Lex-Libras: Morphosyntactic Model of the Brazilian Sign Language to Support a Context-based Machine Translation Process

Antônio M. Silva[1][a], Tanya A. Felipe[2][b], Laura S. García[3][c], Diego R. Antunes[4][d] and André P. Guedes[3][e]

[1]*Engineering Department, State University of Pará, Belém, PA, Brazil*

[2]*National Institute for Deaf Education, Rua das Laranjeiras, 232, Rio de Janeiro, RJ, Brazil*

[3]*Department of Informatics, Federal University of Paraná, Curitiba, PR, Brazil*

[4]*Informatics Department, Federal University of Technology, Ponta Grossa, PR, Brazil*

Keywords: Lex-Libras, Morphosyntactic Rules, Translation, Brazilian Sign-language.

Abstract: Brazilian Sign Language (BSL–Libras) is the preferential language of Deaf communities in Brazil. The Human-Computer Interaction Architecture in Sign Language (HCI-SL) was proposed, which will offer user-system interaction in BSL. In 2015 this architecture had its formal model developed, together with the phonological sign decomposition. The work described here advanced proposing morphological rules. Differently from American Sign Language, Libras has a group of verbs that inflect, so translators need to represent this phenomena in the automatic generation process and reflect it in its output, the 3D Avatar. Related Portuguese-BSL translators available have not yet been able to generate correct BSL sentences with regard to these verbs. This paper presents the Lex-Libras modeling process, a set of rules in the form of a Context-Free Grammar capable of describing the morphosyntactic level of BSL. These rules compose the morphosyntactic model of architecture in the Brazilian Portuguese to Brazilian Sign Language semiautomatic translation process through an avatar.

1 INTRODUCTION

Deaf people in many countries have their own Sign Language (SL) as part of their cultures. The SL are the natural and preferred languages of deaf communities in several countries (Felipe, 2006), and many of them are also used by Indigenous communities in which there are deaf Indians. However, although the SL are also inserted in other cultures predominantly of listeners who use auditory-oral languages, the Deafs still face difficulties that often relate to the processes of teaching and learning and with computational tools that could be an alternative to true school inclusion and citizenship.

For (Ferreira and García, 2018), an accessible environment to deaf students should offer SL interpreted in similar ways to those expressed in the real context. There are some related translation services from Brazilian Portuguese (PT-BR) to BSL available in (Araújo, 2012), (Lima, 2015), (Felipe, 2013) and (Felipe, 2014) - TLibras whose objective was building and intelligent Libras avatar [1], and (De Martino et al., 2017), but they still do not present an adequate treatment for certain grammatical aspects of BSL what leads to the generation of ungrammatical signs.

Built in order to represent Sign Language in real context, the HCI-SL (Garcia et al., 2013), seeks to offer tools that enable the interaction of deaf people with the technology in SL. In addition, the formal computational model for the representation of sign languages (CORE-SL) (Antunes, 2015) stated a set of linguistic-computational requirements for the proper functioning of the architecture. Continuing

[a] https://orcid.org/0000-0003-2012-0491
[b] https://orcid.org/0000-0002-4772-2364
[c] https://orcid.org/0000-0002-6750-5965
[d] https://orcid.org/0000-0001-7098-2597
[e] https://orcid.org/0000-0001-5449-5393

[1]This adjective is being used to differentiate avatars as computational animation, from the avatar that will generate signals from a linguistic knowledge of two languages with equivalence rules from a semi-automatic translation process.

Silva, A., Felipe, T., García, L., Antunes, D. and Guedes, A.
Lex-Libras: Morphosyntactic Model of the Brazilian Sign Language to Support a Context-based Machine Translation Process.
DOI: 10.5220/0010442402670274
In *Proceedings of the 23rd International Conference on Enterprise Information Systems (ICEIS 2021) - Volume 2*, pages 267-274
ISBN: 978-989-758-509-8

this research path, the authors of the present paper built the morphosyntactic rules that will solve the grammatical problems found in related literature. The main problem this research aimed to solve was the context-dependency of the Sign Language. An example of this dependency is the sign for the verb ”fall”, which varies according to the sentence subject.

The aim of this article is therefore to present the Lex-Libras, a rule-based transcription model that establishes the link between phonological and morphosyntactic levels. This model can provide context-based Sign Language generation and, in so, more quality and assertiveness to automatic and semi-automatic translation processes. As a result, in addition to contributing to the translation process that considers the context, Lex-Libras proved the hypothesis of (Antunes, 2015) that it is possible to use the phonological level as an input to the morphological level through a bottom-up approach [2].

