

Signing avatars: making education more inclusive

José Mario De Martino¹ · Ivani Rodrigues Silva² · Carmen Zink Bolognini³ ·
Paula Dornhofer Paro Costa¹ · Kate Mamhy Oliveira Kumada² · Luis Cláudius Coradine⁴ ·
Patrick Henrique da Silva Brito⁴ · Wanessa Machado do Amaral¹ · Ângelo Brandão Benetti⁵ ·
Enzo Telles Poeta¹ · Leandro Martin Guertzenstein Angare¹ · Carolina Monteiro Ferreira¹ ·
Davi Faria De Conti³

Published online: 3 November 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract In Brazil, there are approximately 9.7 million inhabitants who are deaf or hard of hearing. Moreover, about 30% of the Brazilian deaf community is illiterate in Brazilian Portuguese due to difficulties to offer deaf children an inclusive environment based on bilingual education. Currently, the prevailing teaching practice depends heavily on verbal language and on written material, making the inclusion of the deaf a challenging task. This paper presents the author's approach for tackling this problem and improving deaf students' accessibility to written material in order to help them master Brazilian Portuguese as a second language. We describe an ongoing project aimed at developing an automatic Brazilian Portuguese-to-Libras translation system that presents the translated content via an animated virtual human, or avatar. The paper describes the methodology adopted to compile a source language corpus having the deaf student needs in central focus. It also describes the construction of a parallel Brazilian Portuguese/Brazilian Sign Language (Libras) corpus based on motion capture technology. The envisioned translation architecture includes the definition of an

Intermediate Language to drive the signing avatar. The results of a preliminary assessment of signs intelligibility highlight the application potential.

Keywords Assistive technology for the deaf and hard of hearing · Deaf accessibility · Bilingual education · Inclusive education · Sign language · Brazilian sign language · Virtual human · Signing avatars

1 Introduction

According to the 2010 Brazilian population census [29], 5.1% of the population of Brazil—about 9.7 million citizens—have some kind of permanent hearing loss (0.3 million people declared being unable to hear; 1.8 million declared severe hearing difficulty even when using a hearing aid; and 7.6 million declared some hearing difficulty even when using a hearing aid). The 2010 Brazilian census also indicates that 30% of those 9.7 million are illiterate and there are no official statistics regarding the number of citizens that are proficient in Brazilian Sign Language (Libras).

According to the census methodology, an illiterate person is a person who cannot or has extreme difficulty in reading and writing in Portuguese. One reason for the high illiteracy rate among deaf people is the high rate of failure in school, which does not occur due to any biological characteristics of deafness, but because of the countless obstacles found in the educational process for deaf students in Brazilian schools. These schools still have a lack of skilled teachers working in bilingual education, instructional materials, or even interpreters whose role is to translate the contents of the school subjects into Libras [25, 58].

✉ José Mario De Martino
martino@dca.fee.unicamp.br

¹ School of Electrical and Computer Engineering, University of Campinas, Campinas, Brazil

² Center for Rehabilitation Research Studies, University of Campinas, Campinas, Brazil

³ Institute of Language Studies, University of Campinas, Campinas, Brazil

⁴ Federal University of Alagoas, Alagoas, Brazil

⁵ Center for Information Technology Renato Archer, Campinas, Brazil

Bilingual education for deaf people, which comprises Libras as their first language and the written mode of Portuguese as their second language, is a civil right established by decree 5.626 and supported by the National Education Plan [8]. However, this type of education has not yet been fully put into practice.

In their research, Silva et al. [58] concluded that most of the instructional materials that are produced and adopted in schools are designed for and developed toward students who can hear, and whose first language is Portuguese. This practice harms the educational process for deaf students. The authors also highlighted the importance of a multiliteracy approach for the education of deaf students. This approach can facilitate the participation of deaf students in classroom activities as well as their understanding of the school subjects, which helps them to better learn how the Portuguese language works. Silva and Kumada [57] argue that Brazilian schools need to change in order to better accommodate the deaf students.

Currently, tools available through information and communication technologies and the Internet support the reading, production, and distribution of text in which different kinds of semiosis work in the making of meaning. Rojo [53], for example, highlights how these new technologies allow for an intensive textual hybridization, which can include written, oral, and imagetic language, sometimes all at once, in gadgets and interfaces through hypertexts and hypermedia. These tools make it possible for students to try new and different literacy praxes.

In light of this new reality and with the goal of making schools more accessible to deaf students, some teachers have already adopted the use of Assistive Technologies. Among those technologies, the use of animated avatars, which translate words or sentences from Portuguese into Libras automatically, are being used as a strategy for easing content learning and the interaction among deaf and hearing students [16]. However, it is necessary to emphasize that the current avatar-based systems available in Brazil are still being developed toward working with general vocabulary, guided by ordinary/commonly used lexemes. Thus, teachers of specific subjects as, for example, mathematics, geography, and the sciences face not only a lack of educational material for deaf people but also a lack of field glossaries and/or technical dictionaries, which include specific terms in Libras for each school subject.

In line with new digital literacy trends [53], coupled with bilingual education (which considers and respects Libras as the deaf students first language and Portuguese as their second language), we emphasize the importance of technological appropriation as an ally for the development of accessible instructional materials for deaf and hearing students.

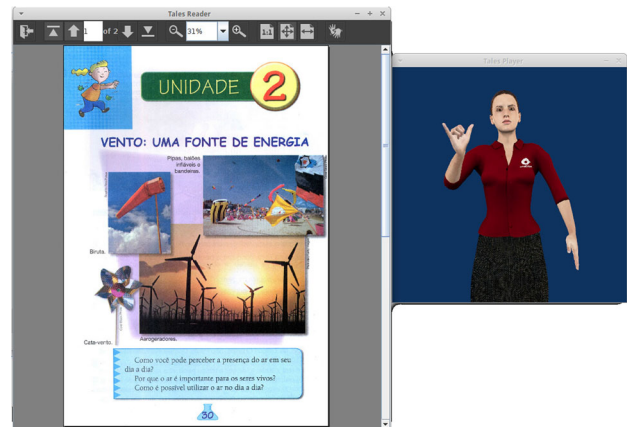


Fig. 1 Illustrative representation of the target application

The work described in this paper involves the scientific and technical collaboration between the Center for Rehabilitation Research Studies, the School of Electrical and Computer Engineering, and the Institute of Language Studies of the University of Campinas, as well as the Institute of Computing of the Federal University of Alagoas. The collaboration between these research groups with different domains of expertise, including Linguistics, Social Science, Pedagogy, and Computer Science, aims to develop an automatic Brazilian Portuguese-to-Libras translation system capable of presenting signed information by means of an animated virtual human, or avatar. The system is being designed especially for translating elementary school textbooks. Figure 1 illustrates the target application envisioned.

In the figure, the text marked by the user is presented by an animated 3D avatar in a separate window that can be moved and resized. The objective is to provide an attractive way for the deaf and hard-of-hearing students to experience written and signed content presented side-by-side on a computer screen. This would allow them to follow the written material as easily as their hearing classmates. It is anticipated that the use of this technology in bilingual educational environments will improve the motivation and performance of deaf students. With the proper pedagogical strategy, the technology can be applied to activities in the classroom and also at home, giving more flexibility and time for students to process and convert information into knowledge. In addition to the educational application, the automatic translation into Libras can also assist the deaf and hard-of-hearing in accessing the Internet and thus opening them up to a whole universe of written information.

This paper presents an overview of the translation system under development. The remainder of the paper is organized as follows: In Sect. 2, we survey prior efforts to developing an automatic translation system from written

text to sign language; Sect. 3 describes the approach followed, Sect. 4 the intelligibility evaluation and finally, Sect. 5 discusses the positive and negative implications of some project decisions already made and challenges yet to be faced.

2 Related work

The automatic translation of written text into sign language without human intervention is a machine translation (MT) problem. Since sign language is a visual language based on manual, facial, and body movements, the result of the translation is typically presented by a signing avatar. A signing avatar is an animated 3D model of a virtual human that presents messages in sign language.

Considering the different translation strategies found in the literature, text to sign language systems can be classified in:

- Rule-based systems, which rely on different levels of linguistic translation rules;
- Data-driven (or corpus-based) systems, which build computational knowledge from examples [64] or from statistical models derived from large bilingual parallel corpora [30];
- Hybrid systems, which combine aspects of the rule-based systems with the data-driven approach [51].

Early MT initiatives for sign language were based on sets of rules derived from a deep knowledge of the lexical, grammatical structure and contrastive characteristics of source and target languages.

Examples of systems that adopted the rule-based approach are ZARDOZ [70], TEAM [74], the Polish Text-into-Gesture Translator (TGT) to Polish sign language (PJM) [65], the South African Sign Language Machine Translation (SASL-ML) project [69], and the many variants developed under the ViSiCAST and eSIGN European projects [21]. More recently, Porta and colleagues proposed a rule-based MT system from Spanish to Spanish Sign Language (LSE) [50]. In the work, the authors highlight that the rule-based approach is advantageous for sign languages since, from a natural language processing point of view, they are still under-resourced and low-density languages in terms of corpora and lexicons. Moreover, the rule-based translation strategy can be considered domain independent.

