributed multimedia applications. This technology makes possible to slice the UHD video in quadrants, that can be transmitted synchronously from different sources.

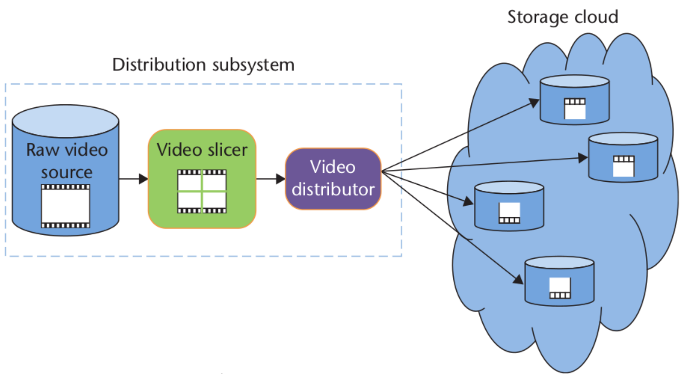

Fig. 2. Schematic view of the Streamer and Player subsystems. (taken from: [Aquino Júnior et al. 2013]).

According to Figure 1, Fogo Player contains four main subsystems: Distribution, Pre-processing, Streamer and Player. These subsystems have a set of software components to address the system requirements. Figure 2 shows a schematic view of the Distribution subsystem.

The Distribution subsystem receives a raw input video source, slices it in multiple quadrants and delivers them to an intermediate storage, allocated in a cloud storage infrastructure (see Figure 2). Then, a Pre-processing subsystem encodes them using the selected video encoder (e.g, H.264 or JPEG2000). The Streamer subsystem (see Figure 1) distributes the content to the clients.

The Player subsystem addresses two different approaches: single or multiple machines. In the single machine scenario, a single machine is used to receive, decode and display the content delivered by the Streamer. On the other hand, we can use multiple machines to manage the received content in the multiple machine approach. In this case, each machine could manage (i.e., receive, decode and display) a piece of video (e.g., a video quadrant). The Player is also responsible to create a management interface that receives the configuration messages needed to synchronize and display of the video.

The Main Coordinator component was created to remain the Fogo Player components synchronized (see Figure 1). It is responsible to preserve the virtual clock used between the Streamer and Player machines, and also allows the synchronization with other services. In our proposal, CineLibras will also be controlled by the Main Coordinator. Further information about Fogo Player can be found at [Aquino Júnior et al. 2013].

4. THE PROPOSED SOLUTION

In this section we describe the proposed solution, named CineLibras. As mentioned in Section 1, CineLibras aims to automatically generate and distribute Brazilian Sign Language (LIBRAS) video tracks in movie theaters. The idea is to generate the LIBRAS video track using the machine translation of the movie subtitle. This LIBRAS video track is then distributed to users using secondary screen devices (e.g. such as smartphone, tablet, among others).

In the proposed solution, the Fogo Player will be used to manage the presentation of the movie. It will be also responsible for signaling the start and end of presentation of the movie to the CineLibras, as well as to provide other information, such as the subtitle media used as input for the generation of the LIBRAS track, the synchronization information between the movie and the Libras track, among others. Beside it is scalable, flexible and has a low cost architecture, the Fogo Player was chosen because it provides an interface to communicate with other systems. This feature makes possible to aggregate new functions to the system, such as the proposed in this article.

The CineLibras will be detailed in Sections 4.1, 4.2 and 4.3. Section 4.1 discusses the conceptual model of the solution, Section 4.2 shows the communication protocol between CineLibras and Fogo Player and Section 4.3 describes the architecture and software components of CineLibras.

4.1 Conceptual Model of the Solution

A movie theater is basically composed of two environments: Control and Exhibition. The Control environment is responsible to control the exhibition of the film. The Exhibition environment is a place where users watch these contents. Figure 3 illustrates the conceptual model of the proposed solution in a movie theater.

According to Figure 3, the CineLibras and Fogo Player subsystems works in the Control environment. In this environment, the Fogo Player manages the system exhibition, and CineLibras generates the Libras video track and distribute for clients in a secondary screen device. In this proposal, a protocol was defined to allow the communication between CineLibras and Fogo Player. This protocol

