


ARTICLE

A machine translation mechanism of Brazilian Portuguese to Libras with syntactic-semantic adequacy

Manuella Aschoff C. B. Lima, Tiago Maritan U. de Araújo, Rostand E. O. Costa*
and Erickson S. Oliveira

Digital Video Applications Lab (LAVID), Computer Science Institute (CI), Federal University of Paraíba (UFPB), João Pessoa, Brazil

*Corresponding author. E-mail: rostand@lavid.ufpb.br

(Received 29 August 2018; revised 9 December 2020; accepted 14 December 2020; first published online 1 February 2021)

Abstract

Deaf people communicate naturally using visual-spatial languages, called sign languages (SL). Although SLs are recognized as a language in many countries, the problems faced by Deaf people for accessing information remain. As a result, they have difficulties exercising their citizenship and access information in SLs, which usually leads to linguistic and knowledge acquisition delays. Some scientific works have been developed to address these problems related to the machine translation of spoken languages to sign languages. However, the existing machine translation platforms have some limitations, especially in syntactic and lexical nature. Thus, this work aims to develop a mechanism for machine translation to Libras, the Brazilian Sign Language, with syntactic-semantic adequacy. It consists of an automatic translation component for Libras based on syntactic-semantic translation rules and a formal syntactic-semantic rule description language. As proof of concept of the proposed approach, we created a specific grammar for Libras translation exploring these aspects and integrating these elements into VLibras Suite, a service for machine translation of digital content in Brazilian Portuguese (BP) to Libras. We performed several tests using this modified version of VLibras to measure the level of comprehension of the output generated by the new translator mechanism. In the computational experiments, as well as in the actual tests with Deaf and hearing users, the proposed approach was able to improve the results of the current VLibras version.

Keywords: Machine Translation; Multimedia Platforms; Formal Rule Description Language; Sign Language

1. Introduction

Deaf people communicate naturally using visual-spatial languages, called sign languages (SLs). The SLs are visual-gestural languages, that is, we perceive them by the eyes and are performed in space with visual articulators: the hands, the body, the movements, and the signaling area. Since SLs are not the natural language of most of the hearing population, and not even a second language, the problems faced by the Deaf in accessing information remain. Consequently, it is a great difficulty for the Deaf to exercise their citizenship and social inclusion through SL.

Several studies have been developed related to the machine translation of spoken languages^a to sign languages^b to minimize the marginalization of the Deaf, such as Araújo *et al.* (2013, 2014), Huenerfauth (2008a, 2004, 2005b), Huenerfauth *et al.* (2008b), Lopez-Ludena *et al.* (2014a, 2014b), and Morrissey and Way (2014).

^aIn this work, when we refer to spoken languages, it means the spoken languages' written modality, not the spoken modality.

^bIn this work, when we refer to sign languages, it means the visual-spatial sign language modality, not writing SL.

Although the Deaf community prefers the translation performed by human interpreters, some studies report that there is an emerging perception of the validity of the use of machine translation through avatars, especially when human interpreters are not available (Huenerfauth 2005a; San-Segundo *et al.* 2008; Boulares and Jemni 2013 and Araújo *et al.* 2013). We reiterated this acceptance by the high number of downloads of this type of tool. In Brazil, for example, the VLibras,^c Prodeaf,^d Rybena^e, and HandTalk^f machine translations together have more than one million downloads.

However, existing solutions have some limitations. Still, they cannot guarantee that the content available to the Deaf will have the same quality as the content provided to the hearing users. These limitations are related to the difficulty in adapting the content to the sign language, respecting the SL's characteristics and grammar, and the low naturalness of the avatars (virtual animated agents) used to represent the output of the machine translation among others (Huenerfauth 2005b).

Another problem is related to the difficulty in dealing with syntactic, lexical, content, and context issues (Costa 2005). Besides, we can emphasize the existence of restrictions imposed by the textual organization in codes with different modalities (spoken-visual). Thus, in the case of translating an oral-auditory language into visual-gestural, they should not maintain the same grammatical form presented in the source language (Quadros 1997).

This work aims to contribute to the development of a mechanism for machine translation of Brazilian Portuguese (BP) for Libras that supports syntactic-semantic adequacy. The proposed approach consists of a “text-to-gloss”^g machine-translation rule-based component and a language modeling for the description of morphosyntactic-semantic rules.

We also developed a prototype of the mechanism along with the language modeling of a specific grammar for machine translation of BP to Libras. Afterward, we incorporated them to the VLibras Suite,^h which we use as the basis for the development of the proposed solution.

The main contributions of this work are as follows:

- A new text-to-gloss machine translation mechanism of Brazilian Portuguese (BP) for Libras that supports syntactic-semantic adequacy;
- The modeling of a formal language to describe the translation rules from Brazilian Portuguese to Libras, making the translation process flexible, configurable, and able to deal with other linguistic aspects;
- A simpler Libras gloss representation to treat the primary sign language specificities described in Section 3, including “special tokens” to incorporate expressions and the specific aspects of sign languages, such as directional verb agreement, incorporation of facial expressions during the signing, and intensification of signs;
- Based on the proposal Libras gloss's representation, our approach makes the translation process from spoken languages into sign languages similar to a translation process between two spoken languages, since it converts to a sequence of tokens in a source to another sequence of tokens in the target language.

^c<https://vlibras.gov.br/>.

^d<http://www.prodeaf.net/>.

^e<http://portal.rybena.com.br/site-rybena/>.

^f<https://www.handtalk.me>.

^gAccording to Bonham (2015), “to gloss means using the written text of the spoken language to label and identify individual signs and concepts of the sign language.” In our work, a Libras gloss uses Brazilian Portuguese words for each sign or phrase that can be labeled. For example, the BP declarative sentence “Você bonito” (You are beautiful) would be transcribed into Libras gloss as “VOCÊ BONITO” (YOU BEAUTIFUL) since Libras does not use the connection verb “SER” (TO BE).

^hThe VLibras Suite is the result of a partnership between the Ministry of Planning, Development, and Management (MP) and the Federal University of Paraíba (UFPB). It consists of a set of tools for machine translation of digital contents (text, audio, and video) to Libras, making computers, mobile devices, and Web platforms accessible to Deaf people. Currently, VLibras is used in several governmental and private websites, among them the leading sites of the Brazilian government (brasil.gov.br), Chamber of Deputies (camara.leg.br), and the Federal Senate (senado.leg.br).

The results obtained in the experiments performed with the proposed solution indicated an improvement in the level of comprehension of the translation by Deaf users compared to the version described in Araújo *et al.* (2014) work, which does not perform syntactic-semantic adequacy. Besides, given the flexibility demonstrated by the proposed approach, it is possible to infer that the strategies developed in this research could be adapted to translate content to other SLs.

The remainder of this article is organized as follows. In Section 2, we describe some works related to the machine translation for SL. In Section 3, we present some critical linguistic aspects of Libras. In Section 4, the mechanism proposed in this work is better detailed, including the architecture of the new translation component and the description of the rules modeling language. In Section 5, we present a proof of concept of the proposed solution containing some examples of morphosyntactic-semantic adequacy rules created from the proposed language, the implementation of the new translation mechanism, and its integration to the VLibras Suite. Finally, we present the results obtained and related some future work in Sections 6 and 7, respectively.

2. Related works

Araújo *et al.* (2013) describe an architecture for machine translation of Brazilian Portuguese contents to Libras in the real-time open domain, such as Digital TV. The system called LibrasTV receives a closed-caption stream as input and generates a Libras window on the TV screen as output. The machine translation is performed in real time and uses some strategies to reduce translation time. These include a Libras dictionary with pre-rendered signs and an efficient mechanism for machine translation of Brazilian Portuguese contents to Libras gloss. The translation mechanism combines morphosyntactic transfer rules, defined by specialists; a statistical method of data compression to classify input tokens (words) in Brazilian Portuguese; and a textual simplification that reduces entry complexity before applying these translation rules. Araújo *et al.* (2014) modify the architecture presented in Araújo *et al.* (2013), expanding the usage scenarios to the Web and Digital Cinema. Human computing strategies are also incorporated to facilitate the creation of Libras signs and the definition of syntactic rules.

The work of Boulares and Jemni (2013) presents a machine translator from English to ASL (American Sign Language) through mobile devices, to make information accessible for Deaf people anywhere. The system is proposed for an open translation domain and uses a rule-based translation strategy. The system receives an SMS (Short Message Service) as input and generates an MMS (Multimedia Message Service) containing an ASL video represented by an avatar as output. The authors report that one of the difficulties of this system is that SMS is not always received quickly, which can lead to a delay in the translation.

Huenerfauth (2004) presents a proposal of machine translation of English texts to an ASL animation and test the boundaries of traditional architectural designs for machine translation. It involves a multi-pathway translation architecture that incorporates direct, transfer-based, and interlanguage-based translation approaches into a single system. He also models semantic-pragmatic elements called “classifier predicates.”

The Huenerfauth (2005a) project uses representations of speech, semantics, syntax, and (sign) phonology adapted for ASL generation. In particular, this system intends to translate contents, use classifiers, and mark the signing space. In Huenerfauth (2008a), he presents a hybrid rule-based machine translation approach combining transfer-based, interlanguage-based, and direct translation. As a mark of the Huenerfauth works, there is always the concern of the modeling classifiers in the sign language translation system.

However, one issue reported by the author (Huenerfauth 2008a) is the difficulty of designing an interlanguage-based solution for an open (or general) domain. Typically, they are only feasible for

specific (or limited) areas. Besides, there are English sentences where in-depth semantic analysis is not required. In this case, we could use a transfer-based translation approach.

The formal model named Partition/Constitute (P/C) is a method for representing a linguistic sign containing several parameters (Huenerfauth 2005b). This formalism is compared to representations used in research related to gesture animation and was used in the development of an English-to-ASL machine translation system.

Morrissey and Way (2014) use a statistical method for translating English into Irish Sign Language (ISL), German Sign Language (DGS), and a gloss annotation in DGS.¹ To perform this task, the authors use four aligned linguistic corpora, which make possible the translation of the 12 pairs of languages. This study also allows the conversion of both: spoken language-to-SL and SL-to-spoken language.