However, from another point of view, sign languages also impose important challenges to the automation of translation by rules. First, the linguistic knowledge of a certain sign language may be limited, which prevents the definition of robust translation rules from source to target language. This is the case, for example, of sign languages

that just recently have been documented and those that are considered emerging sign languages [36, 44, 55, 66].

Second, sign languages comprehend visual communication mechanisms that are not modeled by simple syntactic or semantic analyses. This is the case, for example, of Classifier Predicates (CP). CPs are sign language mechanisms used to visually describe scenarios and sequences of events. For example, in the sentence, “the boy climbed the tree and fell,” the signer may use hand and arm configurations to represent a tree in the space and use the “boy” sign to simulate that he is climbing the tree and then falling. Without abandoning the rule-based paradigm, Huenerfauth [28] proposes the modeling of CPs through the association of lexicalized syntactic structures with 3D spatial directives for the signing avatar at the cost of limiting the rules domain.

The technological evolution that resulted in the increase in computational power, inexpensive storage, and greater access to motion capture equipment leveraged the building of parallel corpora for sign languages and the exploration of data-driven MT approaches. The data-driven approaches can be divided into two main strategies: Example-Based Machine Translation (EBMT) and Statistical Machine Translation (SMT).

In analogy to humans, EBMT systems construct translations of new unseen content in the source language by making use of previously seen translation examples, rather than by performing a deep linguistic analysis. For sign languages, a database of examples is typically constructed through the annotation of a video of a sign language interpreter. Glosses, sometimes accompanied by interrogative or negative markers, are frequently used to annotate the signs presented by the interpreter. The corresponding annotation of the video in the target language provides the missing information to build a database of examples. Examples of EBMT systems for sign language are the English into Sign Language of the Netherlands (NGT) [45], the Arabic into Arabic Sign Language (ArSL) described in [1], and the automatic translation system into JSL (Japanese Sign Language) [67].

In SMT systems, a translation model is typically trained over a large bilingual sentence-aligned corpus. The model is capable of determining translation probabilities based on metrics derived from the corpus, such as frequencies of word/gloss co-occurrences, relative source and target word/gloss positions and sentence lengths. Among the initiatives to implement SMT systems into sign language, we highlight the works from Bungeroth and Ney [11], based on a parallel corpus for German Sign Language (DGS), the Czech TTSL system, and the work from Othman and Jemni [48], which implemented an SMT from English written text to American Sign Language (ASL) gloss.

While EBMT and SMT approaches can be considered independent of a deep linguistic analysis of the target language, their translation performance is heavily dependent on the diversity of sentences and the size of the corpus. As pointed out by Porta et al. [50], for SMT systems in particular, when data are scarce or in sparse domains, induction of correct translation hypotheses is difficult or impossible. Among the strategies studied to improve the translation result in situations where the corpus is not sufficiently large, Massó and Badia [41] suggest the use of morpho-syntactic information from the sentences result. Recent initiatives to build large parallel corpora are also reported in the literature [23, 62].

Hybrid approaches arise as an alternative adopted by systems that aim to take advantage of the data-driven approach without losing too much translation quality. This is the case for the present work, which adopts Falibras, a TTSL system from Brazilian Portuguese into Libras [10]. Since its first release in 2002 [14], the system has undergone many cycles of development and improvement. Currently, Falibras is characterized as a hybrid approach that combines rule-based syntactic transfer and data-driven approaches [15] using knowledge derived from a translation memory [9] and a database of previously translated terms and segments used as a reference for similar translation cases [61]. Additionally, to disambiguate terms that can be translated in different ways, Falibras also implements a statistical model that takes into account the context and frequency of the occurrence of the terms from previous translations. Details of the implementation are provided in Sect. 3.

Another example of a hybrid system is described by San Segundo et al. [54]. The system differs from Falibras because it includes a speech recognition module capable of translating Spanish speech to Spanish sign language (LSE). Our approach is focused on text-to-sign language translation.

Finally, it is also important to observe that MT systems for sign language are typically designed to resolve accessibility problems for deaf people regarding real services or relevant information, such as postal office services [17], driver's license renewal service [54], access to disaster messages [40], weather forecast information [71], and public transport passenger information [56]. Such systems focus on the modeling of the signs and sentences that frequently occur in the contexts in which they are specialized.

Similarly, the present work focuses on the deployment of an application for the interactive translation into Libras of elementary school textbooks. In the present version, the corpus is built based on the contents of a science textbook for young children (see Sect. 3.1).

3 Our approach

Figure 2 presents an overview of the Brazilian Portuguese to Libras TTSL system under development. In the figure, the rectangles with rounded corners represent processes, and the cylinders represent main databases. As depicted in the figure, the system is composed of two main modules, namely the Translation Module and the Animation Module.

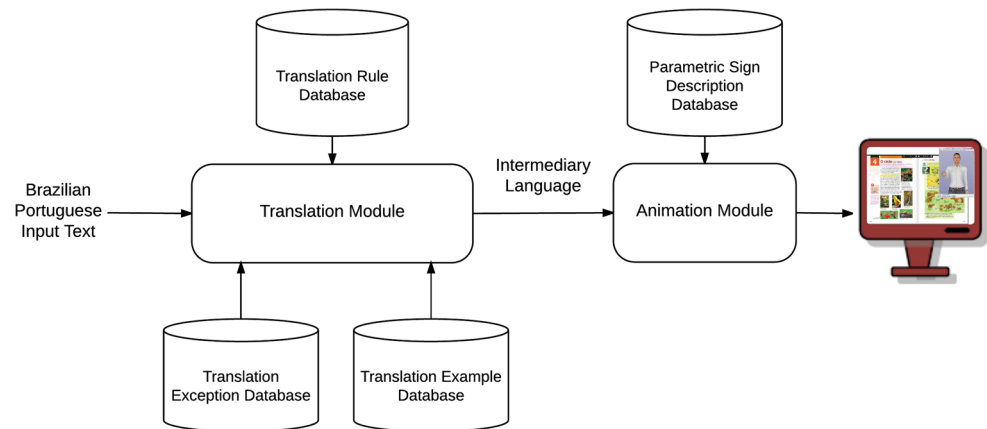
The system receives as input text in Brazilian Portuguese to be presented by the signing avatar. Based on a set of rules and the knowledge obtained from a translation memory, the Translation Module analyzes the input text and converts it into an intermediary representation—the Intermediary Language. The Intermediary Language is the input for the Animation Module which maps symbols of the Intermediary Language to signs in Libras.

The signs in the Animation Module are stored in a database using a parametric description (see Fig. 2). The parametric description includes information like the configuration and orientation of the hands, place of articulation, and the type of the movement that defines a template sign. During synthesis, each template sign can be modified to handle word inflections informed by the Intermediary Language. A word inflection can result in changes in sign parameters associated, for example, with intensity, speed, repetition, and facial expression. The signing avatar is a three-dimensional (3D) geometric model that is driven by sign descriptions and an algorithm to concatenate the sequence of signs.

The following sections describe in detail our approach for the implementation of the proposed architecture. Section 3.1 describes the approach for building a parallel Brazilian Portuguese/Libras corpus which is used to feed the databases depicted in Fig. 2. Section 3.2 details the implementation of the Translation Module. The Intermediary Language is described in Sect. 3.3 and the Animation Module is described in Sect. 3.4.

3.1 Parallel Brazilian Portuguese/libras corpus

Our approach for implementing a Brazilian Portuguese to Libras translation system is anchored in the analyses of a parallel Brazilian Portuguese/Libras corpus. The analyses of the corpus allow us to identify general rules of translation and exceptions to these rules, accumulate examples of translation, and model the production of signs. The translation rules are stored in the Rules Database and the exceptions in the Example of Exceptions Translation Memory (see Fig. 3). Accordingly, examples of translation are collected in the Examples of General Translation. The Parametric Sign Description Database contains descriptions of how to produce signs in the format expected by the Animation Module.

Fig. 2 System architecture**Fig. 3** Mocap session

The guidelines for building a parallel Brazilian Portuguese/Libras corpus are aligned with the primary goal of deploying an application for the interactive translation into Libras of elementary school textbooks, as illustrated in Fig. 1.

3.1.1 Methodology for compiling a Brazilian Portuguese corpus

Considering the target application (see Fig. 1), an important step in our methodology was to understand the realities faced by the deaf community, in particular deaf children in the first years of their school life, regarding the support material available for their special needs.