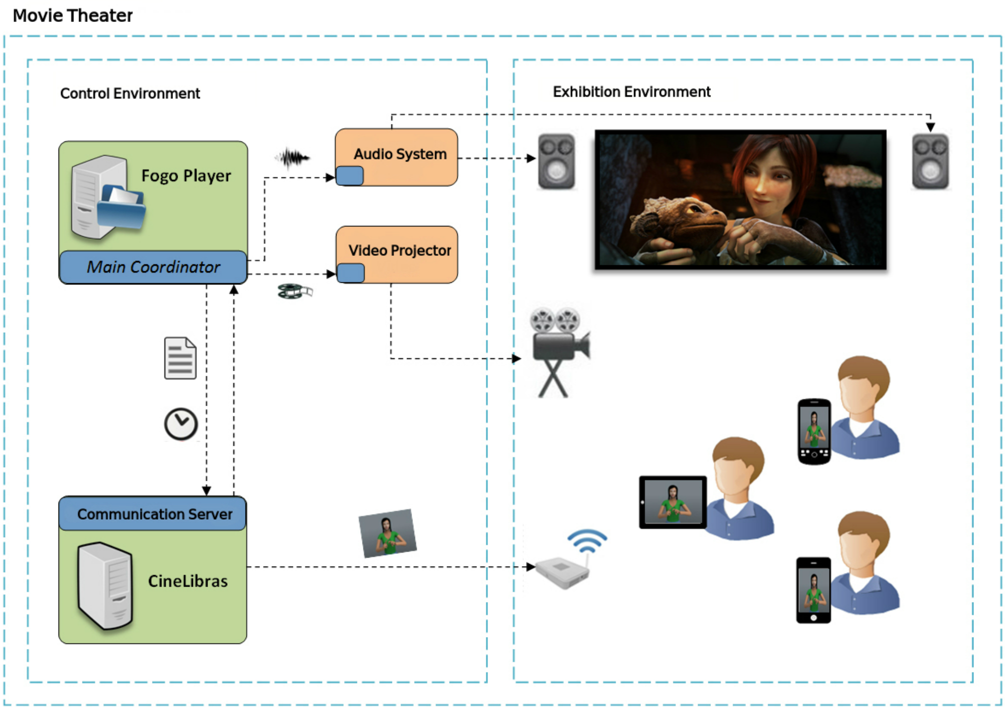

Fig. 3. Conceptual model of the proposal solution.

allows Fogo Player and CineLibras communicate to coordinate system execution, managing the exhibition of the contents; to exchange the subtitle information for generating the LIBRAS video track; to synchronize the LIBRAS video track and the movie content, among others. Further information about the protocol are presented in Section 4.2.

The synchronization between LIBRAS video track and the movie content is made calculating a PTS (Presentation Time Stamp) for each subtitle. As mentioned in Section 1 and Section 4, the subtitle files are used to generate the LIBRAS video tracks. The process includes software components to extract the subtitles, to machine translate the subtitle text to an intermediate sign language representation (i.e., a sequence of glosses), to synthesize the LIBRAS video track from these sequence of glosses and, finally, distribute the accessible video track for clients. The distribution of accessible videos is made using a Wireless Local Area Network (WLAN), based on IEEE 802.11g specification.

To receive and display the Libras track, the mobile devices must connect to the CineLibras WLAN and have a video player with support to a MPEG-2 Transport Stream (MPEG-TS). Thus, deaf users can watch the LIBRAS translation in their devices, without disturbing non-deaf people. It is important to point out that, in this solution, the LIBRAS signs are represented by a 3D virtual animated agent (i.e., a 3D avatar).

4.2 Communication Protocol

As mentioned in Section 4.1, CineLibras and Fogo Player use a protocol to communicate. This protocol allows them to exchange messages, share information and resources, among others. More specifically, CineLibras receives the subtitle track (used by the machine translation), some control commands (e.g, play and stop), synchronization information, among others, using this protocol. The general structure of messages is shown in Figure 4.

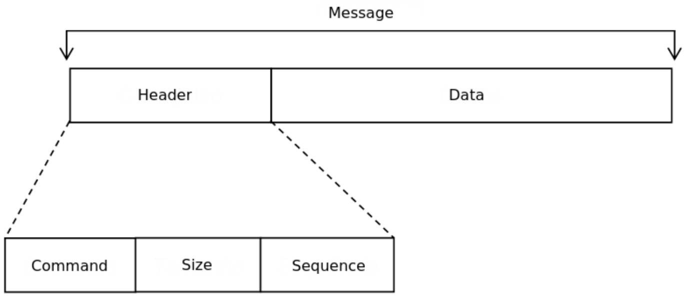

Fig. 4. Structure of messages used in communication between CineLibras and Fogo Player.

According to Figure 4, a message consists of two blocks: Header and Data. The Header block includes the following fields: Command, Size and Sequence. The Command field represent the type of message. The Size field represent the size of the message in bytes, and the Sequence field represents is used to identify the number of the message when it needs to be fragmented. The Data block is used to encapsulate the message data (e.g., the synchronization information, the subtitle files, among others).

Figure 5 shows a sequence diagram depicting the exchange of messages between Fogo Player and CineLibras using the proposed protocol.

According to Figure 5, initially, CineLibras requests authentication on Fogo Player, using the IDENTIFY command associated with its UID (Unique Identification) - SERVICE_LIBRAS. In Fogo Player, when this authentication message is received, it checks the UID and, if valid, sends a response message