2 RELATED WORKS

The aim of the TLibras Project (Felipe, 2003) was to develop a translator from PT-BR to BSL. It also aimed at the creation of phonological rules for the generation of signals. However, the team responsible for creating the avatar and programming the phonological level for the inclusion of signals could not finish this step. Therefore, the team that worked with the morphosyntactic level, created a translation process that was named PUL0 - (Oralizer UNL-LIST from Portuguese) (Felipe, 2003) a system of automated and unidirectional translation from an oral-auditive language, PT-BR, into the linear representation (Libras Script for Translation – LIST) of a gestural-visual language. Yet, the team could not test this translation system. Its goal was to convert a sentence originally produced in PT-BR to a specialized transcription of BSL. That would be signaled by through an intelligent avatar. Such avatar would signal from the phonological configurations, that is, from the five parameters of contrast: hand configuration (HC), location (LOC), hand arrangement (HA), directionalityy-contact (C), and non-manual expressions (NME).

Another initiative was Prodeaf [3], a text and voice translation proposal from PT-BR to BSL. The translation is made by two tools: a mobile application and a plugin for websites. Currently, HandTalk team [4], which was purchased the Prodeaf project. HandTalk was derived from a project called Falibras, a computer system that converts texts and audios into BSL. HandTalk performs machine and semi-automated translation through a website translator and an app.

Another translation proposal was the VLibras (Araújo, 2012), which consists of a set of tools aiming to translate digital content (text, audio, and video). According to the research group working with this application, the system presented syntactic and semantic deficiencies, and, therefore, an automatic translation component was developed, which brings a formal language of syntactic-semantic translation rules and also a set of grammar rules (Lima, 2015). The research related to the morphosyntactic and discursive issues of this avatar-translator is still ongoing.

The research of (De Martino et al., 2017) and (Paiva, 2019) presented a rule-based machine translation system, which has a translator module that analyzes the inputs and then converts them into an intermediate representation, called an “intermediate language” which serves as the input for the animation module.

A morphosyntactic evaluation of these translators was presented by (Silva, 2020). The author, together with a Brazilian Sign Language expert, considered an PT-BR input and observed if the output was generated according to the grammar of BSL. From the analysis of these translators, it is possible to see that, although they have achieved their objectives, they still generate signs or phrases that are considered ungrammatical by the deaf community. We hypothesize, therefore, that such incongruity occurs due to the lack of inclusion of phono-morphosyntactic-semantic-discursive equivalence rules for the units in both languages.

3 RESEARCH CONTEXT

The HCI-SL Architecture (Garcia et al., 2013) aims to provide the development of an integrated environment for the Deafs. Such an environment has been built through methodological strategies and services that can correctly solve, both from the linguistic and computational points of view, issues related to the computational treatment of SL and, consequently, assist in the elimination of the social barrier of access to information and knowledge suffered by the Deafs. The rules presented in Lex-Libras is a new component of the HCI-SL and will integrate its CORE-SL formal

[2] In this approach, we start modeling at the lowest level - the phonological level, to, then, move up to the highest levels of the language.

[3] http://prodeaf.net/

[4] https://www.handtalk.me/br

model(Antunes, 2015), (Antunes et al., 2015).

For the CORE-SL project (Antunes, 2015), the author stated a formal structure for computational representation of the signals. This representation considered all layers of the HCI-SL (Garcia et al., 2013). In addition to the computational approach, linguistic-computational aspects and requirements were designed to provide adequate support to the architecture. Through CORE-SL, it was possible to create a set of sample signals which was more suitable for training and evaluating the processes involved in the translation. That is, the database can be built from the parameters which were specified in the model, enabling a small set of signals that can exemplify several cases of SL through phonemes (Antunes et al., 2015). Figure 1 shows the linguistic model, which was extended with the Lex-Libras component through the research presented in the next sections.

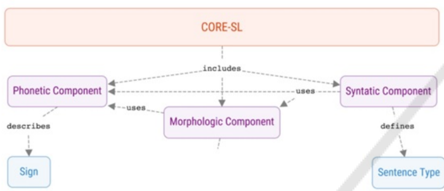

Figure 1: CORE-SL layered model.

4 LEX-LIBRAS

4.1 Methodological Steps

For the development of the present research, we followed these steps: 1. Study of the Morphology of BSL, especially regarding the verbs and their inflection systems. This research aimed at creating a conceptual basis for the computational model. 2. Review of Related Works to understand the context in which the research is carried out. 3.Meetings with a computational linguist specialized in BSL for the Morphosynthatic Evaluation of the PT-BR - BSL translators (Silva, 2020), to verify how the programs generate lexical items from PT-BR input; 4- Definition of the Computational Model in tree format; 5- Meetings with a computational linguist specialized in BSL, for the evaluation of the proposed model; 6- Adjustments in the model and definition in the format of a Context-Free Grammar (CFG).