López-Ludeña *et al.* (2014a) propose the generation of Spanish Sign Language (LSE) gloss with a rule-based translation strategy. The system architecture consists of an analyzer, a transfer module, and a generator. The analyzer is responsible for identifying the morphological dependencies in the spoken language and the SL, generating a dependency tree that represents the functional relations between the words. The transfer module maps the dependencies in the spoken language and the SL and performs the lexical (lexical-morphological and semantic) and structural replacements. Finally, the generator is responsible for the ordering of words based on the LSE dependency tree, generating an LSE gloss. It is important to emphasize that the López-Ludeña *et al.* (2014a) solution is for a restricted area, specifically for attending some public services (e.g., withdrawal of driver's license).

The hybrid translation strategy (rules-based and statistical) used by the López-Ludeña *et al.* (2014b) had higher accuracy than rules-based and statistical approaches. However, the authors make a caveat that the dependence of a bilingual corpus may represent a bottleneck for its adoption, since LSE is a modern language and not yet fully consolidated. Another observation verified from the validation of their system is that it has higher accuracy in the translation of short sentences, that is, with few signs.

Rayner *et al.* (2016) present an open platform for rule-based speech-to-sign translation. The platform Lite Speech2Sign aims to combine components for speech recognition, translation, and sign language animation. It is a similar approach used in VLibras. This kind of modularization of the sequence of processes involved (speech-to-text, text-to-gloss, and gloss-to-sign) allows us to incorporate advances in specific technologies used at each stage of the process. The fundamental differences between the platforms are translation approaches and the sign language animation strategy. Lite Speech2Sign adopts transfer-based machine translation and synthesis from form notation, whereas VLibras uses rule-based with syntactic-semantic adequacy and handcrafted animation. While the first strategy allows greater ease in the prototyping new simple translation grammars, the second one can provide a better result.

In the opposite direction, we can cite several initiatives translating continuous sign language to text, such as the SignSpeak Project (2010) and, more recently, the work of Kim and O'Neill-Brown (2019). The overall goal of SignSpeak is to develop new vision-based technology for translating continuous sign language to text to improve their Deaf communication community with the hearing people and the other way.

Kim and O'Neill-Brown research aim the challenges of building an automated sign language recognizer (ASLR), due to the lack of sufficient quantities of annotated ASL-English parallel corpora for training, testing, and developing. In particular, the authors explore a range of different techniques for automatically generating synthetic data from existing datasets to improve the accuracy of ASLR. In common with our approach, both projects consider a textual representation of the recognized signs as an intermediate step before message translation and rendering.

¹Proposed by Daniel Stain of RWTH Aachen such as Stein *et al.* (2010).

As identified by Kim and O'Neill-Brown (2019) and Morrissey and Way (2014), considering the difficulty of having annotated sign language-spoken language parallel corpora, especially for general (or open) domain, in this paper, we chose to develop a rule-based translation approach from BP to Libras, instead of a statistical or neural approach. Besides, unlike the works of Huenerfath (2005a, 2005b, 2008a) and López-Ludeña *et al.* (2014a, 2014b), which have solutions addressed to limited (or specific) domains, we designed our approach addressing general domain texts from BP into Libras, treating some syntactic-semantic adequacy.

Additionally, we also model a formal language to describe the translation rules from Brazilian Portuguese to Libras, and a simple string-based notation for Libras gloss. This notation helps simplify the sign language animation implementation, which has to display a sequence of signs (simple or compound) from the gloss. This is an alternative path with respect to notations like the one proposed by Huenerfath, which requires to render a 3D animation coordinating multiple channels (e.g., right and left hand, head tilt, eye gaze, among others) and makes its implementation more complex. We discuss this issue in more detail in Section 5.2.

Finally, the approach by Araújo *et al.* (2013, 2014) presents some limitations, not addressing some essential aspects of a machine translation solution to Libras, such as the treatment of verbal tense, semantic appropriateness of the use of adverbs of intensity and negation, treatment of connecting verbs, among others. As the main contribution of this work, our proposal for a new machine translation from BP into Libras mechanism seeks to address these aspects. The comparison between our proposal and the proposal by Araújo *et al.* will be further detailed in Section 4.

3. Linguistic aspects of libras

Libras is a natural human language. It has grammatical rules common to the Deaf people in Brazil, linguistic variation according to their specific community (dialects), and its own phonological, morphological, syntactic, and semantic-pragmatic system. Besides, Libras is a cultural expression of the Brazilian Deaf community.

Concerning the phonological aspects, the minimum units of their linguistic system are the quiremas.^j The quirema is a formational unit of the sign, that is, it is composed of the articulation between the parameters of the sign language. The sign consists of five basic parameters (Figure 1), that are handshape (HS) (Subfigure 1D), location (L) (Subfigure 1C), hand movement (M) (Subfigure 1E), orientation (Or) (Subfigure 1A), and non-manual aspects of signs (NM) (Subfigure 1B) (Quadros and Karnopp 2004).

In terms of sign language morphology, the complexity of morphemes is a result of non-concatenated processes, that is, the root is enriched with movements and contours in the signing space (Quadros and Karnopp 2004).

The syntactic aspects of Libras consider the sign execution space because the syntactic relations make use of the pronominal and nominal system for this purpose. According to Felipe (2007), since SLs are gestural-visual, they use the three-dimensionality of the space as a grammatical element. Non-manual aspects are also relevant in Libras syntax, since utterances with verbs with an agreement in their formation must have this mark (Quadros and Karnopp 2004). However, it is not only in this situation that the non-manual aspects must appear. We also use them in marking the location of referents.

The semantic and pragmatic aspects are related to the linguistic meaning of the utterance. In Libras, the syntactic elements of the “verb” influenced them. According to Silva (2006), the verb is also quite influential in the semantic relations of the language. It happens due to the importance of the meaning of the verb to understand its behavior and to predict its syntactic properties. Thus, we classify the verbs considering their morphological and semantic characteristics in Libras.

^jQuirema is a term proposed by Stokoe (1980) to identify the formational units of the signs, similar to the phonemes of the spoken languages.

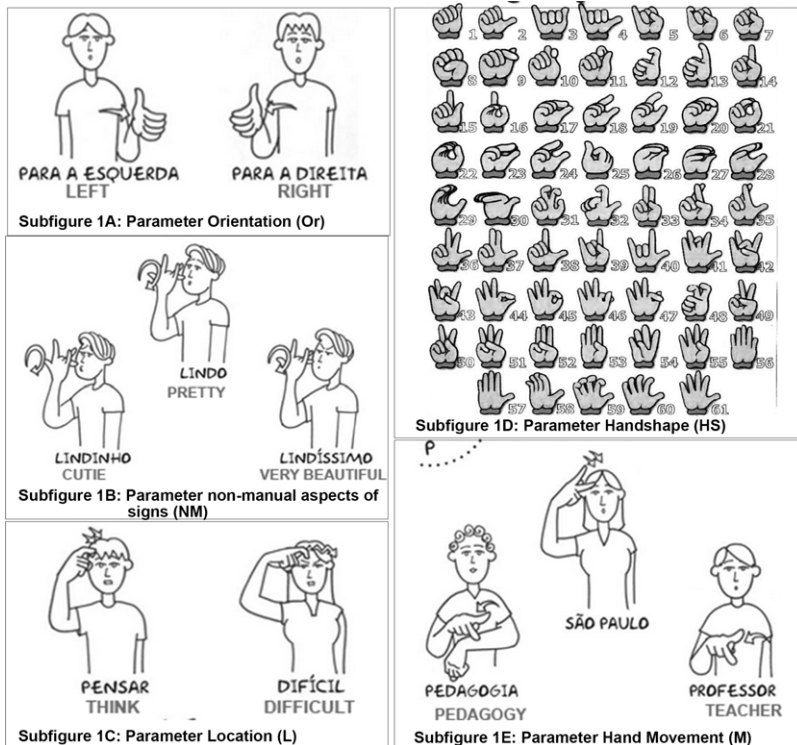


Figure 1. Parameters of the sign language (Choi *et al.* 2011).

4. The proposed approach

As mentioned in Section 2, the VLibras platform proposed by Araújo *et al.* (2014) has some limitations in its machine translation component, especially regarding the handling of syntactic and semantic issues. Araújo *et al.* (2013) translation strategy designed their solution to translate general domain content efficiently (quickly). To perform this task, it combines statistical compression methods to classify input tokens, textual simplification strategies to reduce the complexity of input text, and a set of morphological and syntactic rules defined by Libras specialists. Figure 2 shows the architecture of the machine translation component described by Araújo *et al.* (2013, 2014).

According to Figure 1, their machine translation component is given in five steps. In the first step, the tokenizer splits the text in BP into a sequence of words (or tokens). Afterward, it classifies the tokens into morphological categories (Morphological-Syntactic Classifier step), using the PPM-C [29], a statistical data compression method based on N-order Markov models. Then, they simplify the text by removing prepositions and articles (Remove Tokens step). Afterward, they replace some words (Lexical Replacement step) to adapt the meaning of the sentence rewritten to Libras. For example, the words *casa* (house), *lar* (home), *habitação* (habitation), and *residência* (residence) in BP have the same sign in Libras: CASA (HOME). Non-inflected gloss verbs replace the BP verbs. Finally, they perform a dactylogogy replacement for fingerspelling the proper names and technical terms. The output generated is a Libras gloss.

Thus, the solution of Araújo *et al.* (2013, 2014). presents some limitations, not addressing some essential aspects of a machine translation solution to Libras, such as the treatment of verbal tense, semantic appropriateness of the use of adverbs of intensity and negation, treatment of copula verbs, among others.