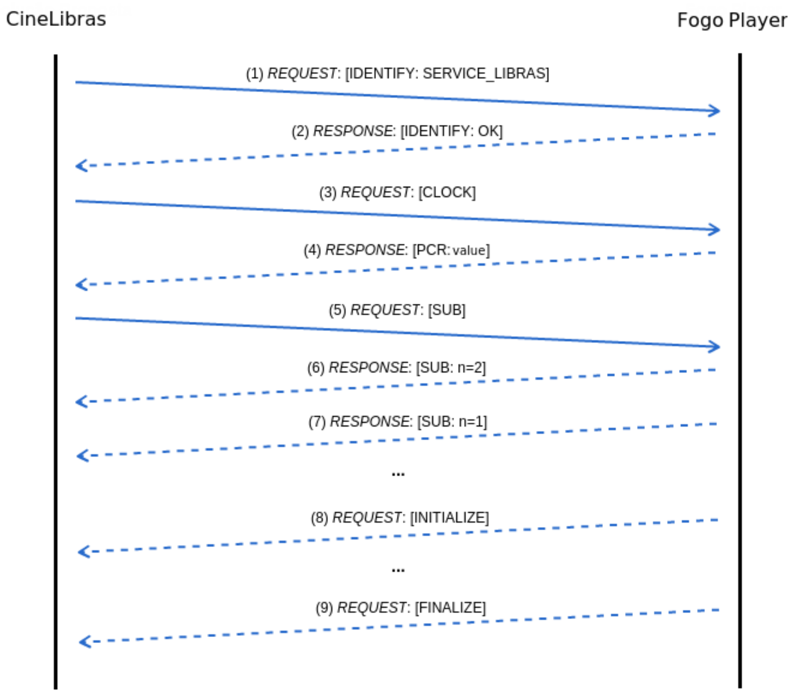

Fig. 5. Exchange of messages between Fogo Player and CineLibras using the proposed protocol.

indicating that the authentication was successful. To perform this task, it uses an "OK" identifier. Otherwise, if the UID is not valid, it sends a message with an "ERROR" identifier. If CineLibras is authenticated, it may requests new information and resources from Fogo Player. These data and resources include the subtitle track of the film, the synchronization information between CineLibras and Fogo Player, among others. Some of these messages are illustrated in Figure 5. In addition, CineLibras receives some control messages from Fogo Player. These messages are used, for example, to indicate the beginning and end of the movie, as well as update notification on the logical clock.

On Fogo Player, the Main Coordinator (see Figure 1) is the component responsible for managing the communication with CineLibras. As will be presented in the Section 4.3, on CineLibras, the component responsible for this is the Communication Server. More details about the CineLibras components will be presented in Section 4.3.

4.3 CineLibras Architecture

As mentioned in Section 1, CineLibras consists of a set of software components responsible for machine translation of movie subtitles and distributing them to users' secondary screen devices on Cinemas. Figure 6 shows the architecture of CineLibras.

According to Figure 6, the CineLibras architecture consists of six main components: (1) Communication Server; (2) Controller; (3) Subtitle Extractor; (4) Translator; (5) Synchronizer; and (6) Distributor. The solution has also a database, called Libras Dictionary. This Dictionary is a repository of MPEG2TS-encoded videos, where each video is a sign in Libras, represented by a 3D-avatar.

The Communication Server component is responsible for exchanging messages with the Fogo Player based on the protocol defined in Section 4.2. The Controller component coordinates the process of generation and distribution of LIBRAS video tracks.

To perform this task, initially, the Controller requests the Communication Server communication to get subtitles track and the logical clock from Fogo Player. When the subtitle track is received, it is delivered to the Subtitle Extractor component that extracts the text from subtitles alongside with their timestamps, i.e., Time-in and Time-Out. In this implementation, we used subtitle files specified by the DCSS (Digital Cinema System Specification).

Based on the logical clock received from Fogo Player and on the subtitle timestamps, the Controller manages when subtitles must be translated and distributed. Then, the Controller sends the subtitle text to the Translator and Synchronizer components, which are responsible, respectively, for the machine translation of the text in Brazilian Portuguese to a sequence of glosses in LIBRAS (i.e., an intermediate representation in Libras), and for converting a sequence of glosses to a LIBRAS video track synchronized with the movie.

The machine translation method used in the Translator component was developed by Araújo [2012]. This machine translation approach combines statistical compression methods to classify words, with textual simplification strategies to reduce the complexity of the input text and a set of syntactic and morphological rules defined by LIBRAS specialists [Araújo 2012].

In order to generate the LIBRAS video track, the Synchronizer queries the LIBRAS Dictionary to associate every gloss token to its corresponding video (a LIBRAS sign). If the sign is not found in the dictionary, it is spelled. This process consists of displaying each letter of a gloss signal individually. For example, in LIBRAS, the word "CINELIBRAS" does not have a corresponding sign. Then, it is spelled and represented as "C-I-N-E-L-I-B-R-A-S". This is the same strategy used by deaf and LIBRAS interpreters to represent words in Brazilian Portuguese that do not have a corresponding sign, such as proper names and technical terms.