A brief survey in the Brazilian state of Sao Paulo involved a sample of ten schools, both public and private, with deaf individuals among their students and interpreters capable of intermediating the interaction between teachers and students. The survey showed that, among the ten schools analyzed, eight schools were unaware of any educational material for teaching deaf children and only two schools knew of a single source material. Of the two schools that had contact with any source material, only one makes use of a source material that is available on the

market today [20], which consists of a book designed for hearing students accompanied by a CD with the translation of its content into Libras. This scenario highlights the accessibility gap that exists between deaf students and educational textbooks. Based on the survey, we decided to build a parallel corpus from the content of a textbook. The task of choosing the textbook was based on two guidelines: the needs of the target group (deaf children), and the educational quality of the book.

Our understanding about the educational needs of deaf children was constructed from the experience of professionals who assist deaf children older than 7 years of age at the Center for Studies and Research in Rehabilitation at the University of Campinas. An important observation reported by these professionals regards the huge difficulties faced by deaf teenagers in understanding the course subjects they are supposed to learn during their later school years (young students aged 16–18). They observed that deaf teenagers hardly understand written Portuguese in the textbooks—not necessarily because they are unable to read or write, but because they have not acquired the background knowledge necessary for understanding the subjects presented. Therefore, it is important to support the reading and writing activities of texts starting in the early school years in order to allow deaf students to gradually obtain the basic knowledge required for comprehending more complex texts.

Based on this diagnosis, we decided to build our corpus by compiling the text from a textbook used during the early school years. Based on its practical merits for life and for the study of disciplines such as biology, physics, and geography, it was decided to select a science textbook designed for the third-grade students (aged 8–9).

In order to select the book to be analyzed, we first defined a set of candidate third-grade science textbooks from a list of books recommended by the Brazilian Ministry of Education and Culture (MEC). Second, we asked ten schools with deaf students which books they currently

use among those recommended by MEC. These books were given greater weight in the analysis. The set of candidate books was then further analyzed according to three criteria: the visual organization of the book, the progression of complexity in the use of the written language, and the translation difficulties associated with the written content.

We adopted the visual organization of the book as a selection criterion because visual cues are extremely important for deaf people due to the fact that they occur through a channel accessible to the deaf [59]. In this sense, the role of visual attention for deaf persons is equal to the importance of auditory attention for hearing persons. According to Siple, when the input information occurs through a visual channel, deaf persons have greater benefits in cognitive and language development. The importance of visual elements in pedagogy for the deaf is stressed by authors such as Reily [52], who state that visual literacy for the deaf is an especially important skill because the use of images favors the learning of concepts and the association of ideas. Researchers from the New London Group, who are concerned with multiliteracies, indicated that broadening the horizons of the school to provide access to digital technologies (PCs, phones, tablets, etc.) is necessary in order to allow students access not only to written texts, but also to multimodal texts [46].

Considering the complexity of the written language of the book, we initially thought that one of our criteria should be to analyze whether the vocabulary was introduced slowly, with explanations important for deaf children. Words like “earth,” “planet,” “world,” and “universe,” for instance, had to be clearly explained and differentiated. Hearing children listen to these words constantly in adults’ conversations or in the media, though we knew that for deaf teenagers, the differences between such words were not very clear, since they are all represented by the same sign in Libras. We also thought that syntax was a point to be observed. We wanted to avoid books with too many long sentences, especially those with long sentences at the beginning of the book. The analyses of the textbooks, however, showed that these criteria were not sufficient in that they were impossible to fulfill. It was not possible to find a book that dealt with the Brazilian Portuguese language in a way that introduced more complex structures after simple structures. At that point, we decided to reconsider the criteria, as such textbooks had been written for children who had Brazilian Portuguese as their mother tongue. Children that have Libras as their mother tongue were not the target group of the schoolbooks. In this context, it is also important to stress that deaf children undertake a more difficult task than others because they are learning Brazilian Portuguese at the same time as they are learning the contents of written textbooks. They have to read about things they do not know in a language that is not

their mother tongue and somehow translate that knowledge into Libras.

3.1.2 Textbook translation

Following the textbook selection, the next step in our methodology was to perform its translation to Libras. The translation process was performed through the close interaction of hearing native Brazilian Portuguese speakers proficient in Libras and deaf individuals having Libras as their first language and reading skills in Brazilian Portuguese. The textbook content was divided in more than 2000 sentences and, seeking to guarantee an acceptable level of standardization, the proposed translations were checked against well-known Libras dictionaries [12, 37]. A major challenge faced during the translation process was that even the most widely known Libras dictionaries in Brazil are still limited to everyday vocabulary, missing signs for many specific scientific and technical terms. In order to avoid fingerspelling, or dactylology, our strategy was to establish a network of deaf and hearing people, teachers, and technical professionals with mastery of Libras, to discuss signs found on field-specific Libras glossaries or sign compedia and, when necessary, to create and validate new scientific or technical signs. At present, the network is composed of 22 individuals and the textbook translation resulted in the creation of more than 160 signs to represent scientific terms.

Besides, the following aspects were also taken into consideration during the translation process:

- Respect for the most common syntactic structure in Libras, that is, the “topic-comment” structure instead of the structure commonly used in Brazilian Portuguese, which is characterized by the order “subject-verb-object.” It is important to emphasize that the structure “subject-verb-object” is also found in Libras [22], though our choice came from more recent studies of linguistic descriptions in Libras pointing that the “topic-comment” structure is currently more recurrent [49].
- The use of Libras classifiers to explain actions usually consists of several words in Brazilian Portuguese. Libras classifiers synthesize long sentences of Brazilian Portuguese through short sentences that incorporate the form and/or movement of the object or subject. In this way, this structure moves away from a more literal translation.
- The presence of polysemy and homonymy in Libras (for example, when we encounter the Brazilian Portuguese words “planet,” “Earth,” “world,” and “universe” whose signs are equally represented in Libras). With Sentences like “After Earth, how many planets

will pass until you reach your destination?” we observed difficulties in interpretation because “Earth” and “planet” have the same sign. In this particular case, we chose to translate “Planet Earth” through the sign for “planet” plus the manual spelling the word “Earth.”

The final result of the translation process was a set of videos of interpreters and deaf individuals signing the translation of the textbook sentences to Libras. In addition, each video was transcribed as a sequence of glosses, which corresponds to the sequence of signs recorded in each video.

3.1.3 Sign acquisition process

Considering the more than 9500 catalogued Libras signs in the largest compendium ever written on the subject [13], and taking into consideration the fact that Libras is a “living” language (constantly changing and evolving), the construction of a sign database is a major challenge to be faced.

The strategy for tackling this problem involves the use of motion capture (MoCap), which enables the capture of the movement of the body, limbs, head, and face in the tridimensional space. The use of MoCap technology to construct sign language corpora has been previously applied, for example, in the works of Gibet et al. [24], for the construction of a French Sign Language (FSL) corpus, and Lu and Huenerfauth [38], for the construction of an American Sign Language (ASL) corpus. MoCap technology enables the recording of natural and accurate description of signs and sentences performed by a signer. Such a realistic description of signs and signed sentences is applied to improve the realism of signing avatar animation. In addition, a relatively large amount of signs can be recorded in each MoCap session, facilitating corpus construction and further expansion.

The motion capture sessions were guided by the video recordings resulting from the textbook translation process (see 3.1.2). The MoCap equipment used is a passive optical technology that consists of 8 infrared cameras and reflective markers that are fixed to the Libras interpreter body. In our configuration, 41 markers were arranged on the upper body of the interpreter. During a recording session, the system is capable of capturing, simultaneously, the markers’ trajectory in space and the video of the signs presented by the interpreter. Figure 3 shows snapshots of a MoCap session. The MoCap setup also included two video cameras positioned one in front of the deaf native signer and the second capturing its side view. The material recorded by the cameras was used during the annotation process, as detailed in Sect. 3.1.4.



Fig. 4 Head-mounted MoCap system to capture facial expressions which have semantic meaning in Libras

Even though composed of eight cameras, the MoCap equipment may lose track of a marker when it does not correctly recognize the same marker in at least two frames captured by different cameras. Thus, after each MoCap session, the captured information has to be manually inspected to eliminate noise and correct tracking flaws. In addition, due to unavoidable joint occlusions of the fingers that occur during Libras signing, the MoCap equipment is not capable of satisfactorily tracking finger movements. Consequently, the joints of the fingers have to be corrected in the post-production process, which uses the video recorded during the MoCap session to retrieve the configurations of the hands. Facial expressions are also an important non-manual semantic component of sign language communication. In order to capture such expressions with accuracy, the head-mounted MoCap system Vicon Cara is being used. The system is equipped with four 1280×720 pixels, 60 fps, cameras (Fig. 4). The facial expressions captured are carefully combined with data from the upper body gestures during the MoCap post-processing phase.

In the last step of the MoCap process, the MoCap data are semi-automatically analyzed in order to generate the parametric sign descriptions that are stored in the associated database (see 3.4.1).

The present version of the built corpus comprehends more than 2000 sentences, approximately 8 h of raw material, corresponding to the complete content of a school science textbook.