Table 1. Comparison between the aspects of translation addressed in our proposal and the solution proposed by Araújo *et al.* (2013, 2014)

| | VLibras with the proposed solution | VLibras version of Araújo <i>et al.</i> |
|--|------------------------------------|---|
| Morphological analysis of BP | X | X |
| Syntactic analysis of BP | X | |
| Lexical replacement of synonymous words | X | X |
| Treatment of verbal tense | X | |
| Treatment of adverbs of intensity and negation | X | |
| Treatment of common nouns of two genera | X | |
| Treatment of copular verbs | X | |
| Qualified elimination of articles and prepositions | X | |

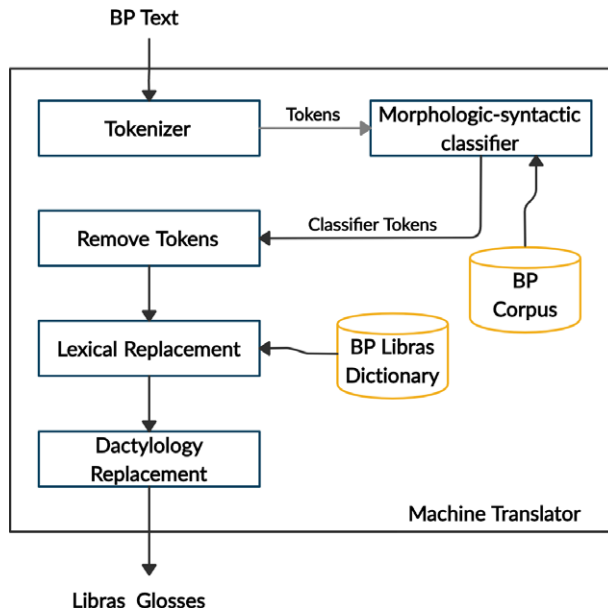


Figure 2. Architecture of the machine translation component defined by Araújo *et al.* (2013, 2014).

To address these limitations, in this work, we propose the development of a new mechanism to perform machine translation to Libras addressing these aspects. We also model a formal language to describe the translation rules from Brazilian Portuguese to Libras, making the translation process flexible, configurable, and able to deal with other linguistic aspects. Othero (2009) corroborates the importance of this type of language when he states that “the study of language syntax may be essential for its computational treatment at several other levels of linguistic description, and usually, the computational implementation work of a language involves the syntactic or morphosyntactic treatment of that language.”

Table 1 makes a summary comparison between the aspects of translation addressed in our proposal and the solution proposed by Araújo *et al.* (2013, 2014).

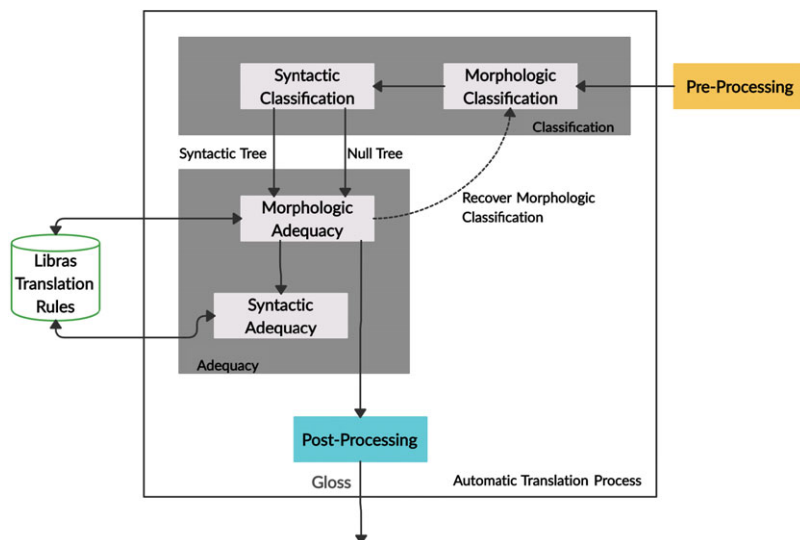


Figure 3. Architecture of the proposed machine translation mechanism.

Figure 3 illustrates the proposed machine translation mechanism. According to Figure 3, it works as follows. Initially, we apply a process of morphological and syntactic classification in the input text in BP. This process may produce two types of output: (1) a syntax tree or (2) a null tree. It produces a syntax tree for sentences recognized by the classifier grammar and generates a null tree when the sentences is unrecognized by the classifier grammar. Then, the Morphological Adequacy and Syntactic Adequacy modules apply a set of translation rules defined by specialists based on the proposed rule description language (see Section 4.1). When the classification module returns a null tree, the Adequacy module uses the Morphological classification (e.g., a POS-tagging process). In the latter case, the component applies only translation rules defined at the morphological level. As a result, the translation component generates a Libras gloss.

Afterward, we apply post-processing in this gloss. This process is responsible for performing some translation refinements, such as the treatment of numbers, plural, among others. The post-processing process is important because it reduces the possibility of fingerspelling^k during the presentation of visual gestures by the Libras generation service. More details about the implementation of a prototype of this machine translation mechanism will be presented in Section 5.2.

An essential contribution of our work is that we use a simpler Libras gloss representation to treat the primary sign language specificities described in Section 3. More specifically, we simplify Libras gloss as a sequence of tokens. Besides, we also include “special tokens” to incorporate expressions and the specific aspects of sign languages, such as directional verb agreement, incorporation of facial expressions during the signing, and intensification of signs. For example, in Libras, the adverb *NÃO* (*NO*) can modify the signing of the verb *TER* (*HAVE*) in a sentence. In this case, our translation rules address this issue, generating a special token *NÃO_TER* (*NOT_HAVE*), which will have a different signing than the *NÃO* (*NO*) and *TER* (*HAVE*) signs separately. Thus, we can represent this modification in the *TER* sign generated by the adverb *NÃO*, using a special sign *NÃO_TER*.

We can also address the treatment of the directional verb agreement using this representation. For example, the sentence “*EU PERGUNTO PARA VOCÊ*” (*I ASK YOU*) is translated to a

^kFingerspelling is the manual spelling of a word. It is generally undesirable because it refers to the textual representation of the word in the spoken language, which does not always make sense for SL’s user.

special Libras gloss “1S_PERGUNTAR_3S” (*1S_ASK_3S*), which represents the directional verb PERGUNTAR (*ASK*) with the directionality of the first person singular (represented by 1s) to the third person singular (represented by 3s).

This simpler representation of Libras gloss has the advantage of making the translation process from spoken languages into sign languages similar to a translation process between two spoken languages. It converts to a sequence of tokens in a source language to another sequence of tokens in the target language. As it uses a more straightforward and closer representation of the current bilingual corpus for machine translation, it could also simplify the creation of bilingual text-to-gloss translation corpus. Consequently, it can also possibly reuse some machine translation techniques currently available for use in this type of problem, when a bilingual text-to-gloss corpus is available.

Another advantage is that it reduces the sign language animation component’s complexity, which converts Libras gloss into a Libras representation (e.g., video or animation). When we use a multiple channel output representation (e.g., right hand, left hand, head tilt, eye gaze), as proposed by Huenerfauth (2005b), we would have to render a 3D animation coordinating multiple channels, which makes this component more complex.

Finally, it is also important to point out that our solution also supports syntactic-semantic adequacy. According to Felipe (2006), the incorporation, simultaneous activity, and directional verb agreement grammatical aspects, treated by our solution, are syntactic-semantic aspects of the sign languages, since they change the form (syntactic) and the meaning (semantics) of the sentence. Additionally, as our solution has a rule description language, we can incorporate translation rules to address other syntactic-semantic aspects.

In Section 4.1, we will present the rule description language developed for describing the translation rules used by our Adequacy module.

4.1. Rule description language

In the rule description language developed in this work, we define a sentence *S* formally as:

$$S = \langle sg_1, sg_2, sg_3, \dots, sg_n \rangle,$$

where sg_1, \dots, sg_n are noun, verbal, adverbial, adjectival, numeral, quantitative, prepositional, determinant, or possessive phrases. Each phrase sg_i is composed of subphrases (non-terminals) and/or words (terminals) and is formally defined as:

$$sg_i = \langle title, specific, action, newpos, newprop, newtoken, newtokenpos \rangle,$$

where *title* represents the morphological and/or syntactic class of the phrase; *specific* is an optional field that indicates the existence of restrictions in the rule application, that is, the rule is only applied to a specific subphrases and/or word; *action* indicates the type of operation to be performed by the rule; *newpos* represents the new position of the subphrase and/or word after the application of the rule. The value “-1” is used when the subphrase and/or word should be removed of the sentence; *newprop* is an optional field that indicates changes in the subphrase and/or word (e.g. incorporate the adverb of negation into adjectives, verbs, and adverbs); *newtoken* is also an optional field that indicates the addition of a new subphrase and/or word to the sentence (e.g. when a sentence does not present the adverb of time in Libras, the signs FUTURE and PAST must be added to indicate a reference to the future or the past tense, respectively); *newtokenpos* is also an optional field that represents the position of the new subphrase and/or word added in the sentence.

From the formalization described above, we define an XML representation to represent the attributes of translation rules. Each rule has a <counter> attribute that represents the number subphrases and/or words involved in that rule. It has also a <active> flag that indicates whether

```

<rule name = "S(VP_NE) ">
  <active>true</active>
  <count>2</count>
  <class>
    <action>invert</action>
  </class>
</rule>

```

Figure 4. Example of the translation rule “S (VP NP) → S (NP VP)” described with the proposed rule description language.

```

<rule name = "VB-R">
  <active>true</active>
  <count>1</count>
  <class>
    <title>VB-R</title>
    <action>change_vb</action>
    <newprop>VB</newprcp>
    <newtoken>FUTURO</newtoken>
    <newtokenpos>next</newtokenpos>
  </class>
</rule>

```

Figure 5. Example of the translation rule with sentence processing in future tense: (VP (VB-R *) → VP (VB * <FUTURE>), where * represents any phrase and/or word that is after the verb.

the solution should apply the rule or not, which facilitates the tests of the translation rules and the Libras specialists’ validation.

An example of the XML representation of a translation rule is shown in Figure 4. It indicates that a Brazilian Portuguese sentence with “Verbal Phrase (VP) + Noun Phrase (NP)” that is in reverse order must be translated to direct order (“Noun Phrase (NP) + Verbal Phrase (VP)”) in Libras.

Another example of a translation rule is shown in Figure 5. It indicates that in a BP sentence in the future tense, the verb must be modified to the infinitive form, followed by the sign FUTURE in Libras.

5. Proof of concept

As a proof of concept of the proposed solution, initially, we modeled a preliminary grammar using the rule description language described in Section 4.1. This grammar consists of a set of formal machine translation rules from Brazilian Portuguese (BP) to Libras, defined by Libras specialists. Afterward, we implemented a prototype of the proposed translation mechanism using this modeled grammar and integrated it with the VLibras Suite (Araujo *et al.* 2014). Sections 5.1 and 5.2 present more details about the modeled grammar and the prototype of the proposed translation mechanism, respectively. Section 5.3 describes the integration of this prototype into VLibras.