In the intervals without dialogue (i.e, when there is no caption to be displayed), the Synchronizer adds a video with the Avatar in a neutral pose. The neutral position was defined by LIBRAS special-

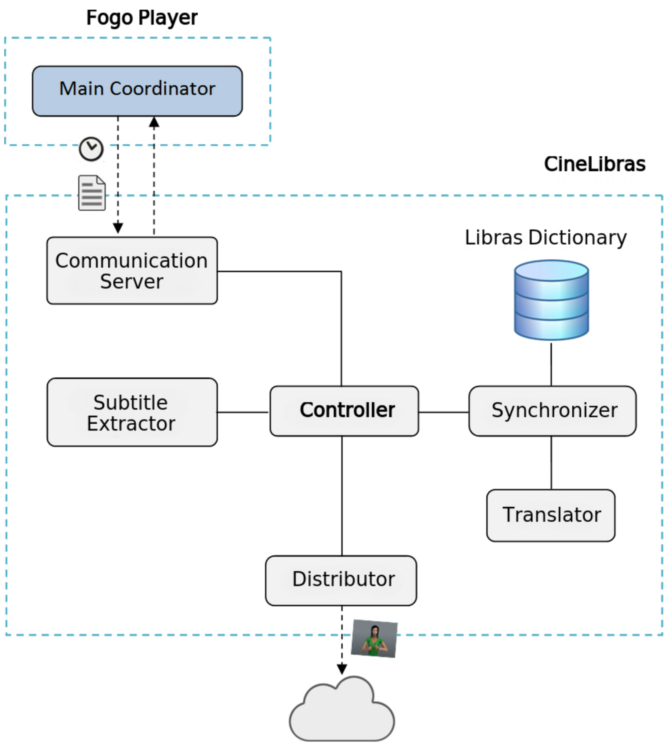

Fig. 6. CineLibras components architecture.

ists, placing the hands and arms extended in a straight line down and with a neutral facial expression (i.e., without applying movement in the facial bones). This approach is also used by LIBRAS interpreters and its use helps to avoid technical issues. Most video players would drop the connection with the Distributor if not receiving any video stream during the silent intervals.

After the generation of the LIBRAS video, the Controller feeds the buffer of the Distributor component with data segments of this video. Then, the Distributor transmits the Libras video segment on a local network (WiFi) previously configured inside the movie theater.

To distribute the Libras videos, we used an approach based on reflectors. This approach was used to keep the maximum synchronization between users secondary screen devices. The idea is to use the same data source and replicate them to all connected devices. Thus, the data is sent periodically in seven TS packet units, each having 188 bytes, totaling 1316 bytes at a time in each IP packet. Using this strategy, we probably reduce the transmission delay between the clients, since the number of bytes transmitted is below the capacity of the Maximum Transmission Unit (MTU) of most datagram based networks, which normally uses a MTU of 1500 bytes.

On the client side, the secondary screen devices must be connected to the CineLibras network to receive the LIBRAS content. In addition, they need to decode and display video streams in MPEG2-TS format. Figure 7 shows some photos of a demonstration of the proposed solution performed in the CineGrid 2012 Workshop, held on San Diego, California, USA.

To evaluate the solution, in Section 5, we present some tests performed with the proposed solution. These tests aimed to evaluate the quality of content generated by the proposed solution as well as to investigate the satisfaction of deaf users on the use of a secondary screen device to access the translation of the movie.

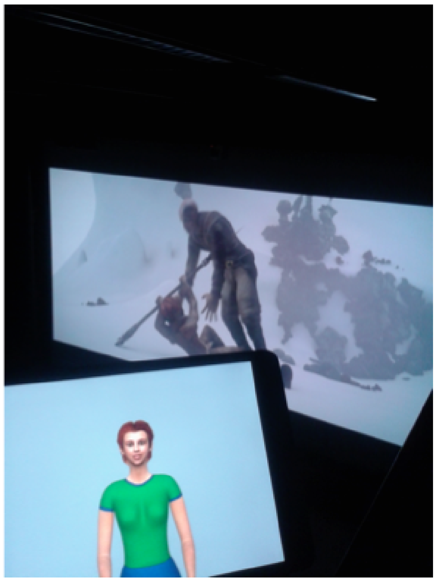

Fig. 7. Demonstration of CineLibras in the CineGrid 2012 Workshop.

5. RESULTS AND DISCUSSIONS

The tests to evaluate the proposed solution were performed with some Brazilian deaf users in two rounds. In the first round, we focused on assessing the intelligibility of the LIBRAS contents generated by the proposed solution and the satisfaction of deaf users in the use of this solution. In the second one, we performed an assessment focused on user experience to complement the first evaluation. These tests are presents in Sections 5.1 and 5.2, respectively.

5.1 Evaluation of the intelligibility of contents

Initially, we performed some tests with Brazilian deaf users to evaluate the intelligibility of the LIBRAS contents generated by the proposed solution as well as the satisfaction of deaf people in use of the solution. The tests were made with 20 (twenty) Brazilian deaf, where 7 were men and 13 were women, aged between 13 to 56 years old, with an average age of 30.4 years. We also observed the educational level of users, which include users with incomplete primary education until higher incomplete education.

The tests were performed at two places, in two different steps. The first step was performed at the Foundation to Support People with Disabilities (FUNAD), located at João Pessoa, Paraíba, Brazil. In this step, 10 Brazilian deaf users watched and evaluated the movie with subtitles in Brazilian Portuguese. In the second step, the tests were performed at the National Institute of Deaf Education (INES), located at Rio de Janeiro, Brazil. This step also involved 10 Brazilian deaf that evaluate movie along with the LIBRAS video track generated by the proposed solution. To follow the movie translation generated by CineLibras, these users used an iPad 2 mobile device as their secondary screen device. They also used the VLC Media Player application to receive the accessible contents generated.