3.1.4 Annotation process

The material recorded by the frontal- and side-view cameras, synchronized with the spatial marker coordinates provided by the MoCap system, was used to annotate and segment the temporal frontiers of each sign for each translated sequence. The annotation schema includes the following tiers:

- English translation of the sentence;
- Brazilian Portuguese translation of the sentence;
- the sequence of glosses (identification of signs in Libras);
- hand configuration identification for both hands;
- facial expression identification;
- the presence of narrative features;
- comments.

The annotation tool used was ELAN (EUDICO Linguistic Annotator) [72]. For each translated sentence, the following synchronized parallel data constitute the Brazilian Portuguese/Libras corpus: a frontal-view camera recording, a side-view camera recording, the signing gestures as MoCap C3D and BVH files and the corresponding ELAN Annotated Format (EAF) files. A more detailed description of the corpus construction is provided in [19].

3.2 Translation Module

The Translation Module processes the information based on the grammatical structures of the source and target languages. Figure 5 presents the activities of the translation process, showing when the databases presented in Fig. 2 are used. The module is composed of a text preprocessor, a grammatical analyzer, three complementary translators, and a sentence post-processor.

As depicted in Fig. 5, the module receives a Brazilian Portuguese text as input to be converted into the Intermediary Language. At the beginning of the conversion process, the input text is preprocessed in order to replace elements of the sentence and then prevent some semantic problems in

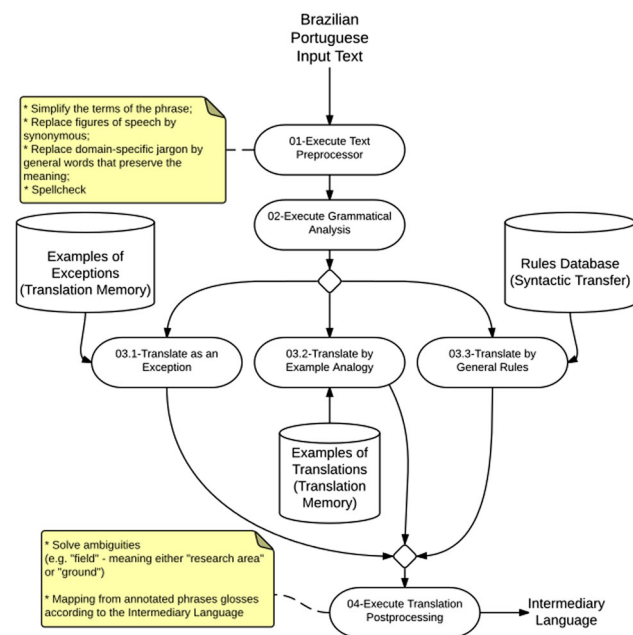


Fig. 5 Translation process

translation. Examples of expressions to be replaced are: figures of speech, domain-specific jargon, and wrongly spelled words. Following that, the module executes a grammatical analysis of the preprocessed text, in order to obtain morphological and syntactical annotations about each word. After that, the module uses the grammatical information to assess whether each sentence fits as an exception previously stored in the Examples of Exceptions Database. Exceptions are restrictive translation scenarios that cannot be accommodated in the Example Translation Database. Examples of exceptions are translations involving the construction of interpretation scenarios. When the input given is not an exception, the text is checked to see whether it corresponds to an example stored in the Examples of Translation Database. Such examples are indexed according to the morphological and syntactic structures of the sentences, in order to facilitate the identification of similar sentences to be translated. The examples in the Example of Translation Database are used to speed up the translation process and improve its quality by providing a translation rule specific for the grammatical structure of the sentence. If there is no match between the text and a translation example, the input is then processed according to standard translation rules stored in the Rules Database.

Finally, after executing one of the translation strategies, the translated sentences go through a post-processing step, in which the translated sentences are mapped to annotated glosses, following the structure of the Intermediary Language. For this however, some semantic ambiguities should also be resolved. For example, in the sentence “Esse um campo complexo da ciência” (This is a complex field of science), the word “field” should be translated to gloss “OCCUPATION-AREA” instead of “GROUND.” Such semantic analysis is done by using a Bayesian Network [18] for estimating the probability of each meaning based on the context provided by the words of the sentences, as well as the words of adjacent sentences. The outputs of the post-processor are sentences in the Intermediary Language (see Figs. 2 and 5), which will be detailed in 3.3.

3.3 Intermediary Language

The Intermediary Language uses glosses to represent signs accompanied by information associated with inflection. The Intermediary Language was devised to provide a straightforward mapping to signs and defines animation commands that are interpreted by the Animation Module in order to control the movements of the avatar.

The structure of the Intermediary Language is presented in Table 1, following the Bakus–Naur Format [43].

A translation (<translation>) is defined by a type (<type>) and a sentence (<sentence>). The type defines whether the sentence is an affirmation, a question,

Table 1 Structure of the translation in the intermediary language

```

<translation> ::= <type>, <sentence>
<type> ::= "affirmation" |
"question" | "exclamation"
<sentence> ::= <word>, <sentence> | <word>
<word> ::= <nonverb_inflection> <non_verb> | <verb_inflection> <verb>
<nonverb_inflection> ::= "" | <number_inf> | <grade_inf> | <grade_inf>
<number_inf> ::= "" | "plural"
<grade_inf> ::= "" | "augmentative" | "diminutive"
<non_verb> ::= <gloss>
<verb_inflection> ::= "" | <inf_spacial> | <inf_aspect> |
<inf_spacial> <inf_aspect>
<inf_spacial> ::= <direction> | <movement> | <position> | ...
<inf_aspect> ::= "fast" | "slow" | "fixed" | "intense" | ...
<verb> ::= <time> <gloss>
<time> ::= "past" | "present" | "future"
<gloss> ::= unique name of the sign or classifier
    
```

or an exclamation. The type information is needed when signing since it may interfere with facial expressions or even with the speed of the sign.

A sentence is composed of a list of words (<word>), separated by commas. A word can be defined by a non-verb inflection (<nonverb_inflection>) followed by a non-verb (<non_verb>) or by a verb inflection (<verb_inflection>) followed by a verb (<non_verb>). When the word is a verb (<verb>), the sign can also be changed according to modifiers found in the sentence (<verb_inflection>). Examples of such modifiers are fast, slow, fixed, and intense, which can interfere with the speed, the repetition of the sign, and the facial expression as well. An analogous organization is applied for non-verb words. Table 1 is a partial representation of the structure of the Intermediary Language and presents enough information to give a general idea of the language.

To illustrate the use of the Intermediary Language, the sentence “Vamos desenhlar as estações do ano?” (Shall we draw the seasons of the year?) is translated in Libras into three signs and represented in the Intermediary Language by the following string of words (see Fig. 6): question <type>, ESTACAO-DO-ANO <non-verb>, present <time> DESENHAR <verb>, present <time> ACEITAR <verb>. The Animation Module handles this text and animates the avatar to produce three signs in Libras: “ESTACAO-DO-ANO” (season of the year), followed by “DESENHAR” (to draw), and then “ACEITAR” (to accept). Moreover, since the sentence is a question, the avatar indicates it by emphasizing a doubt in its facial expression.

Another example can be seen in the sentence “No inverno chove muito” (It rains a lot in winter) which is represented in the Intermediary Language by (see Fig. 7): affirmation <type>, INVERNO <non_verb>, intense <inf_aspect> present <time> CHOVER <verb>. In this case, the Animation Module produces the animation

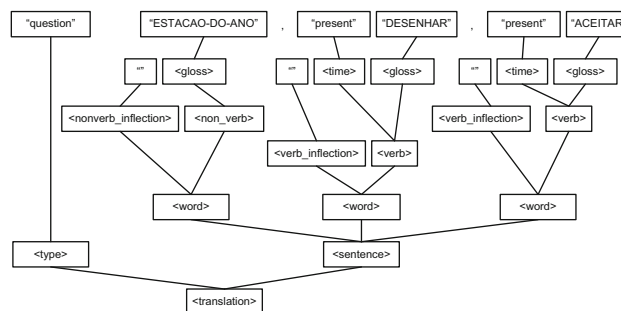


Fig. 6 Example “Vamos desenhlar as estações do ano?”

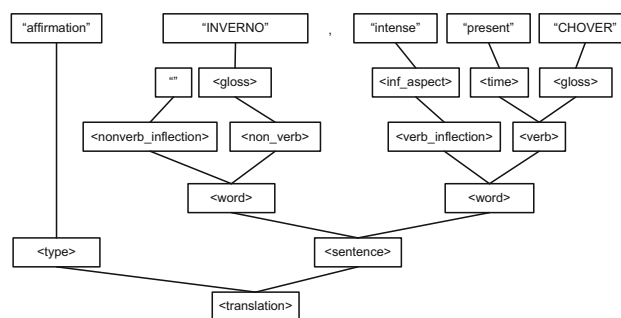


Fig. 7 Example “No inverno chove muito”

of two signs: “INVERNO” (winter), “CHOVER” (to rain). However, the sign for the verb “CHOVER” is modified to indicate the intensity of the “CHOVER” (intense <inf_aspect>).