5.1. Definition of a grammar from BP to libras

Table 2 presents the grammar we modeled to translate content from Brazilian Portuguese to Libras. As mentioned in Section 4, we use a Libras gloss representation, which includes “special

Table 2. Grammar proposed using the rule description language

| Translation rules | Ex: BP → Libras gloss |
|--|--|
| (1) Direct order of the sentence | |
| S (VP NP) → S (NP VP) | BP: Estão descontraídos os garotos. (<i>Are relaxed the boys.</i>) Libras: GAROTOS ESTAR DESCONTRAÍDO. (<i>BOYS ARE RELAXED.</i>) |
| (2) Delete connecting verb | |
| VP (V AP) → VP (AP) | BP: Você é bonito (<i>You are beautiful</i>) Libras: VOCÊ BONITO (<i>YOU BEAUTIFUL</i>) |
| (3) Delete preposition | |
| PP (P _x DP) → PP (DP) {em, de, por, para} | BP: de menino (<i>of boy</i>) Libras: MENINO (<i>BOY</i>) |
| PP (P _x AdvP) → SP (AdvP) | BP: por aqui (<i>around here</i>) Libras: AQUI (<i>HERE</i>) |
| (4) Processing common noun of two genders | |
| DP (DA _(f) SN _c) → SN < MULHER> | BP: o artista (<i>the artist</i>) |
| DP (DA _(m) SN _c) → SN < HOMEM> | Libras: ARTISTA HOMEM (<i>ARTIST MAN</i>) |
| (5) Delete article | |
| DP (DA NP) → NP | BP: a mulher (<i>the woman</i>) |
| DP (DA PossP) → PossP | Libras: MULHER (<i>WOMAN</i>) |
| DP (DA NumP) → NumP | BP: os dois porcos (<i>the two pigs</i>) Libras: DOIS PORCOS (<i>TWO PIGS</i>) |
| (6) Processing verb with adverb of time^a | |
| VP (VB-*SAdv _(t) P) → VP (VB SAdv _(t) P) | BP: Ontem trabalhei (<i>yesterday worked</i>) |
| VP (SAdv _(t) VB-*) → VP (Adv _(t) P VB) | Libras: ONTEM TRABALHAR (<i>YESTERDAY WORK</i>) |
| (7) Processing verb in the present, future, and past | |
| VP (VB-P DP) → VP (VB DP) | BP: Beboágua (<i>Drink water</i>) |
| VP (VB-P PP) → VP (VB PP) | Libras: BEBER ÁGUA (<i>DRINK WATER</i>) |
| VP (VB-P AP) → VP (VB AP) | BP: Estudarei Libras (<i>Will study Libras</i>) |
| VP (VB-P DP) → VP (VB DP < FUTURO>) | Libras: ESTUDAR LIBRAS FUTURO (<i>STUDY</i> |
| VP (VB-P PP) → VP (VB PP < FUTURO>) | <i>LIBRAS FUTURE</i>) |
| VP (VB-P AP) → VP (VB AP < FUTURO>) | BP: Aprendi inglês (<i>Learned english</i>) |
| VP (VB-P DP) → VP (VB DP < PASSADO>) | Libras: APRENDER INGLÊS PASSADO (<i>LEARN</i> |
| VP (VB-P PP) → VP (VB PP < PASSADO>) | <i>ENGLISH PAST</i>) |
| VP (VB-P AP) → VP (VB AP < PASSADO>) | |
| (8) Processing adverbs of intensity(i) and negation(n) | |
| VP (VB Adv _(i)) → VP (V _{Mi}) | BP: trabalhar muito (<i>work very</i>) |
| VP (Adv _(n) VB _(n)) → VP (VN) | Libras: TRABALHAR_MUITO (<i>WORK_VERY</i>) |

Table 2. Continued

| Translation rules | Ex: BP → Libras gloss |
|---|--|
| VP (Adv _(n) VB) → VP (V _{Mn}) | BP: não ter (<i>no have</i>) |
| AP (Adv _(i) Adj) → AP (Adj _{Mi}) | Libras: NÃO_TER (<i>NO_HAVE</i>) |
| AP (Adj Adv _(i)) → AP (Adj _{Mi}) | BP: muito bonito (<i>very beautiful</i>) |
| AP (Adv _(n) Adj) → AP (Adj _{Mn}) | Libras: MUITO_BONITO (<i>VERY_BEAUTIFUL</i>) |
| AdvP (Adv Adv _(i)) → SAdv(Adv _{Mi}) | BP: Não bonito (<i>no beautiful</i>) |
| AdvP (Adv Adv _(i)) → SAdv(Adv _{Mi}) | Libras: NÃO_BONITO (<i>NO_BEAUTIFUL</i>) |
| AdvP (Adv _(i) Adv) → SAdv (Adv _{Mi}) | |

^oP, R, D (present, future and past).

tokens” to address the specific aspects of sign languages, such as directional verb agreement, incorporation of facial expressions during the signing, intensification of signs, among others.

According to Table 2, the rule (2) intends to eliminate the Brazilian Portuguese connecting verb. In BP, the connecting verbs are not significant and do not indicate action. These verbs appear in the verbal phrase S(VP) that have a verb (VB-*) followed by an adjectival phrase (AP). In this case, the corresponding VP in Libras will be formed only by the AP (See Table 2, rule 2).

Another rule example is the elimination of prepositions. In Libras, some prepositions are not signing and consequently must be eliminated in the translation process. We identify this rule as (3) in Table 2, which indicates that: when we have a prepositional phrase (PP) composed by a specific preposition (P_x) and a definite phrase (DP) or an adverbial phrase (AdvP), the (P_x) is excluded.

When we have a common noun of two genders, which represents nouns having the same writing in both males and females, we can apply the rule (4). When dealing with a common noun of two genders, the sign HOMEM (MAN) or MULHER (WOMAN) is added in the Libras interpretation to sign the gender. Then, the rule (4) indicates that whether a masculine (DA_(m)) or definite feminine article (DA_(f)) is followed by a common noun of two genders (SN_c) in BP, we remove the definite article and add a HOMEM or MULHER sign to represent the gender, respectively.

5.2. Implementation of the proposed translator mechanism

To validate and develop a proof of concept of the proposed approach, we implement the proposed machine translation mechanism and integrate it into VLibras (Araujo *et al.* 2014). We chose VLibras because it is an open-source BP-to-Libras translator, designed for three usage scenarios (Digital TV, Web, and Digital Cinema), and performs translation for an open and general domain.

We developed the proposed translation component in the Python programming language, according to the architecture presented in Figure 2. The Aelius tool¹ (Alencar 2010), an open-source POS-tagger library for Brazilian Portuguese, performs the morphosyntactic classification of the input sentences. The Morphological Classification step generates a list of tuples, where each tuple is composed of the word and its respective morphological classification. From this list, it performs a Syntactic Classification. As mentioned in Section 4, this process may produce a syntax tree (if the classifier recognizes the sentence) or a null tree (if the classifier has unrecognized the sentence). The null tree can happen in two situations: the sentence structure is not adequate in BP or due to a limitation of the grammar used by the parser.

Afterward, the Morphological and Syntactic Adequacy modules apply some translation rules defined by Libras specialists based on the grammar and language presented in Sections 5.1 and 4.1, respectively. In the scope of this proof of concept, we define some rules for the treatment

¹<http://aelius.sourceforge.net/>.

Table 3. Some examples of the use of translation rules in the proposed translation component. The first column shows the input sentence in BP and the output generated in Libras gloss. The second column represents the translation rules applied to generate that output

| Input in BP → Output in Libras gloss | Translation Rules applied |
|---|---|
| Ele comprou o carro → ELE COMPRAR CARRO PASSADO (He bought the car) → (HE BUY CAR PAST) | DP (DA NP) → NP VP (VB-D DP) → VP (VB DP < PASSADO>) |
| Ele comprará o carro → ELE COMPRAR CARRO FUTURO (He will buy the car) → (HE BUY CAR FUTURE) | DP (DA NP) → NP VP (VB-R DP) → VP (VB DP < FUTURO>) |
| Ele compra o carro → ELE COMPRAR O CARRO (He buys the car) → (HE BUY CAR) | DP (DA NP) → NP VP (VB-P DP) → VP (VB DP) |
| Ele comprou o carro ontem → ELE COMPRAR CARRO ONTEM (He bought the car yesterday) → (HE BUY CAR YESTERDAY) | DP (DA NP) → NP VP (VB-*SAdv _(t) P) → VP (VB SAdv _(t) P) |
| Ontem ele comprou o carro → ONTEM ELE COMPRAR CARRO (Yesterday he bought the car) → (YESTERDAY HE BUY CAR) | DP (DA NP) → NP VP (SAdv _(t) VB-*) → VP (Adv _(t) P VB) |
| A presidente morreu → PRESIDENTE MULHER MORRER PASSADO (The president died) → (PRESIDENT WOMAN DIE PAST) | DP (DA NP) → NP DP (DA _(f) SN _c) → SN < mulher> VP (VB-D DP) → VP (VB DP < PASSADO>) |
| O presidente morreu → PRESIDENTE HOMEM MORRER PASSADO (The president died) → (PRESIDENT MAN DIE PAST) | DP (DA NP) → NP DP (DA _(m) SN _c) → SN HOMEM VP (VB-D DP) → VP (VB DP < PASSADO>) |

of verbal tense, treatment of common nouns of two genders, elimination of articles, semantic appropriateness of the use of adverbs of intensity and negation, treatment of connecting verbs, among others. We present some examples of the application of these rules in Table 3.

Finally, we apply the post-processing module to the treatment of numerals, plural terms, among others. More specifically, it performs the following steps:

- Synonym treatment: performs simple substitution of synonyms, using a list of synonyms in BP whose signing is the same in Libras. Ex: RESIDÊNCIA → CASA (RESIDENCE → HOME).
- Treatment of cardinal numbers: performs substitution of the number writing for its algebraic representation. Ex: DOZE → 12 (TWELVE → 12).
- Treatment of ordinal numbers: performs substitution of the algebraic writing (and the extended writing of the numbers from 10^o) to the long writing and referencing only the sequence of numbers from 0 to 9. Ex: 12^o or DÉCIMO SEGUNDO → PRIMEIRO SEGUNDO (12^o or TWELFTH → FIRST SECOND).
- Treatment of gender: performs substitution of pronouns, adjectives, and adverbs for their corresponding masculine version, since the signing in Libras is similar. Ex: BONITA → BONITO (BEAUTIFUL → BEAUTIFUL).
- Treatment of plural terms: performs substitution of pronouns, adjectives, and adverbs in the plural for their corresponding singular representation, since Libras signing does not modify them in plural form. Ex: MEUS → MEU (MY → MY).