The evaluation process was the same in both steps. Initially, we presented the main objectives of the test to the deaf users. Then, users filled out a form with some personal information (age, education level, gender, among others), and were invited to watch the movie Sintel. Sintel is a 15-minute animation movie in 4K resolution (i.e., UHD) created by Blender Foundation.

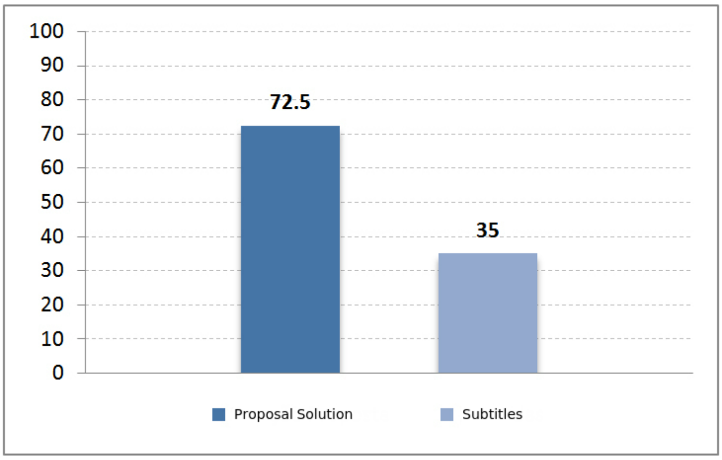

Fig. 8. Results of the intelligibility tests with Brazilian deaf users.

After watching the movie, the users were invited to answer some questions. These questions were divided in two parts. In the first part, users answered four questions related to the movie content to assess their level of comprehension. In these questions, users have to select which of four alternatives (A, B, C or D) is related to the content presented, where only one of the alternatives is correct. For all questions, the fourth alternative (D) represented a "I do not know" option, which was included to prevent users randomly choose one of the options when they did not know the correct answer.

The second part of the form was applied only for the deaf group that watched the movie with the LIBRAS video track generated by the proposed solution on a mobile device. The purpose of was to investigate the level of satisfaction of the users when using the the secondary screen device to watch the LIBRAS translation. Afterwards, users answered two questions related to: (1) the level of difficulty using a mobile device; and (2) the level of complexity to watch the movie and the LIBRAS track in the mobile device at the same time. These questions were answered rating these aspects on a 1-to-6 scale. During tests, LIBRAS interpreters mediated the communication with the users. Figure 8 shows the results of the intelligibility tests.

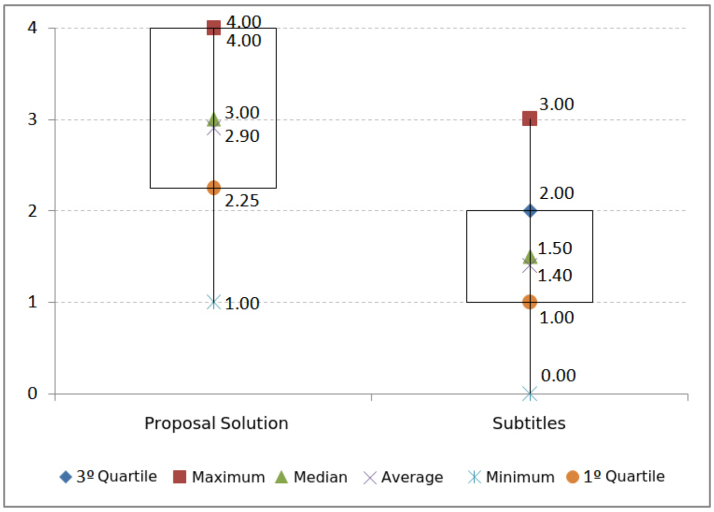

Fig. 9. Dispersion analysis of the results between the investigated approaches.

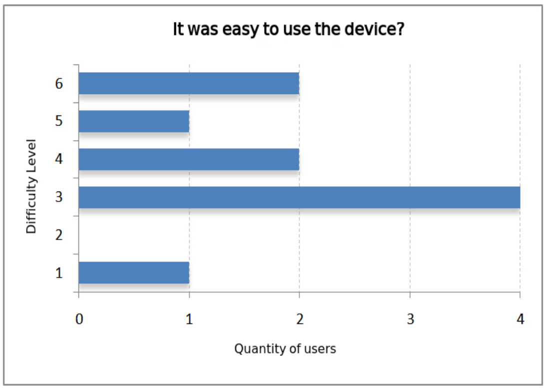

Fig. 10. Results of the evaluation of the level of difficulty using a mobile device.

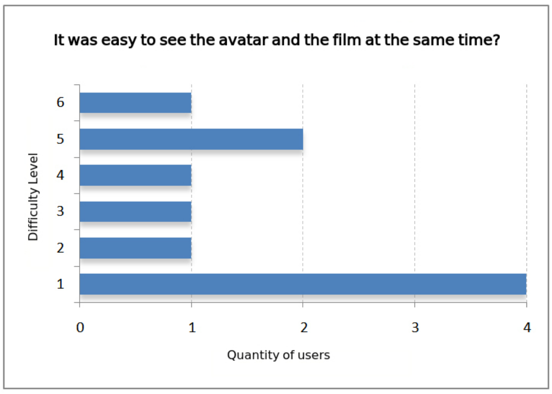

Fig. 11. Level of difficulty to watch the two screens simultaneously.