In the Intermediary Language, each individual sign is represented by a unique name, or gloss.

3.4 Animation Module

The Animation Module is responsible for animating the avatar and presenting the resulting animation on a



Fig. 8 The three-dimensional model of the avatar (from left to right): as presented to the user; polygon mesh; and control skeleton

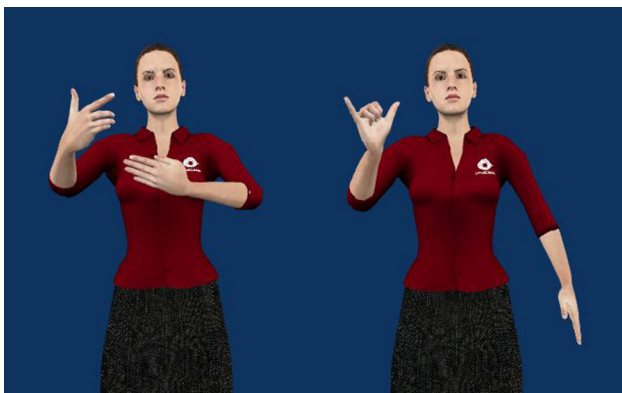


Fig. 9 Signing avatar

computer screen. Control of the movements that are performed by the avatar is based on the information provided by the Intermediary Language, which specifies a sequence of signs accompanied by inflection information, when appropriate. The inflection information is used to modulate parameters of the production of the sign, such as velocity, spatial amplitude of the movement, and number of repetition.

The visible part of the signing avatar comprises a three-dimensional polygonal mesh (Fig. 8, center) with associated skin, clothing, and hair textures to create a realistic representation of a human being (Fig. 8, left). Figure 9 shows the avatar as seen by the user.

The actual movement of the avatar is controlled by its skeleton (Fig. 8, right). The skeleton is a system of virtual bones and joints arranged hierarchically to simulate the existing joints of the human body. The movement of each bone influences the deformation of predefined regions of the polygonal mesh to create the illusion of natural body movement. As with the body animation, the facial animation of the virtual three-dimensional model is also controlled by a set of joints (Fig. 8, right). The information required for the handling of the bone-and-joint system is

retrieved from XML files containing the parametric description of the signs stored in the database of the Animation Module (see Fig. 2). Details of the format and structure of the XML sign description are discussed in 3.4.1.

3.4.1 Sign description

There is a wide range of objective and subjective elements that together generate meaning in a gestural sequence in sign language. Stokoe Jr. [63], in his pioneering work, enumerated the components of the signs, analyzing the morphological particles of the gestures correlated to the phonemes of the oral languages. This approach was further developed by Liddell and Johnson [35], incorporating components such as position, orientation and hand configuration, movement, trajectories, and locations and points of contact, as well as non-manual elements. Rhythm, cadence, and pauses also contribute to the effectiveness of message transmission in sign language.

Despite previous efforts, there is no consensus on the structure of sign languages [42]. Linguists and researchers have not yet agreed upon which information should be considered relevant and recorded in transcriptions. Unlike spoken language, which for hundreds of years has been represented by a quasi-phonological system—the alphabet—sign language still lacks a widely accepted writing system.

The proposed approach is based on the sign description format described in Amaral et al. [4] and Amaral and De Martino [3]. The format is based on XML language and describes the relevant characteristics of signs needed to synthesize the animation of the avatar. Essentially, the format considers that a sign or a sequence of signs can be described by a sequence of keyframes, or poses, and the transitions between them. The main features of the sign description format can be summarized as follows:

- The format structures the characteristics of the signs hierarchically;
- the format uses textual notation widely used for computational processing (XML-based), thus facilitating implementation;
- the format can represent sequentiality and concurrency in a sign;
- the format is capable of expressing the concept of symmetry, which leads to a more succinct sign description;
- the format describes the configuration of the hands, location points, and signing space, and other relevant features from a geometric point of view;
- the format also allows for the representation of non-manual, facial, and body expressions.

According to Johnson and Liddell [31–34], a sign can be described as a composition of two stages: In a first moment, we can observe hands, face, and body stationed in space, and then they move until they reach a new moment of stagnation. The format represents these two stages by the notation elements “pose” and “movement,” which respectively describe the configuration of the joints of the avatar skeleton in the start position followed by movement away from this initial state.

It is important to note that textual description of the subtle movements performed during signing is not trivial task. However, tests involving deaf people using a prototype implementation of the avatar have shown that our approach produces an intelligible sequence of signs. More details about the description of signs and the intelligibility test can be found in Amaral [2], Amaral et al. [4], and Amaral and De Martino [3].

3.4.2 Concatenating signs

As stated previously, the articulation of avatar joints is made by interpolating translations and rotations of bones in the control skeleton between two or more keyframes. If performed without further refinement, the resulting animation may appear overly artificial, which may hinder the understanding of the content. To prevent this type of problem, the joint interpolation will use mathematical models retrieved from MoCap data using techniques known as curve fitting and curve approximation.

The keyframes extracted from the description database (see Fig. 2) constitute atomic signs; to create a sentence, these have to be concatenated. A naive approach would be to interpolate a sign’s final keyframe values to the following sign’s first keyframe values. The computational cost of this approach is minimal, but the result is not visually appealing since it does not consider any physiological constraints and thus renders the movement mechanical and unnatural. We approach this problem by analyzing motion capture data from movements between signs in order to create a mathematical model that can emulate hand, arm, body, and facial movements more naturally.

The study of human arm movement is a vast one in neurophysiology. Much research has been conducted in order to understand how this movement is planned, as in Gottlieb et al. [26] and Soechting et al. [60]. This study is relevant not only in movement analysis, for instance in increasing a baseball’s velocity [27], but also in computer graphics in order to better predict the movement of a character’s arm [73]. While there is not yet a consensus on the existence of a universal model able to predict arm movement, some researchers found a minimum torque-change model that has, with some accuracy, predicted planar arm movement [68]. Others have shown

experimental results indicating a correlation of torques in arm joints to achieve a final grasp orientation [39], but still other studies have shown that, for wider or different range, of movements than those presented in previous works which has led to a hypothesis of a law of movement, such a law does not apply. Specifically in Soechting et al. [60] the experimental results did not conform to Donders’ Law, which predicts that the final position of the hand in space is achieved by a unique posture of the elbow and angles of the shoulder.

What is interesting about these studies is that each tend to deal with a very specific arm trajectory—a baseball throw [27]; a constrained movement of the arm starting at 18 predefined locations and ending at 5 predefined targets [60]; a reaching to fixed objects [39]—which may explain the different results. These experiments have also been conducted using a movement that starts and ends at rest; that is, the initial and final velocity is zero (in the avatar, movement usually does not end or start at zero velocity). Considering the specific needs of our application, we decided to construct a transition model derived from the analyses of the motion capture data from our parallel corpus.

Currently, we are studying the MoCap data, adjusting it to different mathematical models and evaluating the results. In the future, we also intend to explore machine learning algorithms to find and tune the models of transition based on initial and final arm postures.

4 Intelligibility evaluation

The final quality of the system presented in Fig. 2, including the accuracy of the translation and the animation aspects that influences the signing avatar intelligibility, is going to be highly dependent on the evolving design of Translation and Animation Module, through intermediate evaluation steps.

Considering this approach, we performed an isolated sign intelligibility test, or ISIT, as an early assessment to check the foundations of our animation strategy. The ISIT seeks to measure how intelligible isolated signs produced by our animation system are. The ISIT was conducted using a set of 18 signs from Brazilian Sign Language (see Table 2). Each sign was presented as a video clip of a real interpreter and animated by an avatar. The conducted experiment focused in evaluating the proper modeling of hand configurations and the graphical presentation of the avatar. For this reason, the sign animation was not driven by motion capture data, but instead it was implemented a key-pose animation approach. Both types of stimuli (video and animation) were presented as single signs alone, without being part of a sentence.