Figures 6 and 7 illustrate how the proposed machine translation module works for the input sentence “Ele comprou o carro” (He bought the car). Figure 6 shows the syntax tree generated

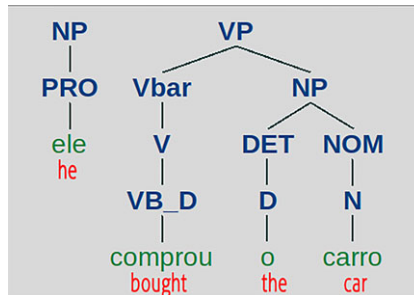


Figure 6. A syntax tree example for the input sentence “Ele comprou o carro” (He bought the car).

```

>>>
>>> from Tradutor import iniciarTraducao
>>> iniciarTraducao( "Ele comprou o carro.")
[[u'ELE', 'PRO'], [u'COMPRAR', 'VB-D'], [u'CARRO', 'N'], ['PASSADO', 'T-VB']]
>>>

```

Figure 7. Output generated by the proposed Machine Translation Process.

by the morphosyntactic classification process. In this case, after this step, the Adequacy module is invoked and applies rules 5 and 7 (see Table 2), followed by the post-processing module. Figure 7 shows Libras gloss generated as output as well as its labeling. We then pass this gloss to the Animation component, which will convert it to a Libras animation or video.

The generated syntactic tree has the sentence (S) initially formed by a nominal phrase (NP - “ELE”) and a verbal phrase (VP - “COMPROU O CARRO”). Since NP is an end node, that is, it does not have any other phrase in its constitution, the morphological classification of the word “ELE” is PRO (PRONOME). VB is not only composed by a terminal node, and therefore its classification generates a terminal node (VBar - “COMPROU”) and another nominal phrase (NP - “O CARRO”). The NP produced from the VP is a terminal node with the letter “O” morphologically classified as a determinant (DET) and then as an article (D), while the word “CARRO” is classified as a name (NOM) and then as a noun (N). Finally, the terminal node VBar is classified morphologically as a verb (V) and later as VB_D (verb in the past), since this morphological class carries information about when the action happened.

It is important to point out the output of this component is just a sequence of signs. As mentioned in Section 4, one of the main reasons for this is to reduce the complexity of the Synthesis (or Animation) component, which converts Libras gloss into a Libras 3D animation, since when we use a multiple channel output representation, we would have to render a 3D animation coordinating multiple channels. Thus, the Animation component has to display a sequence of signs (simple or composite) from Libras gloss, preventing it from making specific adaptations (e.g., including non-manual expressions) in the presented signs.

5.3. Integration of the proposed translator into VLibras

The integration of the proposed translator component into VLibras is illustrated in Figure 8. We inserted the new translator component after the components that filter and extract the text in BP (i.e., the Speech Recognition and Subtitle Extractor components), and before the Synchronizer component. Thus, the translator component receives an input text in BP and generates a Libras gloss sent to the Synchronizer component.

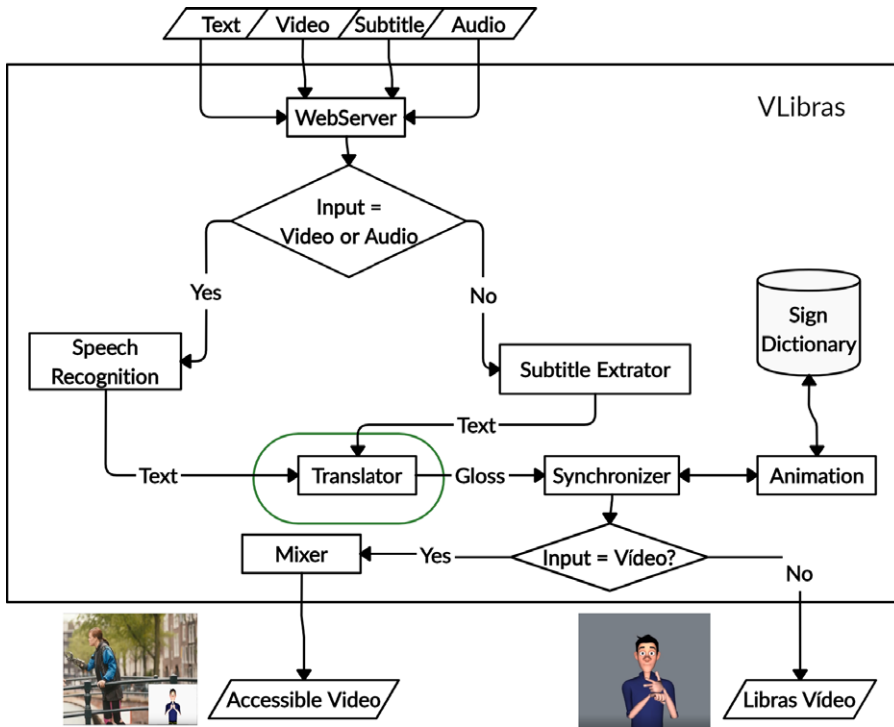


Figure 8. Schematic view of the integration of the proposed translation component into VLibras.

More specifically, the system works as follows. Initially, a user submits a digital content (text, video, subtitle, or audio) to the VLibras service. A web server receives this content and sends it to a component that processes it and extracts a text in BP from it. For example, if the content is an audio or a video, it triggers a speech recognition component to generate an input text in BP. Otherwise, if the content is a subtitle, it triggers the Subtitle Extractor to extract the text from this subtitle. Then, it sends the text to the proposed translation mechanism, which converts it to Libras gloss. Then, this gloss is sent to the Synchronizer, which, together with the Animation component, is responsible for generating a Libras video or animation.

Afterward, if the input stream is a video, Libras video can be mixed in the original video, generating an accessible copy of this video, that is, the user video with a Libras window. Then, we return this accessible video to the user. Another option is only to generate the Libras video or animation as output, usually when the input is a text, an audio, or a subtitle.

It is also important to point out that we currently implemented this proposed machine translation mechanism and put it in production on VLibras.^m Currently, VLibras is installed on thousands of Brazilian websites and is accessed by millions of users daily. It also helps to show the feasibility of the solution and its contribution in a real use scenario. VLibras’ current sign dictionary has about 15,000 signs and a significant portion of them consisting of “special glosses” that address the specific aspects of sign language mentioned in Sections 3 and 4.

In Section 6, we will present some results taken from the proposed approach to evaluate the translated contents.

^mThe VLibras source code is available at <http://gitlab.sgd.nuem.gov.br/vlibras/vlibras-library/vlibras-translate>.

6. Results and discussions

6.1. Computational tests

Initially, we performed some computational tests to evaluate the machine translation contents generated by the proposed solution. These tests were performed using the WER (Word Error Rate) (Niessen *et al.* 2000) and BLEU (Bilingual Evaluation Understudy) (Papineni *et al.* 2002) metrics. According to Melo, Matos, and Dias (2014), these metrics are intended to evaluate machine translation systems with greater economy, speed, and language independence than evaluations performed manually. Besides, we also chose these metrics because they were also used in some related works (San-Segundo *et al.* 2008; Su and Wu 2009; Araujo 2012).

To perform this test, we used the same 69 sentences selected by Araujo *et al.* (2014) to evaluate their work. Thus, we could compare our results with previous work.ⁿ Afterward, we calculate the WER and BLEU values from our solution using the reference translations generated by the Libras interpreters and perform a comparison with the results of Araujo *et al.* (2014).^o

Table 4 shows the WER and BLEU values for the two solutions. To generate the BLEU score, we use the SacreBLEU reference implementation proposed by Post (2018).

According to Table 4, in these tests, the BLEU and WER scores for the VLibras version with the proposed translator mechanism were better than those taken with the version of Araújo *et al.* (2014). Concerning the BLEU metric, we can observe a improvement of around 58 percentage points, considering the sentences used in the test. Concerning the WER metric, we can observe that there was a reduction of the error rate by more than 50 percentage points for the translation using the proposed approach.

Table 5 presents some sentences extracted from the 69 sentences from Bosque and their translation using the proposed solution and using the version by Araújo *et al.* (2014). Analyzing these sentences, we can observe some improvements implemented by the solution proposed in comparison with the version proposed by Araújo *et al.* (2014).

According to Table 5, we can notice that in sentences in the present tense (e.g., sentences 1 and 4), there is a high similarity of the translations between the two versions of the translator. However, the translation of the proposed solution can correctly address the elimination of articles and adequately handle the symbol referring to the identifier of the Brazilian currency (R\$ - real), which has its sign and a relevant and semantic load.

In sentences 2, 3, and 5, we can also observe that our approach's translation rules allow us to eliminate the copular verb, while the translator by Araújo *et al.* (2014) incorrectly maintains the copulate verb. Besides, it also maintains the form of the most usual phrase, that is, sentences in the Subject-Verb-Object (SVO) structure. In sentence 3, we can also note that the translator of the proposed solution can identify the verb tense and allows the signing of the semantic content that identifies when the action occurred from the application of predefined rules. On the other hand, the translation generated by the translator of Araújo *et al.* (2014) indicates that the action is taking place in the present tense. Finally, in sentence 5, we can observe the translator of the proposed solution, differentiate between the preposition (DE) and the conjugated verb (DE - DAR in VP), and correctly eliminate the preposition since this it is an accessory term and therefore not flagged in Libras.

ⁿThe 69 sentences used in the evaluation of the Araujo *et al.* (2014) were randomly selected from the corpus Bosque (<https://www.linguateca.pt/Floresta/principal.html>), which was developed by the project "Synthetic Forest." The corpus has 9368 sentences and 186,000 words in BP, extracted from the Brazilian newspaper "Folha de São Paulo." Linguists morphologically and syntactically classified all the sentences. These 69 sentences were translated by two Libras interpreters to Libras gloss, following some defined criteria (e.g., write cardinal numbers in numerical form; write out the ordinal numbers, among others). The vocabulary used in these 69 sentences has 1.106 different words.