According to Figure 8, we observed that Brazilian deaf users have serious difficulties to understand the movie content when the information was a subtitle in Brazilian Portuguese. In this case, the average rate of correct questions was 35%. On the other hand, the users that watched the movie along with the LIBRAS track generated by CineLibras had an average rate of correct questions of 72.5%.

Observing the dispersion of results obtained in the comprehension test (presented in the Figure 9 graph), we conclude that the group that evaluated the proposed solution had little variation in their results. The median, first and third quartile values of this distribution were 3.0, 2.25 and 4.0, respectively. The dispersion of results obtained from the use of subtitles shows that deaf people had a low average rate of correct questions. The values of the median, first and third quartile (1.5, 1.0 and 2.0, respectively) were lower than the correspondent values for the proposed solution. For this group, a user hit 75% of the questions, but two users do not have any hit, (0%), which illustrates the difficulty of some deaf to understand text information.

Figures 10 and 11 presents the results for the second part of the form. According to Figure 10, we can observe that most deaf users answered there was no major difficulties in using a mobile device technology. To this question, the average rate was 3.8 with a standard deviation of 1.55.

Fig. 12. Moments during the intelligibility tests with CineLibras.

Regarding the level of complexity to watch the two screens simultaneously, the results (see Figure 11) show that users have some difficulties to watch the Libras track in the secondary screen. According to the answers provided by the users, the average difficulty was rated as 3.1. However, it is important to point out that they had a good result in the comprehension tests. Figure 12 shows the deaf users using CineLibras during testing at the National Institute of the Deaf (INES), held in Rio de Janeiro, Brazil.

As stated in Section 2, most related work [Tambascia et al. 2012; Sony 2012; West et al. 2009] involves text-based solutions to provide access to provide access of deaf to movie theaters. However, in the tests with the proposed solution, we observed that the deaf users really have difficulty understanding text-based content (they have an average hit rate of 35% in comprehension tests), and have better understanding of Libras content, if that content had been generated automatically and in a secondary screen device (average hit rate of 75% in comprehension tests). Therefore, it is a good indication that the proposed solution generates more understandable contents than related solutions [Tambascia et al. 2012; Sony 2012; West et al. 2009].

Thus, the proposed solution can be seen as an effective, low cost and practical solution to address the issue of access of deaf people to Digital Cinemas.

To complement this evaluation with respect to the user satisfaction, in the next section, we will present another round of tests focused on evaluate the user experience with our solution.

5.2 User Experience with the Proposed Solution

After the first round of tests described in Section 5.1, we planned a new round of tests to evaluate more deeply the user experience with the proposed solution. In this step, a new set of 10 (ten) Brazilian deaf users took part on the evaluation. This group consists of 7 (seven) men and 3 (three) women, aged between 11 and 20 years. Among them, 3 (three) users said they go to the Cinema regularly, and 9 (nine users) said they use smartphones to access social networks.

This round of tests were performed on November 17, 2014 at the Municipal School of Basic Education Índio Piragibe, in João Pessoa, Paraíba, Brazil. During the evaluation process, one deaf collaborator with lip-reading skills was used to make easier the communication between users and evaluators.

In addition, in this new round of experiments, we adjust the experiment, developing a new form, defining the evaluator's role during the tests, and the evaluation method to be applied. Initially, the form was developed in Brazilian Portuguese and then it was translated to LIBRAS, allowing users to see the questions in their natural language of communications. The type of answer options in the questionnaire was also changed. In this new round, instead of choosing an option in a numerical scale, the users had to select a graphical representation (emoticons) to assess their experience with CineLibras. These emoticons represent two alternatives (Yes or No).

Table I. Qualitative evaluation of the result of CineLibras.

Question	**Answer**	
	Yes	**No**
Do you like to use the application with the movie?	100%	0%
It was easy understand the movie?	30%	70%
The translation was correct?	20%	80%
Would you use the application again?	100%	0%
The avatar facial expression is good?	100%	0%

Initially, a member of the team explained to the users the main goal of the evaluation: investigate whether CineLibras could help, or not, deaf people in their daily activities. After this presentation, the users were also invited to watch the Sintell movie (also used in the first round of tests). After the movie presentation, they were invited to fill the proposed form. In this case, for each question presented, there was a movie with LIBRAS representation and the two alternatives (Yes or No) graphed. Table 1 presents some results collected from the users.

According to Table I, users liked to use the application, the avatar facial expressions were considered good and they would use the application again to watch a movie. However, they had some difficulties to understand the movie content (only 30% said that it is easy to understand the content), and the translation had also problems (80% of them related problems with the translation).