Table 2 Signs used in the isolated sign intelligibility test or ISIT

Sign	Short description
AMANHÃ (Tomorrow)	One-handed sign; articulated at the side of the forehead; middle finger starts by touching the side of the forehead, the hand is moved away from the forehead, and simultaneously the middle finger is moved downward
ANDAR (To Walk)	One-handed sign; articulated in front of the body; alternate movement of the index and middle fingers
APLAUDIR (To Applaud)	Two-handed sign; articulated near the side of the head; mirrored rotation of the wrists
ÁRVORE (Tree)	Two-handed sign; articulated in front of the body; rotation of the wrist of the dominant hand.
BARATA (Cockroach)	One-handed sign; articulated at the center of the forehead; alternate movement of the index and middle fingers
BICICLETA (Bicycle)	Two-handed sign; articulated in front of the body; alternate circular movement of the arms
BOM_ DIA (Good Morning)	One-handed sign; articulated in front of the mouth; the hand starts by touching the mouth and the finger tips touching each other, the hand is moved away from the mouth and the hand is opened; the hand assumes a pointing position and is moved in an arc from one side of the body to the other
BRASIL (Brazil)	One-handed sign; articulated in front of the body; downward oscillating movement
CACHORRO (Dog)	One-handed sign; articulated in front of the mouth; inward and outward movement of the hand
COMER (To Eat)	One-handed sign; articulated in front of the mouth; simultaneous downward movement of the index, middle, ring and little fingers
COMPUTADOR (Computer)	Two-handed sign; articulated in front of the body, mirrored circular movement of the hands
CORAÇÃO (Heart)	One-handed sign; articulated by touching the left side of the body; downward and upward movement of the opened hand
DEUS (God)	One-handed sign; articulated on the side of the body; upward movement of the arm
DIFÍCIL (Difficult)	One-handed sign; articulated on the forehead; movement from on side of the forehead to the other
FEIJÃO (Bean)	Two-handed sign; articulated in front of the body; rotation of the wrist of the dominant hand that is rrasping the index finger of the other hand
FUTEBOL (Soccer)	Two-handed sign; articulated in front of the body; alternated downward-upward movement of the hands
MOTOCICLETA (Motorcycle)	Two-handed sign; articulated in front of the body; alternated rotation of the wrists
SURDO (Deaf)	One-handed sign; articulated around the head; circular movement going from the ear to the mouth

The ISIT is in accordance with the guidelines of biomedical research involving human beings of the National Health Council of Brazil. It was approved by the Human Research Ethics Committee of the School of Medical Sciences of the University of Campinas.

Eighteen signs allowed us to keep the total duration of the test to around 20 min. We selected a set of commonly used and well-known signs. The selected signs cover a broad variety of hand configurations and movements from Brazilian Sign Language.

Eighteen video clips of a real interpreter and 18 animations of the avatar, corresponding to each one of the 18 signs were presented to 33 participants, aged between 12 and 35, 16 of whom were deaf and 17 of whom were not deaf; all were Libras signers.

The participants had no previous knowledge of the research and the goals of the test. The animations were always presented to the participant before the video of the real interpreter to avoid inference from the video about the signs presented by the avatar. The order of the presentation of the signs was randomly shuffled for each new participant. After evaluating the signs generated by the avatar, the user also evaluated the signs articulated by the real human interpreter.

A software tool was especially developed to present the animations and videos to the participants, and then accept and store the answers. The tests were conducted in a Core 2 Duo 1.6 GHz with 4 GB of RAM notebook computer with a GeForce 8600 M graphics processing unit with 256 MB of RAM memory. Each participant evaluated the material in an isolated room, without contact with other participants and without access to any additional aids such as dictionaries, books or Internet sites. Before starting the evaluation, a brief introduction explaining the test was presented to the participants by an instructor—orally (for the hearing individuals) and signed. The software tool was designed to present each animation/video only once. To start a new animation/video, the participant had to click a command button on screen using a mouse. After the presentation of an animation/video, the participant was asked to type in a text box the meaning of the sign, or leave it empty if he/she did not recognize the sign. After advancing to the next animation/video, it was not possible to return and change a previous answer. After collecting the input of the participants, the answers were analyzed to identify and group the answers with similar meaning as, for instance, “andar” and “caminhar” (to walk and to go on foot). Typographical errors and misspelled words were also corrected during this analysis.

Table 3 presents the number of correct answers for each sign and the related intelligibility scores. In the table, the columns under the heading “# Correct Answers” present the number of correct answers for the videos of the real interpreter (“Video” column) and for the animations (“Animation” column). The scores presented in the columns listed under “Intelligibility Scores” are calculated by dividing the number of correct answers by the number of participants. The last row of the table presents the total number of correct answers for all signs. The average intelligibility scores are calculated by dividing the total

Table 3 ISIT results

Sign	# Correct answers		Intelligibility scores	
	Video	Animation	Video	Animation
AMANHÃ	28	25	0.848	0.757
ANDAR	29	29	0.879	0.879
APLAUDIR	33	31	1.000	0.939
ÁRVORE	33	33	1.000	1.000
BARATA	32	30	0.970	0.909
BICICLETA	33	33	1.00	1.00
BOM_DIA	29	28	0.879	0.848
BRASIL	33	33	1.000	1.00
CACHORRO	33	32	1.000	0.970
COMER	33	33	1.000	1.000
COMPUTADOR	32	32	0.970	0.970
CORAÇÃO	27	27	0.818	0.818
DEUS	29	24	0.879	0.727
DIFÍCIL	33	33	1.00	1.00
FEIJÃO	31	31	0.939	0.93
FUTEBOL	24	30	0.727	0.909
MOTOCICLETA	33	33	1.000	1.000
SURDO	33	33	1.000	1.000
Total/Average	558	550	0.939	0.926

number of correct answers by the total number of all answers (33 participants \times 18 signs = 594).

The results show that, for the majority of the tested signs (12 out of 18), the intelligibility scores of the avatar are equal to or greater than the rates for the video of the real interpreter. The average intelligibility score for the videos is 0.939 and score of the animations is 0.92%. A non-parametric statistical Mann–Whitney U -test was applied to test the null hypothesis $\mu_v = \mu_a$, where μ_v is the intelligibility score for the video of the real interpreter and μ_a is the intelligibility score for the avatar. The Mann–Whitney U -test resulted in $Z = 0.52204$ and p value = 0.60164, meaning that the difference between μ_v and μ_a is not statistically significant. Considering the alternative hypothesis $\mu_v > \mu_a$, the Mann–Whitney U -test resulted in $Z = 0.52204$ and p -value = 0.69918, meaning that statistically it is not possible to state that the intelligibility score of the video is greater than that of the avatar.

Besides the 18 signs, the fingerspelling of two names by the avatar was also evaluated. We only tested two situations together with the 18 isolated signs to keep the duration of test to around 20 min. The results for the dactylology yielded 32 correct answers for the name UNICAMP and 30 for ALICIA, from a total of 33 answers. The results suggested that the avatar’s fingerspelling mechanism provides meaningful information.

5 Concluding remarks

Following the framework for action put forward by UNESCO's "Salamanca World Conference on Special Needs Education," Brazil instituted a model of inclusive education through a law which established guidelines and foundations for national education [6]. This law recommends that deaf students be offered opportunities to be included in the mainstream school system. Moreover, the Brazilian government recognized Brazilian Sign Language as a language of the communities of deaf people in Brazil [7] and established that the education of the deaf deserves a bilingual approach in which the teaching of deaf students through their first language, Libras, is ensured. Additionally, schools must teach deaf children the written form of Brazilian Portuguese as a second language [8].

However, despite favorable legislation, in practice the inclusion model left some issues open regarding the implementation of the bilingual approach in schools, resulting in many schools lacking teachers with bilingual training to deal with deaf students. Other aggravating factors which may be mentioned refer to a lack of professional Libras translators and interpreters to accompany deaf students in the classroom, as well as a lack of appropriate teaching materials for deaf education.

The present project brings us back to the traditional discussion about pedagogical questions among those who work with minority groups, the deaf among them; that is, bilingual teaching. We can hardly speak about bilingual education in a context in which both languages in contact have different values in society and one has much more space than the other. This leads to the conclusion that deaf children fail in school not because they are deaf, but because their language is different from the official language of the school.

In this context, the present work focuses on the translation of educational material aiming to improve the bilingual education experience for deaf children, facilitating the comprehension of written Portuguese and fostering the mastery of sign language. Among the contributions of the present work, we highlight first the implementation of a methodology to compile a source language corpus which not only focuses on their suitability for automatic translation but also includes relevant criteria for the education of deaf children. Second, the project promoted the construction of a comprehensive parallel Brazilian Portuguese/Libras corpus, with more than 2000 sentences, which is also contributing to advance the studies on Libras, including the systematic description of its grammar. Third, we presented a text-to-sign language translation architecture based on an Intermediary Language to drive a signing avatar. Finally,

we have shown that the proposed animation methodology is capable of conveying intelligible signing.

The components of the proposed architecture and the integration between Translation and Animation Modules are in the prototyping and evaluation phases. The Intermediary Language is being evolved to include the indication of non-manuals, such as facial expressions, which convey important information for adequate signing intelligibility. In addition, we plan to evaluate the quality of the translation component considering two complementary perspectives: software engineering and language effectiveness. From the point of view of software engineering, two quality attributes will be considered [47]: performance and scalability, i.e., the system's ability to meet efficiently to an increasing demand of simultaneous accesses. From a linguistic point of view, the translation component will be evaluated using the Corpus Linguistics approach [5], which provides a means to evaluate the translation efficiency by using translation examples (corpus) as benchmarks. Finally, we are also working on the definition of a comprehensive evaluation protocol in order to assess the system from the perspective of deaf students.

Acknowledgements This research is being supported by CNPq/MCTI-SECIS, Grant Number 458691/2013-5, and by Capes/SDH/MCTI, Grant Number 88887.091672/2014-0.