^oThe reference translations generated by the Libras interpreters, the output generated by our solution, and the output generated by the VLibras solution of Araujo *et al.* (2014) for the 69 sentences used in this tests are available at http://drive.google.com/drive/folders/1jNqIaucVXL6BIvabk770BUdFCAC_W99t?usp=sharing.

Table 4. BLEU and WER test results

| | | VLibras with the proposed solution | VLibras version of Araújo <i>et al.</i> |
|------|---|------------------------------------|---|
| BLEU | – | 76.11% | 18.10% |
| WER | – | 31.70% | 85.23% |

Table 5. Comparison of glosses generated from the Bosque phrases translated by Araújo *et al.* (2014) and the proposed solution

| Input in BP | Output in Libras gloss Araújo <i>et al.</i> (2014) | Output in Libras gloss VLibras Proposed Solution |
|---|---|--|
| O projeto original do governo destinava ao TSE R\$ 334,9 milhões <i>(The original government project allocated R\$ 334.9 million to TSE)</i> | PROJETAR ORIGINAL O GOVERNAR DESTINAR A TSE 334,9 MILHÕES <i>(PROJECT ORIGINAL THE GOVERNMENT ALLOCATE THE TSE 334.9 MILLIONS)</i> | PROJETO ORIGINAL GOVERNO DESTINAR PASSADO TSE R\$ 334,9 MILHÕES <i>(PROJECT ORIGINAL GOVERNMENT ALLOCATE PAST TSE R\$ 334.9 MILLIONS)</i> |
| Aqui só joga quem está bem <i>(Only those who are well play here.)</i> | AQUI JOGAR QUEM SÓ ESTAR BEM <i>(HERE PLAY WHO ONLY BE WELL)</i> | AQUI SÓ JOGAR QUEM BEM <i>(HERE ONLY PLAY WHO WELL)</i> |
| Ele só não jogava porque não estava bem <i>(He just didn't play because he wasn't doing well)</i> | ELE SÓ JOGAR PORQUE NÃO ESTAR BEM NÃO <i>(HE ONLY PLAY BECAUSE NO BE WELL NO)</i> | ELE SÓ NÃO JOGAR PASSADO PORQUE NÃO BEM PASSADO <i>(HE ONLY NO PLAY PAST BECAUSE NO WELL PAST)</i> |
| Essa divisóo gera algumas distorções terríveis <i>(This division generates some terrible distortions)</i> | ESSA GERAR ALGUMAS DIVISÃO DISTORÇÕES TERRÍVEIS <i>(THIS GENERATE SOME DIVISION DISTORTIONS TERRIBLE)</i> | ESSA DIVISÃO GERAR ALGUMAS DISTORÇÕES TERRÍVEIS <i>(THIS DIVISION GENERATE SOME DISTORTIONS TERRIBLE)</i> |
| Os três estão presos desde 30 de julho de 93 <i>(The three have been in prison since July 30, 93.)</i> | OS ESTAR PRESOS 3 DESDE 30 DAR JULHO DAR 93 <i>(THE PRISON 3 SINCE 30 GIVE JULY GIVE 93)</i> | 3 PRESOS DESDE 30 JULHO 93 <i>(3 PRISON SINCE 30 JULY 93)</i> |

However, this result alone is not enough to conclude that the translation proposal is good or bad, but we can observe some signs of improvement in translation quality with the proposed approach.

6.2. Tests with users

According to Su and Wu (2009), the evaluation based on computational metrics is not enough to evaluate the quality of translation for sign languages since these languages are gestural-visual. Besides, the computational metrics of string comparison are not enough to assess automatic translations of sign languages since they have syntactic, semantic, and pragmatic information that is transmitted by facial and body expressions, for example. Thus, we also evaluated our solution with users.

The central aspect studied by the experiment was the level of comprehension from the Libras users' point of view, whether Deaf or hearing (interpreters and Libras teachers). To perform this

task, we confronted the accessible contents generated by the proposed solution with the contents generated by the VLibras version of Araújo *et al.* (2014).

The experiment was performed using an electronic form. It consists of a set of questions related to the translation of five sentences (in total) previously and randomly selected from the Bosque linguistic corpus (Freitas, Rocha, and Bick 2008).

The questions were multiple-choice, each with four alternatives (A, B, C, or D) and were designed to assess whether users had understood the content of the five sentences translated into Libras, such as “How much did it cost?”, “When did it happen?”, “Did he do it? Why?”, among others. The alternatives A, B, and C represented possible answers to the question, where only one of the options is correct. For all questions, the fourth alternative (D) is a “I do not know” option, included preventing users from randomly choosing one of the options when they did not know the correct answer.^P

Since we have performed the experiments with Deaf and hearing users, they did not have access to the written representation of the sentences even in Brazilian Portuguese or in Libras gloss. We did this to prevent users who understand the written representation in Brazilian Portuguese (or Libras gloss) from being able to answer questions even if they do not understand Libras signing. In other words, as they only had access to Libras sentence, they would only be able to answer the questions if they understood its signing in Libras.

We recorded the Libras interpretation of all questions and alternatives using a human interpreter. We machine-translated the five selected sentences using the VLibras in both versions: (1) original (Araújo *et al.* 2014) and (2) with the proposed translation mechanism. The output of the translation for each Bosque sentence was a video signed using the VLibras avatar in both versions. Thus, for each sentence, the users saw a Libras video presented by the VLibras avatar containing the translation of the sentence using one version of VLibras, followed by the questions and alternatives in BP and Libras signed by a human interpreter.

Then, we generated two versions of the form: (1) with the translation of the five sentences using the original version and (2) with the translation of the five sentences using the proposed version. We inserted these two forms into an electronic page, which was responsible for managing the distribution of the forms to the users in a balanced way. Besides, the users accessed the form with controlled access (using Google login and password), to prevent them from responding to the form more than once.

6.2.1. Assessment of the level of comprehension

The experiment to evaluate the content’s level of comprehension had spontaneous participation of 23 Libras users. We recruit these users through invitations sent by social networks, and through sensitization carried out on visits to universities (Federal University of Paraíba - UFPB), schools (Piragybe Municipal Elementary School and Municipal School of Elementary Education Dumerval Trigueiro Mendes and the State School of Elementary and Middle School Maria Geny), support organizations (Association of the Deaf of João Pessoa—ASJP and Association of Disabled and Relatives—ASDEF) and other organizations (Foundation Integrated Support Center for the Disabled) to support Deaf people in the city of João Pessoa in Brazil. These users randomly received the forms containing the questions and the translations of the five sentences generated by one of the VLibras version. However, we do not identify the information about the treatment he was evaluating (1) or (2) in the electronic form.

Although 23 users performed the tests, the group of users was heterogeneous and composed of 12 Deaf, and 11 hearing users (interpreters and Libras teachers), where 15 were women and 8 were men. They ranged in age from 15 to 45 years, with a mean age of 31.9 years.

^PA translated English version of the questionnaire can be accessed through the following link: <http://twixar.me/2Thn>.

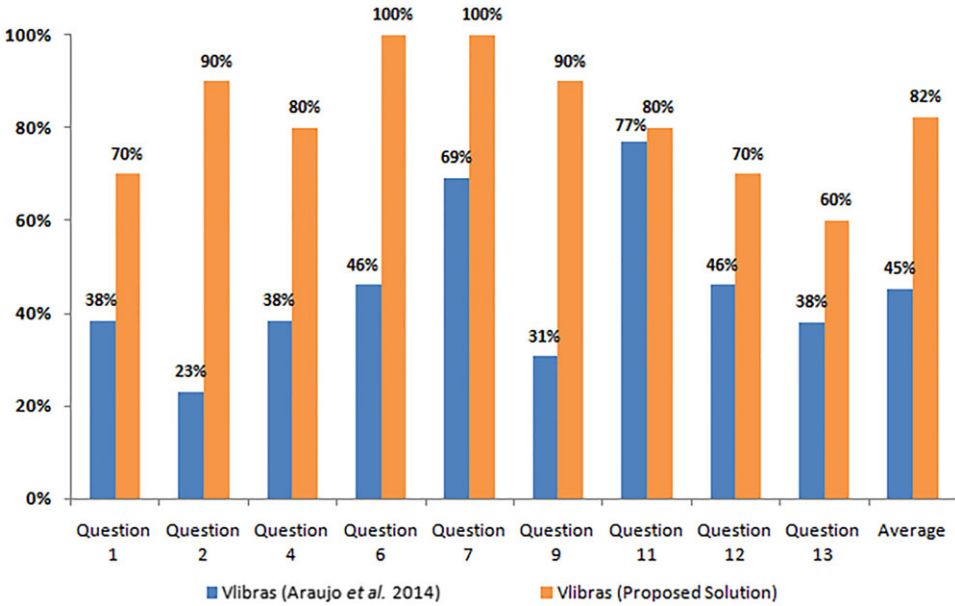


Figure 9. Content comprehension test results.

Figure 9 presents the results of the content comprehension test. According to Figure 9, we can observe that users that saw the translation of the five sentences performed by the proposed solution had a relatively higher hit rate for all evaluated contents, except for question 11.

The results also show that users that saw the translation performed by the proposed solution having an average hit rate of 82% with a standard deviation of 14%. In contrast, users that saw the translation performed by the VLibras described in Araújo *et al.* (2014) having an average hit rate of 45% with a standard deviation of 17%. The difference between the average hits between the two groups was approximately 35%.

An interesting aspect observed in the test is that although VLibras version of Araújo *et al.* (2014) did not address the common noun of two genders in the translation, the results show that 46% of users performed the test with this version answered the corresponding question (question 6) correctly. One possible explanation is that despite the lack of treatment of this issue, this version of the translator does not exclude the masculine article “O” (THE) of the sentence. Therefore, maybe some users who see the signing of the “O” (THE) article may associate it with a masculine noun. However, it is not the correct way to present the information in Libras.