One possible explanation for the issue to understand the movie content is the exhibition of Libras video on a secondary screen device. In this case, users have to switch focus from the main screen to the mobile device, and vice versa. This information was collected informally, after the test, when a deaf students said: "(...) during an action scene, for example, it is difficult to change the focus of the main scene and look at the mobile device." However, we believe that it can be an ergonomic problem, and not a limitation of the solution. Maybe, if the secondary screen device was positioned in the same region of the movie subtitle, users can watch the Libras contents more easily. Another alternative would be the exhibition of Libras video in a glass. Further investigation on the ergonomics of the solution is a proposal for future work.

5.3 Comparing Accessible Mobile Technologies

Continuing the second round of tests, we also compared CineLibras with two machine translation solutions developed to provide accessible information for deaf users on mobile devices, Handtalk [HandTalk 2012] and Prodeaf [ProDeaf 2014]. The Handtalk is an application for device mobile developed with the purpose of translating Portuguese to Libras contents. The translating of this contents can be made through text entries, voice recognition, or through Optical Character Recognition (OCR) for text found in images [HandTalk 2012]. The Prodeaf has the same features as the Handtalk. One

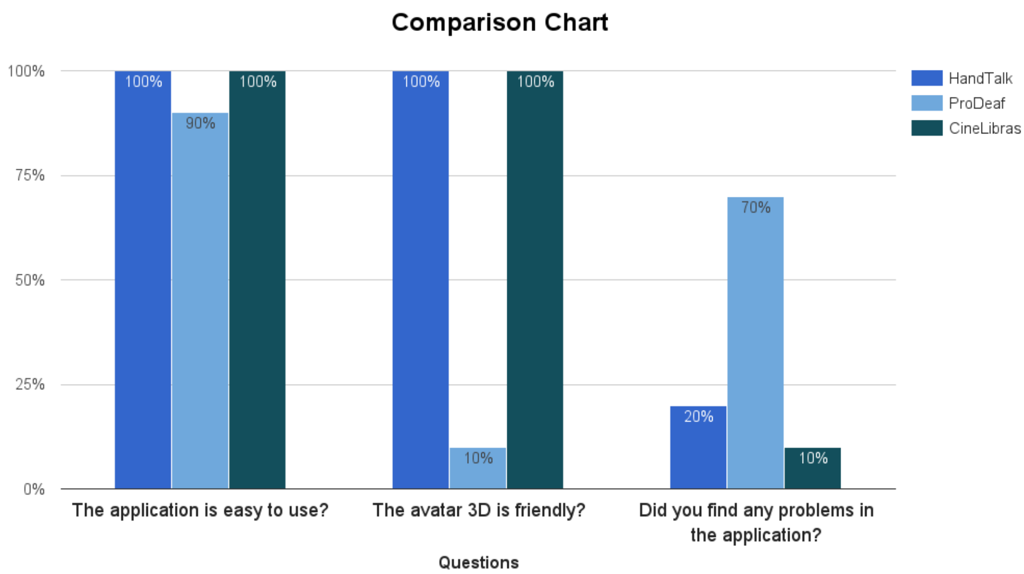

Fig. 13. Comparing the leading mobile accessibility tools for the deaf.

difference between the applications is that the Prodeaf locally has a portion of the dictionary, where the goal is to speed the recovery of signals. The result of the comparison between the Handtalk, Prodeaf and CineLibras is shown in Figure 13.

As shown in Figure 13, deaf users answered three questions about the applications used during the experiment: (1) *The application is easy to use?*; (2) *The avatar 3D is friendly?*; and (3) *Did you find any problems in the application?*. All applications were considered easy to use.

Because their hearing impairment, deaf usually become more demanding on aspects of visual representation, such as, GUI designs, combination of colors, correlation between objects, and others. The second point investigated in this comparison was regarding the avatar (used for LIBRAS signing) appearance. According to Figure 13, the ProDeaf [ProDeaf 2014] application had rejection, hitting only 10% of acceptance.

One of the critical points of this research was related to the problems that deaf people found while using the applications. Regarding this item, 20%, 70% and 10% of the users mentioned problems with [HandTalk 2012], [ProDeaf 2014] and CineLibras, respectively.

6. FINAL REMARKS

This article presented a technological solution to improve accessibility for deaf in movie theaters. The proposal is that LIBRAS video tracks be automatically generated from the movie subtitles and delivered as a video stream in a secondary screen device, allowing deaf users to watch the translation of the movie in their natural language of communication without disturbing non-deaf people.

From this proposal, we have developed a prototype of the proposed solution that was used in some tests involving Brazilian deaf users. The results of these tests show that the deaf using the prototype had a better understanding of the movie than the ones reading subtitles. Thus, the solution presented in this article was able to help them in this scenario.

Finally, for future studies it is an option to investigate the use of different devices for projecting the LIBRAS window. Moreover, it is intended to investigate the automatic generation from audio tracks and movies with subtitles in other languages.

REFERENCES

Anuja, K., Suryapriya, S., and Idicula, S. M. Design and Development of a Frame Based MT System for English-to-ISL. In *Proceedings of the World Congress on Nature & Biologically Inspired Computing*. Coimbatore, India, pp. 1382–1387, 2009.