References

1. Almohimeed, A., Wald, M., Damper, R.I.: Arabic text to arabic sign language translation system for the deaf and hearing-impaired community. In: Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies, Association for Computational Linguistics, Stroudsburg, PA, USA, SLPAT'11, pp. 101–109 (2011)
2. Amaral, W.M.: Sistema de transcrição da língua brasileira de sinais voltado à produção de conteúdo sinalizado por avatares 3D. PhD thesis, Faculdade de Engenharia Elétrica e de Computação—Universidade Estadual de Campinas (2012)
3. Amaral, W.M., De Martino, J.M.: Towards a transcription system of sign language for 3D virtual agents. In: Tarek, Sobh (ed.) Innovations in Computing Sciences and Software Engineering, pp. 85–90. Springer, Netherlands (2010)
4. Amaral, W.M., Angare, L., De Martino, J.M.: Sign Language 3D Virtual Agent. In: Proceedings of the 5th International Multi-Conference on Society, Cybernetics and Informatics (IMSCI 2011), International Institute of Informatics and Systemics, Orlando, Florida, vol. I, pp. 93–97 (2011)
5. Baker, M.: Corpus linguistics and translation studies: implications and applications. In: Baker, M., Francis, G., Tognini-Bonelli, E. (eds.) Text and Technology: In Honour of John Sinclair, pp. 233–248. John Benjamins, Amsterdam (1993)
6. Brasil: Lei no. 9.394. In: Diário Oficial União, Imprensa Oficial República Federativa do Brasil, Brasília, DF, pp. 1–9 (1996)
7. Brasil: Lei no. 10.436. In: Diário Oficial União, Imprensa Oficial República Federativa do Brasil, Brasília, DF, p. 23 (2002)

8. Brasil: Decreto no. 5.626. In: Diário Oficial União, Imprensa Oficial República Federativa do Brasil, Brasília, DF, pp. 28–30 (2005)
9. Breda, W.L., Caon, D.R.S., Brito, P.H.S., Tavares, O.L., Menezes, C., Coradine, L.C., Albuquerque, F.C.: SOTAC: a software for knowledge-based automatic translation. In: Proceedings of the 9th IFIP World Conference on Computers in Education, Springer, Bento Gonçalves, RS, Brazil, vol. 1, pp. 1–10 (2009)
10. Brito, P.H.S., Franco, N., Coradine, L.C.: Falibras: Uma Ferramenta Flexível para Promover Acessibilidade de Pessoas Surdas. In: Anais del XVII Congreso Internacional de Informática Educativa (TISE 2012), Universidad de Chile, Santiago, Chile, pp. 87–96 (2012)
11. Bungeroth, J., Ney, H.: Statistical sign language translation. In: Workshop on Representation and Processing of Sign Languages, LREC vol. 4, 105–108 (2004)
12. Capovilla, F.C., Raphael, W.D.: Dicionário Enciclopédico Ilustrado Trilíngue da Língua de Sinais Brasileira *Libras*, 3rd edn. EDUSP, São Paulo, SP, Brazil (2008)
13. Capovilla, F.C., Raphael, W.D., Mauricio, A.C.: Novo Deit-libras: Dicionário Enciclopédico Ilustrado Trilingue da Língua de Sinais Brasileira *Libras*. EDUSP, São Paulo, SP (2009)
14. Coradine, L.C., Albuquerque, F.C., Brito, P.H.S., Silva, R.L., Silva, T.F.L.: Sistema Falibras: interpretação animada, em Libras, de palavras e expressões em português. In: Anais do III Congresso Ibero-Americano de Informática na Educação Especial (CIIEE 2002), Fortaleza/CE, Brazil, vol. 1 (2002)
15. Coradine, L.C., Tavares, O.L., Albuquerque, F.C., Ribeiro, M.N., Breda, W.L., Calado, I.A.A.R., Smarsaro, A.D.: Sistema Falibras: Um intérprete virtual como ferramenta de apoio pedagógico à educação de surdos. In: Anais do VII Congresso Ibero-Americano de Informática Educativa Especial (CIIEE 2007), Mar del Plata-Argentina, vol. 1, pp. 1–10 (2007)
16. Côrrea, Y., Vieira, M.C., Santarosa, L.M.C., Biasuz, M.C.V.: Aplicativos de tradução automática como facilitadores da comunicação entre surdos e ouvintes. In: Santarosa, L.M.C., Conforto, D., Vieira, M.C. (eds.) Tecnologia e Acessibilidade: Passos em direção à inclusão escolar e sociodigital. Evangraf, Porto Alegre (2014)
17. Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M., Abbott, S.: Tessa, a system to aid communication with deaf people. In: Proceedings of the Fifth International ACM Conference on Assistive Technologies, ACM, pp. 205–212 (2002)
18. Darwiche, A.: Bayesian networks. *Commun. ACM* **53**(12), 80–90 (2010)
19. De Martino, J.M., Costa, P.D.P., Benetti, A., Rosa, L.A., Kumada, K.M.O., Silva, I.R.: Building a brazilian portuguese-brazilian sign language parallel corpus using motion capture data. In: Proceedings of Workshop of Corpora and Tools for Processing Corpora, pp. 56–63 (2016)
20. Editora Moderna (Org): Projeto Pitanguá, 1st edn. Moderna, São Paulo (2005)
21. Elliott, R., Glauert, J.R., Kennaway, J., Marshall, I., Safar, E.: Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. *Univ. Access Inf. Soc.* **6**(4), 375–391 (2008)
22. Felipe, T.A.: Introdução à Gramática de Libras. Ministério da Educação e do Desporto. Secretaria de Educação Especial. Educação especial: Deficiência Auditiva, Brasília, DF (1997)
23. Forster, J., Schmidt, C., Hoyoux, T., Koller, O., Zelle, U., Piater, J.H., Ney, H.: RWTH-PHOENIX-weather: a large vocabulary sign language recognition and translation corpus. In: LREC, pp. 3785–3789 (2012)
24. Gibet, S., Courty, N., Duarte, K., Naour, T.L.: The SignCom system for data-driven animation of interactive virtual signers: methodology and evaluation. *ACM Trans. Interact. Intell. Syst.* **1**(1), 6:1–6:23 (2011)
25. Goês, M.C.R.d., Tartuci, D.: Alunos surdos e experiências de letramento. In: Farnades E, Dorzial A, Lodi ACB (eds.) Letramento, Bilinguismo e Educação de Surdos, 2nd edn. Mediação Editora, Porto Alegre, pp. 289–302 (2015)
26. Gottlieb, G.L., Song, Q., Almeida, G.L., Hong, D.A., Corcos, D.: Directional control of planar human arm movement. *J. Neurophysiol* **78**, 2985–2998 (1997)
27. Hirashima, M., Kudo, K., Watarai, K., Ohtsuki, T.: Control of 3D Limb Dynamics in unconstrained overarm throws of different speeds performed by skilled baseball players. *J. Neurophysiol.* **97**(1), 680–691 (2007)
28. Huenerfauth, M.: A multi-path architecture for machine translation of English text into American sign language animation. In: Proceedings of the Student Research Workshop at HLT-NAACL 2004, Association for Computational Linguistics, Stroudsburg, PA, USA, HLT-SRW '04, pp. 25–30 (2004)
29. IBGE: Censo Demográfico 2010—Características gerais da população, religião e pessoas com deficiência. Instituto Brasileiro de Geografia e Estatística, Ministério do Planejamento, Orçamento e Gestão e Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro, Brasil (2010)
30. Imamura, K., Okuma, H., Watanabe, T., Sumita, E.: Example-based machine translation based on syntactic transfer with statistical models. In: Proceedings of COLING, pp. 99–105 (2004)
31. Johnson, R.E., Liddell, S.K.: Toward a phonetic representation of hand configuration: the fingers. *Sign Lang. Stud.* **12**(1), 5–45 (2011)
32. Johnson, R.E., Liddell, S.K.: A segmental framework for representing signs phonetically. *Sign Lang. Stud.* **11**(3), 408–463 (2011)
33. Johnson, R.E., Liddell, S.K.: Toward a phonetic representation of signs: sequentiality and contrast. *Sign Lang. Stud.* **11**(2), 241–274 (2011)
34. Johnson, R.E., Liddell, S.K.: Toward a phonetic representation of hand configuration: the thumb. *Sign Lang. Stud.* **12**(2), 316–333 (2012)
35. Liddell, S.K., Johnson, R.E.: American sign language: the phonological base. *Sign Lang. Stud.* **64**, 195–278 (1989)
36. Linde-Usiekiewicz, J., Czajkowska-Kisil, M., Łacheta, J., Rutkowski, P.: A corpus-based dictionary of polish sign language (PJM). In: Proceedings of the XVI EURALEX International Congress: The User in Focus (2014)
37. Lira, G.A., de Souza, T.A.F.: Dicionário da língua brasileira de sinais. <http://www.acessobrasil.org.br/libras/> (2008). Accessed August 2016 2008
38. Lu, P., Huenerfauth, M.: Collecting a motion-capture corpus of American sign language for data-driven generation research. In: Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, Association for Computational Linguistics, Stroudsburg, PA, USA, SLPAT '10, pp. 89–97 (2010)
39. Marotta, J.J., Medendorp, W.P., Crawford, J.D.: Kinematic rules for upper and lower arm contributions to grasp orientation. *J. Neurophysiol.* **90**(6), 3816–3827 (2003)
40. Martin, P.J.M., Belhe, S., Mudliar, S., Kulkarni, M., Sahasrabudhe, S.: An Indian Sign Language (ISL) corpus of the domain disaster message using Avatar. In: Proceedings of the Third International Symposium in Sign Language Translations and Technology (SLTAT-2013), pp. 1–4 (2013)
41. Massó, G., Badia, T.: Dealing with sign language morphemes in statistical machine translation. In: 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, Valletta, Malta, pp. 154–157 (2010)