As mentioned in Section 5.2, Libras does not have verbal tense, which we incorporate by adding a denotative time sign (e.g., PAST, FUTURE) in the sentence. Questions 2, 7, 11, and 13 dealt with the perception of verbal time expressed in the sentence. The sentence evaluated in Question 2 did not have an adverb of time in its structure. Consequently, the treatment proposed in this work produces an increase of 65% in the average rate of correct answers.

We also generated a boxplot with these data to analyze the dispersion of the results of this test (see Figure 8). According to Figure 10, in the comprehension tests performed with the proposed solution, the median, the first, and the third quartile values were around 89%, 72%, and 97%, respectively. It indicates that 50% of users hit more than 89% of the questions.

In the tests performed with the VLibras version of Araújo *et al.* (2014), the median, the first, and the third quartile were, respectively, around 44%, 33%, and 67%. It means that less than 50% of users had a hit rate higher than 50%.

Finally, Table 6 presents the evaluation of the users concerning the quality of the translation, using a scale from 1 to 4. According to Table 6, we can observe that the two solutions obtained

Table 6. Evaluation of quality of translation. Sentence 1–5 represents the five sentences used in the test, and the last column (Mean) represents the mean value for all sentences

| | Metrics | Sent.1 | Sent.2 | Sent.3 | Sent.4 | Sent.5 | Mean |
|---------------------------------|-----------|--------|--------|--------|--------|--------|------|
| VLibras (Araujo <i>et al.</i>) | Mean | 2.77 | 2.54 | 2.38 | 2.54 | 2.38 | 2.52 |
| | Std. dev. | 0.93 | 1.05 | 1.12 | 0.97 | 1.12 | 0.93 |
| Proposed solution | Mean | 2.80 | 2.60 | 2.40 | 3.10 | 2.80 | 2.74 |
| | Std. dev. | 0.79 | 1.07 | 1.26 | 1.10 | 1.03 | 0.77 |

Table 7. Mean and confidence interval of the test groups. N represents the number of test users, Mean represents the mean value of correct answers, Std dev. represents the standard deviation, and Std error represents the standard error of the mean

| | Group statistics (correct answers) | | | |
|---------------------------------|------------------------------------|------|----------|-----------|
| | N | Mean | Std dev. | Std error |
| VLibras (Araujo <i>et al.</i>) | 13 | 4.08 | 2.290 | 0.635 |
| Proposed solution | 10 | 7.40 | 1.897 | 0.600 |

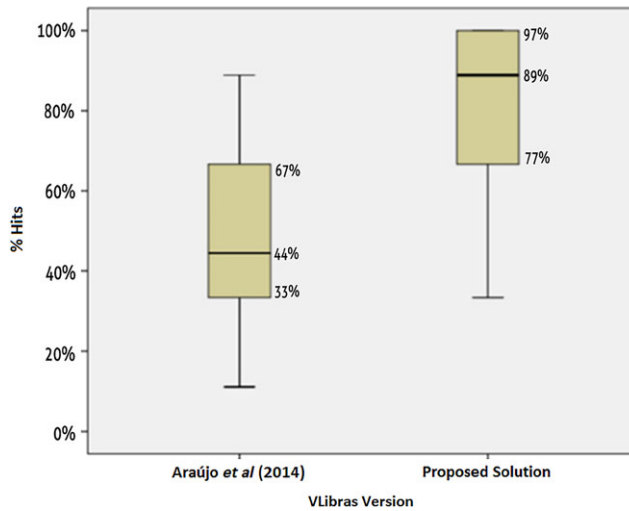


Figure 10. Boxplot of the content comprehension tests.

a moderate mean value for all the sentences. According to the users’ point of view, it is probably an indication that the two solutions generated content in Libras of moderate quality, with a slightly higher mean value for the proposed solution. However, it is essential to note that the difference between the mean value of the two solutions is small, which probably indicates that it is not statistically representative.

One possible explanation for this result would be the difficulty of interpreting the quality of the translation given to its subjectivity. Besides, according to Guimarães (2014), Deaf people have a problem evaluating using numerical scales. The author suggests that evaluation tests with Deaf users should prioritize binary questions for better understanding.

In the results demonstrated in Figures 6–8 and Table 6, we can observe that the proposed solution had better results in the comprehension tests with users. However, we can not say that there is

Table 8. Student’s t-test results for independent samples

| Independent samples test | | | | |
|---|---------------------------|-------------------------|-----------------------------|--------|
| | | Equal variances assumed | Equal variances not assumed | |
| Levene’s test for equality of variances | F | 0.432 | – | |
| | Sig | 0.518 | – | |
| | t | –3.708 | –3.803 | |
| | df | 21 | 20.842 | |
| | Sig(2-tailed) | 0.001 | 0.001 | |
| t-test for equality of means | Mean Diff | –3.323 | –3.323 | |
| | Std Err Diff | 0.896 | 0.874 | |
| | Confidence interval (95%) | Lower | –5.187 | –1.459 |
| | | Upper | –5.141 | –1.505 |

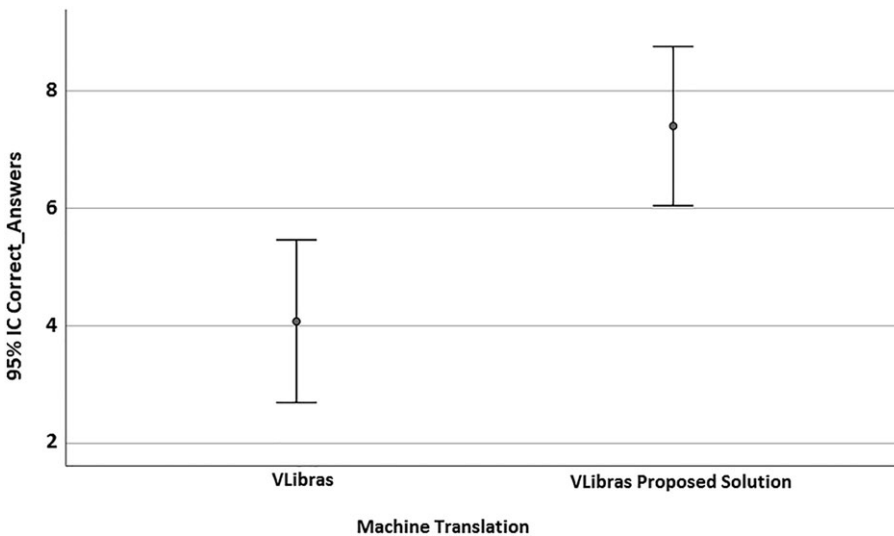


Figure 11. Statistical representation of the mean and confidence interval of the compared groups.

a significant improvement in results. To verify whether there is a statistically significant improvement, we apply a Student’s t-test. We chose this test because we can use it to evaluate if there is a statistically significant difference between the means of two paired or independent samples.

- H_0 : There is no statistically significant difference between the means of adjustment for the comprehension tests using the translation carried out by VLibras (Araujo *et al.* 2014) and the embedded translator of the proposed solution;
- H_1 : There is a statistically significant difference between the means of adjustment for the comprehension tests using the translation carried out by VLibras (Araujo *et al.* 2014) and the embedded translator of the proposed solution;

Figure 11 shows the statistical representation of the means and confidence interval of the results obtained in the comprehension tests with users. We can observe that there is no overlap of confidence intervals, so the averages of these groups are probably distant.

Table 7 shows that the 13 users who answered the test using VLibras (Araujo *et al.* 2014) averaged 4.08 of the questions. In contrast, those who answered the form using the proposed solution averaged 7.4 of the questions. We can also note that the standard deviation and the standard error of the mean obtained by the proposed solution are smaller to the results obtained with the VLibras approach (Araujo *et al.* 2014).

Analyzing the test of equality of variance presented in Table 8, we can note that the SIG is higher than 0.5. Thus, we can conclude that variance is homogeneous, so SIG (two-tailed) is used to verify the test's significance. As SIG (two-tailed) < 0.5 , we can state that the comprehension using the translation generated by VLibras embedded with the proposed solution is better than the version of VLibras (Araujo *et al.* 2014).

7. Conclusions

In this work, we proposed a new machine translation strategy for Libras, incorporating syntactic and semantic aspects. As contributions of this work, we can list the modeling of a formal language to describe translation rules from BP to Libras; the definition of a translation grammar of BP to Libras based on the modeled language; a more straightforward Libras gloss representation which allows representing the main specific aspects of Libras more directly; and the development and integration of a translation component into the VLibras (2014).

Our goal was to improve the translation level of comprehension VLibras, from the treatment of some syntactic and semantic aspects. To verify the proposal's quality and adequacy, we performed some computational tests using the BLEU and WER metrics. The results showed that there was a consistent improvement in the translation quality for the sentences evaluated. Besides, we also performed some tests with Libras users, Deaf or hearing (interpreters and Libras teachers) to verify if the aspects treated promoted an improvement in the comprehension of the contents generated by the proposed solution.

The results obtained for the BLEU 4-grams metric and the WER metric, 43% and 31%, respectively, help evaluate how the free-text translation (open-general domain) is a nontrivial task. However, we can observe that there is a considerable improvement in the level of comprehension of the contents by the users compared with the translation mechanism proposed by Araújo *et al.* (2014).

We believe in MT's importance for the Deaf community, whether we use it for assimilation or dissemination, and could play a complementary role for the digital inclusion of Deaf people, in addition to the other available resources. Following Araújo *et al.* (2014), "The main idea is to reduce barriers to access information, especially when interpreters are not available." By structuring rules that improve the quality of machine translation of such systems, it is envisaged that they can more effectively assist the Deaf community in improving their communication and access to information.

The use of a formal rule description language of spoken languages to SLs may also help to aggregate and disseminate the knowledge of linguists in the field, which may help in the emergence of new formal translation rules and morphosyntactic-semantic adequacy.

As mentioned in Section 5.3, we can also point out that our implementation is currently running in VLibras production service, which is installed on thousands of Brazilian websites and is accessed by millions of users daily. It also helps to show the feasibility of the solution and its contribution in a real use scenario. VLibras' current sign dictionary has about 15,000 signs and a significant portion of them consisting of "special glosses" that address the specific aspects of sign language mentioned in Sections 3 and 4.