Aquino Júnior, L., Gomes, R., Silva Neto, M., Duarte, A., Costa, R., and Souza Filho, G. A Software-based Solution for Distributing and Displaying 3D UHD Films. *IEEE MultiMedia* 20 (1): 60–68, 2013.

Araújo, T. M. U. *Uma Solução para Geração Automática de Trilhas em Língua Brasileira de Sinais em Conteúdos Multimídia.* Ph.D. thesis, Universidade Federal do Rio Grande do Norte (UFRN), Brazil, 2012.

Elkins, C. Creating an Interactive Experience in the Movie Theater. http://the2ndscreen.tv/creating-an-interactive-experience-in-the-movie-theater/, 2013.

Ferreira, F. L. S., Lemos, F. H., Botelho Neto, G. P., de Araújo, T. M. U., and de Souza Filho, G. L. Providing Support for Sign Languages in Middlewares Compliant with ITU J.202. In *IEEE International Symposium on Multimedia*. João Pessoa, Brazil, pp. 149–156, 2011.

Fotinea, S.-E., Efthimiou, E., Caridakis, G., and Karpouzis, K. A Knowledge-based Sign Synthesis Architecture. *Universal Access in the Information Society* 6 (4): 405–418, 2008.

Haddon, L. and Paul, G. Design in the IT Industry: the role of users. In R. Coombs, K. Green, A. Richards, and V. Walsh (Eds.), *Technology and the Market: demand, users and innovation*. Edward Elgar, Cheltenham, England, pp. 201–215, 2001.

HandTalk. HandTalk. http://www.handtalk.me, 2012.

Huenerfauth, M. Generating American Sign Language Animation: overcoming misconceptions and technical challenges. *Universal Access in the Information Society* 6 (4): 419–434, 2008.

Huenerfauth, M., Zhao, L., Gu, E., and Allbeck, J. M. Evaluating American Sign Language Generation through the Participation of Native ASL Signers. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility*. Tempe, USA, pp. 211–218, 2007.

Lemos, F. H., Ferreira, F. L. S., Botelho Neto, G. P., de Araújo, T. M. U., and de Souza Filho, G. L. Uma Proposta de Protocolo de Codificação de LIBRAS para Sistemas de TV Digital. In *Proceedings of the Brazilian Symposium on Multimedia and the Web*. Florianópolis, Brazil, 2011.

Menezes, D. C. and Cavalcante, T. C. F. Compreensão de Textos Escritos: um estudo com adolescentes surdos. *Distúrbios da Comunicação* 20 (1): 29–38, 2008.

Morrissey, S. *Data-driven Machine Translation for Sign Languages*. Ph.D. thesis, Dublin City University (DCU), Dublin, Ireland, 2008.

ProDeaf. ProDeaf. http://prodeaf.net, 2014.

San-Segundo, R., Barra, R., Córdoba, R., D'Haro, L. F., Fernández, F., Ferreiros, J., Lucas, J. M., Macías-Guarasa, J., Montero, J. M., and Pardo, J. M. Speech to Sign Language Translation System for Spanish. *Speech Communication* 50 (11-12): 1009–1020, 2008.

San-Segundo, R., Montero, J. M., Córdoba, R., Sama, V., Fernández, F., D'Haro, L. F., López-Ludeña, V., Sánchez, D., and García, A. Design, Development and Field Evaluation of a Spanish into Sign Language Translation System. *Pattern Analysis and Applications* 15 (2): 203–224, 2012.

San-Segundo, R., Montero, J. M., Macías-Guarasa, J., Córdoba, R., Ferreiros, J., and Pardo, J. M. Proposing a Speech to Gesture Translation Architecture for Spanish Deaf People. *Journal of Visual Languages & Computing* 19 (5): 523–538, 2008.

Sony. Sony Digital Cinema. http://us.digitalcinema.sony.com/pro/products/digital-cinema, 2012.

Tambascia, C. A., Rocha, R., Runstein, F. O., and Costa, H. M. Solução para Comunicação e Interação com Deficientes Auditivos em Sala de Aula. In *Encontro Nacional de Informática e Educação*. Cascavel, Brazil, pp. 115–122, 2012.

Veale, T., Conway, A., and Collins, B. The Challenges of Cross-modal Translation: English-to-sign-language translation in the Zardoz system. *Machine Translation* 13 (1): 81–106, 1998.

Wauters, L. N. *Reading Comprehension in Deaf Children: the impact of the mode of acquisition of word meanings*. Ph.D. thesis, Radboud University, Nijmegen, Netherlands, 2005.

West, L. L., Adler, E. W., Jo, J., Stewart, J. M., and Wood, J. W. Display Design Trade-offs for a Wireless Mobile Captioning System. `https://www.researchgate.net/publication/237431444_Display_Design_Trade-offs_for_a_Wireless_Mobile_Captioning_System`, 2009.

Zhao, Z., Kipper, K., Schuler, W., Vogler, C., Badler, N., and Palmer, M. Machine Translation System from English to American Sign Language. In J. S. White (Ed.), *Envisioning Machine Translation in the Information Future*. Lecture Notes in Artificial Intelligence, vol. 1934. Springer, pp. 54–67, 2000.