42. McCleary, L., Viotti, E.: Transcrição de dados de uma língua sinalizada: um estudo piloto da transcrição de narrativas na língua de sinais brasileira (LSB). In: Salles, H. (ed.) *Bilinguismo e surdez: Questões linguísticas e educacionais*, pp. 73–96. Cânone Editorial, Goiania (2007)
43. McCracken, D.D., Reilly, E.D.: Backus-Naur form (BNF). In: Ralston, A., Reilly, E.D., Hemmendinger, D. (eds.) *Encyclopedia of Computer Science*, 4th edn, pp. 129–131. Wiley, Chichester (2003)
44. Meir, I., Sandler, W., Padden, C., Aronoff, M.: *Emerging Sign Languages*, vol. 2. Oxford University Press, Oxford, pp. 267–280 (**Chapter 18**) (2010)
45. Morrissey, S., Way, A.: An example-based approach to translating sign language. In: *Second Workshop on Example-Based Machine Translation*, Dublin City University, vol. 16, pp. 109–116 (2005)
46. New London Group: A pedagogy of multiliteracies: designing social futures. In: Cope, B., Kalantzis, M. (eds.) *Multiliteracies: Literacy Learning and the Design of Social Futures*, pp. 9–38. Macmillan, South Yarra (2000)
47. Offutt, J.: Quality attributes of web software applications. *IEEE Softw.* **19**(2), 25–32 (2002)
48. Othman, A., Jemni, M.: Statistical sign language machine translation: from English written text to American sign language gloss. *Int. J. Comput. Sci. Issues* **8**(3), 65–73 (2011)
49. Pereira, M.C.d.C., Choi, D., Vieira, M.I., Gaspar, P., Nakasato, R.: *Libras: conhecimento além dos sinais*, 1st edn. Pearson Prentice Hall, São Paulo (2011)
50. Porta, J., López-Colino, F., Tejedor, J., Colás, J.: A rule-based translation from written Spanish to Spanish Sign Language glosses. *Comput. Speech Lang.* **28**(3), 788–811 (2014)
51. Probst, K., Levin, L., Peterson, E., Lavie, A., Carbonell, J.: MT for minority languages using elicitation-based learning of syntactic transfer rules. *Mach. Transl.* **17**(4), 245–270 (2002)
52. Reily, L.H.: As imagens: o lúdico e o absurdo no ensino de arte para pré-escolares surdos. In: Silva, I.R., Kauchakje, S., Gesueli, Z.M. (eds.) *Cidadania, surdez e linguagem: desafios e realidades*, pp. 161–192. Plexus Editora, São Paulo (2003)
53. Rojo, R.: Gêneros discursivos do Círculo de Bakhtin e multiletramentos. In: Rojo, R.H.R. (ed) *Escol@ conectada: os multiletramentos e as TICs*, Parábola Editorial, São Paulo, pp. 13–36 (2013)
54. San-Segundo, R., Montero, J.M., Córdoba, R., Sama, V., Fernández, F., D'Haro, L.F., López-Ludeña, V., Sánchez, D., García, A.: Design, development and field evaluation of a Spanish into sign language translation system. *Pattern Anal. Appl.* **15**(2), 203–224 (2012)
55. Schermer, T., Koolhof, C.: The first national Dutch Sign Language (NGT) Dictionary in book form: Van Dale Basiswoordenboek Nederlandse Gebarentaal. In: Dykstra, A., Schoonheim, T. (eds.) *Proceedings of the 14th EURALEX International Congress*, pp. 1555–1564. Fryske Akademy, Leeuwarden/Ljouwert, The Netherlands (2010)
56. Schnieder, L., Tschare, G.: Passenger information using a sign language avatar–individual travel assistance for passengers with special needs in public transport. *Int. Transp.* **67**(1), 8–10 (2015)
57. Silva, I.R., Kumada, K.M.O.: Representações sobre o contexto multilíngue da surdez. *Interdisciplinar–Revista de Estudos em Língua e Literatura* **19**, 99–114 (2013)
58. Silva, I.R., Nogueira, A.S., Hildebrand, H.R., Kumada, K.M.O.: O uso de jogos eletrônicos no processo ensino-aprendizagem de surdos. In: Valle, L.E.L.R., Matos, M.J.V.M., Costa, J.W. (eds.) *Educação digital: a tecnologia a favor da inclusão*, pp. 213–238. Penso, Porto Alegre (2013)
59. Siple, P.: Attentional resources and working memory: a new framework for the study of the impact of deafness on cognition. In: Emmorey, K., Lane, H. (eds.) *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, pp. 115–134. Lawrence Erlbaum Associates, London (2000)
60. Soechting, J.F., Buneo, C.A., Herrmann, U., Flanders, M.: Moving effortlessly in three dimensions: does Donders' law apply to arm movement? *J. Neurosci.* **15**, 6271–6280 (1995)
61. Somers, H., Diaz, G.F.: Translation memory vs. example-based MT: what is the difference? *Int. J. Trans.* **16**, 5–33 (2004)
62. Stein, D., Schmidt, C., Ney, H.: Analysis, preparation, and optimization of statistical sign language machine translation. *Mach. Transl.* **26**(4), 325–357 (2012)
63. Stokoe Jr, W.C.: Sign language structure: an outline of the visual communication system of the American deaf. *Studies in Linguistics: Occasional Papers* **8**, pp. 1–78 (**Revised 1978**) (1960)
64. Sumita, E., Iida, H., Kohyama, H.: Translating with examples: a new approach to machine translation. In: *Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language*, vol. 3, 203–212 (1990)
65. Suszczańska, N., Szmal, P., Francik, J.: Translating polish texts into sign language in the TGT system. In: *20th IASTED International Multi-Conference Applied Informatics AI*, Innsbruck, Austria, pp. 282–287 (2002)
66. Sze, F., Woodward, J., Tang, G., Lee, J., Cheng, K., Mak, J.: Sign language documentation in the Asia-Pacific region: a deaf-centered approach. In: *Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, Language Resources and Evaluation Conference (LREC) 2012*, pp. 155–158 (2012)
67. Umeda, S., Azuma, M., Uchida, T., Miyazaki, T., Kato, N., Inoue, S., Hiruma, N.: Automatic translation system to JSL (Japanese Sign Language) about weather information. In: *Proceedings of the Sign Language Translation and Avatar Technology (SLTAT 2015)* (2015)
68. Uno, Y., Kawato, M., Suzuki, R.: Formation and control of optimal trajectory in human multijoint arm movement. *Biol. Cybern.* **61**, 89–101 (1989)
69. Van Zijl, L., Baker, D.: South African sign language machine translation system. In: *Proceedings of the 2nd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, Association for Computing Machinery-ACM, pp. 49–52 (2003)
70. Veale, T., Conway, A.: Cross modal comprehension in ZARDOZ an English to sign-language translation system. In: *Proceedings of the Seventh International Workshop on Natural Language Generation*, Association for Computational Linguistics, pp. 249–252 (1994)
71. Verlinden, M., Tijsseling, C., Frowein, H.: A signing Avatar on the WWW. In: Wachsmuth, I., Sowa, T. (eds.) *Gesture and Sign Language in Human-Computer Interaction. Lecture Notes in Computer Science*, vol. 2298, pp. 169–172. Springer, Berlin (2002)
72. Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., Sloetjes, H.: Elan: a professional framework for multimodality research. In: *Proceedings of LREC, Fifth International Conference on Language Resources and Evaluation*, pp. 1556–1559 (2006)
73. Xu, H.Y., Li, D., Tang, S.: Kinematic inbetweening along implicit curves for motion animations. In: *Proceedings of the 16th International Conference on Artificial Reality and Telexistence-Workshops*, IEEE Computer Society, Washington, DC, USA, ICAT '06, pp. 1–4 (2006)
74. Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N.I., Palmer, M.: A machine translation system from English to American sign language. In: *Proceedings of the 4th Conference of the Association for Machine Translation in the Americas on Envisioning Machine Translation in the Information Future*, Springer-Verlag, London, UK, AMTA '00, pp. 54–67 (2000)