As a proposal for future work, we aim to develop a web system where Libras specialists can describe translation rules in a more user-friendly way (e.g., through a friendly and intuitive interface). Besides, we intend to broaden the grammar of the proposed approach. We also plan to expand the proposed description language to describe machine translation rules for other SL, such as American Sign Language (ASL), Spanish Sign Language (LES), among others.

The computational cost is an ever-present challenge in rule-based approaches. To address this issue, we intend to adopt a strategy inspired by the research of Bacardit, Burke, and Krasnogor (2009). In the model proposed by Bacardit, instead of speeding up the operations for a given representation, the representation itself is modified to eliminate irrelevant computations. Large-scale experiments performed by the same authors support the observation that in real-world datasets containing a large number of attributes, only a small fraction of these are expressed in a given rule. In certain domains, this fraction can be as low as 5% or less of the available attributes. In our case, we will organize the rules in a decision tree that optimizes the path of comparisons.

Finally, we also intend to extend and evolve the tests with users, especially to identify which aspects of machine translation can be improved to make the content more understandable.

References

- Alencar L.F. (2010). Aelius: uma ferramenta para anotação automática de corpora usando o nltk. (*Aelius: a tool for automatic corpora annotation using nltk*). In *IX Encontro de Linguística de Corpus*. Campinas, SP, Brasil: Mercado das Letras, pp. 216–219.
- Araujo T.M.U. (2012). Uma solução para geração automática de trilhas em Língua Brasileira de Sinais em conteúdos multimídia (*A Solution for Automatic Generation of Tracks in Brazilian Sign Language in Multimedia Content*). PhD Thesis, Universidade Federal do Rio Grande do Norte, Natal, RN, Brazil.
- Araujo T.M.U., Ferreira F.L.S., Silva D.A.N.d.S., Lemos F.H., Neto G.P., Omaia D., de Souza Filho G.L. and Tavares T.A. (2013). Automatic generation of brazilian sign language windows for digital tv systems. *Journal of the Brazilian Computer Society* 19(2), 107–25, Springer, New York, NY, USA.
- Araujo T.M.U., Ferreira F.L.S., Silva D.A.N.d.S., Oliveira L.D., Falcão E.L., Domingues L.A., Martins V.F., Portela I.A.C., Nobrega Y.S., de Souza Filho G.L., Tavares T.A. and Duarte A.N. (2014). An approach to generate and embed sign language video tracks into multimedia contents. *Information Sciences* 281, 762–780, Elsevier, New York, NY, EUA.
- Bacardit J., Burke E.K. and Krasnogor N. (2009). Improving the scalability of rule-based evolutionary learning. *Memetic Computing* 01(01), 55–67, Springer, New York, NY, USA.
- Boulares M. and Jemni M. (2013). Toward a mobile service for hard of hearing people to make information accessible anywhere. In *International Conference on Electrical Engineering and Software Applications (ICEESA)*. Nova Jersey, USA: IEEE, pp. 1–5.
- Bonham M.E. (2015). *English to ASL Gloss Machine Translation*. Master Thesis, Brigham Young University, Provo, USA.
- Choi D., Vieira M.I.d.S., de Oliveira P.R.G. and Nakasato R. (2011). Libras: conhecimento além dos sinais (*Libras: Knowledge Beyond the Signs*). São Paulo, Brazil: Person Prentice Hall.
- Costa W.C.O. (2005). O texto traduzido como re-textualização (*The text translated as re-textualization*). *Cadernos de Tradução* 2(16), 25–54, UFSC, Florianópolis, SC, Brasil.
- Dreuw P., Forster J., Gweth Y., Stein D., Ney H., Martinez G., Llahi J.V., Crasborn O., Ormel E., Du W., Hoyoux T., Piater J., Moya J.M. and Wheatley M. (2010). SignSpeak: scientific understanding and vision-based technological development for continuous sign language recognition and translation. In *Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*. Valletta, Malta: Universität Hamburg, pp. 65–72.
- Felipe T. (2007). Descrição da língua de sinais: Desafios teóricos e práticos (*Sign language description: Theoretical and practical challenges*). In *Anais do Congresso Internacional do INES*. Rio de Janeiro, Brasil: Instituto Nacional de Educação de Surdos.
- Felipe T. (2006). Os processos de formação de palavra na Libras (*The processes of word formation in Libras*). *Educação Temática Digital* 7(2), 199–216, Unicamp, Campinas, SP, Brazil.
- Freitas C., Rocha P. and Bick E. (2008). Floresta sinta(ct)ica: bigger, thicker and easier. In *Computational Processing of the Portuguese Language*. New York, NY, USA: Springer, pp. 216–219.
- Guimarães A.P.N. (2014). Recomendações para avaliação da experiência de usuário em aplicativos móveis para surdos (*Recommendations for Evaluating the User Experience in Mobile Applications for the Deaf*). Master Dissertation, Universidade Federal da Paraíba, João Pessoa, Brazil.
- Huenerfauth M. (2004). A multi-path architecture for machine translation of english text into american sign language animation. In *Proceedings of the Student Research Workshop at HLTNAACL*, vol. 4. Boston, USA: Association for Computational Linguistics, pp. 25–30.

- Huenerfauth M.** (2005a). American sign language generation: multimodal nlg with multiple linguistic channels. In *Proceedings of the ACL Student Research Workshop*. Ann Arbor, USA: Association for Computational Linguistics, pp. 37–42.
- Huenerfauth M.** (2005b). Representing coordination and non-coordination in an american sign language animation. In *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility*, vol. 7, New York, NY, USA: ACM, pp. 44–51.
- Huenerfauth M.** (2008). Generating american sign language animation: overcoming misconceptions and technical challenges. *Universal Access in the Information Society* 6(4), 419–434, Springer, New York, NY, USA.
- Huenerfauth M., Zhao L., Gu E. and Allbeck J.** (2008). Evaluating american sign language generation through the participation of native asl signers. *ACM Transactions on Accessible Computing* 1(1), 211–218, ACM, New York, NY, USA.
- Kim, J. and O’Neill-Brown, P.** (2019). Improving American sign language recognition with synthetic data. In *Proceedings of Machine Translation Summit XVII Volume 1: Research Track*, vol. 19. Dublin, Ireland: European Association for Machine Translation, pp. 151–161.
- Lopez-Ludena V., Gonzalez-Morcillo C., Lopez J., Ferreiro E., Ferreiros J. and San-Segundo R.** (2014a). Methodology for developing an advanced communications system for the deaf in a new domain. *Knowledge-Based Systems* 52, 240–252, Elsevier, New York, NY, USA.
- Lopez-Ludena V., Gonzalez-Morcillo C., Lopez J., Ferreiro E., Ferreiros J. and San-Segundo R.** (2014b). Translating bus information into sign language for deaf people. *Engineering Applications of Artificial Intelligence* 32, 258–269, Elsevier, New York, NY, USA.
- Melo F.R., Matos H.C.d.O. and Dias E.R.B.** (2014). Aplicação da métrica bleu para avaliação comparativa dos tradutores automáticos bing tradutor e google tradutor (*Application of the metric bleu for comparative evaluation of automatic translators bing translator and google translator*). *E-Scripta* 5(3), 33–45, UNIABEU, Belford Roxo, Brasil.
- Morrissey S. and Way A.** (2013). Manual labour: tackling machine translation for sign languages. *Machine Translation* 27(1), 25–64, Springer, New York, NY, USA.
- Niessen S., Och F.J., Leusch G. and Ney H.** (2000). An evaluation tool for machine translation: fast evaluation for machine translation research. In *Proceedings of the Second International Conference on Language Resources & Evaluation (LREC), Athens, Greece*. Paris: France European Language Resources Association (ELRA), pp. 39–45.
- Othero G.A.** (2009) A gramática da Frase em Português: algumas reflexões para a formalização da estrutura frasal em português (*The Grammar of the Portuguese Phrase: Some Reflections for the Formalization of the Phrasal Structure in Portuguese*). Porto Alegre, Brazil: EdiPUCRS.
- Papineni K., Roukos S., Ward T. and Zhu W.J.** (2002). BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*. Boston, USA: ACL, pp. 311–318.
- Post M.** (2018). A call for clarity in reporting BLEU scores. In *Proceedings of the Third Conference on Machine Translation: Research Papers*. Boston, USA: Association for Computational Linguistics, pp. 186–191. doi: [10.18653/v1/W18-6319](https://doi.org/10.18653/v1/W18-6319).
- Quadros R.M.** (1997). *Educação de Surdos - Aquisição da Linguagem (Deaf Education - Language Acquisition)*. Porto Alegre, Brazil: Editora Artes Medicas.
- Quadros R.M. and Karnopp L.B.** (2004). *Língua de sinais brasileira: estudos linguísticos (Brazilian Sign Language: Language Studies)*. Porto Alegre, Brazil: Artmed Editora.
- Rayner E., Bouillon P., Gerlach J., Strasly I., Tsourakis N. and Ebling S.** (2016) An OpenWeb platform for rule-based speech-to-sign translation. In *54th Annual Meeting of the Association for Computational Linguistics (ACL)*. Boston, USA: Association for Computational Linguistics. Retrieved from <https://archive-ouverte.unige.ch/unige:87596>.
- San-Segundo R., Montero J.M., Macas-Guarasa J., Cordoba R., Ferreiros J. and Pardo J.M.** (2008). Proposing a speech to gesture translation architecture for spanish deaf people. *Journal of Visual Languages and Computing* 19(5), 523–538, Elsevier, Boston: USA.
- Silva M.P.M.** (2006). A Semântica como Negociação dos Significados em Libras (*Semantics as Negotiation of Meanings in Libras*). *Trabalhos em Linguística Aplicada* 45(2), 255–269, Unicamp, Campinas, SP, Brasil.
- Stein D., Forster J., Zelle U., Dreuw P. and Ney H.** (2010). Rwthphoenix: analysis of the German sign language corpus. In *Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, vol. 4. Valletta, Malta: Universität Hamburg, pp. 225–230.
- Stokoe W.C.** (1980). Sign language structure. *Annual Review of Anthropology* 9, 365–3690, JSTOR, Redwood City, USA.
- Su H.Y. and Wu A.H.** (2009). Improving structural statistical machine translation for sign language with small corpus using thematic role templates as translation memory. *IEEE Transactions on Machine Translation, Audio, Speech, and Language Processing* 17, 1305–1315, IEEE, Nova Jersey, EUA.