4.2 The Morphosyntactic Level

In BSL, as in any other natural language, the formal aspects (phonological, morphological and syntactic levels) that relate to the aspects of meaning (semantic and pragmatic) are related or projected. In the SL, the sign corresponds to the lexical item in oral-auditory languages. Within its rules for the creation of new signs, there is a combination of minimal units with meaning (morphemes) with each other, a process that gives rise to new signs.

According to (Felipe, 2006) the five phonological parameters of BSL, hand configuration (HC), location (LOC), hand arrangement (HA), directionality-contact (C), and non-manual features (NME) can be morphemes at the morphosyntactic-discursive level. They can also be classified into two types, as in oral-auditory languages:

- Lexeme: they are the units of lexical meaning, the “roots”. In Brazilian Portuguese verbs: achar, andar, amar (to buy, to walk, to love) - the lexemes are {compr-} {and-} {am-};
- Grammeme [5]: are the designations, prefixes, suffixes, infixes and gender desinence: {-o}, {-a}; verbal endings: thematic vowels: {-a}, {-e}, {-i}; personal-number ending: {-o},{-s}, { }, {-mos}, {-is}, {-m};

Identifying these minimal components is useful for computational processing, both in oral-auditory and gestural-visual languages.

4.3 Verbal Inflection in BSL

Verbal inflection can occur through five processes: inflection for the person of speech, inflection for the verbal aspect (Finau, 2004), inflection for gender, inflection for the locative, and verbs with frequency inflection or modal cases (Felipe, 1998a), (Felipe, 1998b), (Felipe, 2002), (Felipe, 2006), (Felipe, 2007). Inflection for the person of speech, inflection for gender and inflection for the locative were selected for this research. Felipe is the Brazilian Sign Language researcher that adapted the American Sign Language (Stokoe, 1960) system to the Brazilian Sign Language, a language which, differently from the American one, has verbal inflexion.

4.3.1 Person of Speech Inflection

Through displacement, the direction of the movement parameter can make rectilinear or semi-circular

[5]These examples of grammemes, refer to the Portuguese Language.

trajectories, thus inflecting the person of speech. The beginning and the ending of the movement mark the subject and object that agree with the verb. The signaling of this verb is carried out through the movement of hand displacement concerning the person of speech.

The person will be signaled by the active hand in direction to the referred person of speech, being the speaker the first person. In the photo [6], the speaker, looking at the interlocutor, performs a movement directed at this interlocutor, which will be transcribed as ”1sPERGUNTAR2s”, translated: “I ask you”.

4.3.2 Gender Inflection

Gender inflection does not refer to male and female aspects in BSL, but, as in some oral-auditory languages, it is a classification for people, animals, objects, plants, or vehicles (Felipe, 2002). In this case, there is the animacy marking: animate (person or animal) or inanimate (objects, plants or vehicle). In classifier verbs, the HC parameter will be performed as a grammeme. That is, as a mark of a verbal agreement with the subject or object of the sentence, depending on the verb [7]. In the example, the type of object defines the form of the verb ”CAIR” (”fall”).

4.3.3 Locative Inflection

The LOC parameter can indicate a type of inflection. the verb ”COLOCAR” (to put)[8], which in addition to being a classifier verb, also agrees for the locative. The locative agreement mark indicates that the direct object of the sentence will have its final point of displacement in the locative presented prior to the direct object. That is why the order is so relevant for this type of verb. Thus, first is the locative sign, then the object sign, and, finally, the verb sign that will end its articulation at the location.

4.4 Conceptual Model

The unique CORE-SL architecture(Antunes et al., 2015) was the insertion of the signal from the phonological level. Though considering the morphological level, CORE-SL did not include a proposal for the connection with the morphological and syntactic - discursive levels; therefore, the modeling of the Lex-Libras complements this model. Thus, the morphological model receives as input the phonological model, defining a morpheme. Such morpheme will provide input to the syntactic model, and, finally, the syntactic will provide input to the phono-morphosyntactic-discursive model. Figure 2 presents the structure.

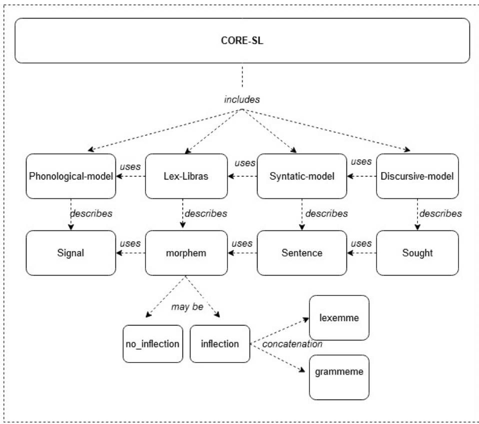

Figure 2: Language model for CORE-SL adapted from the addition of the morphosyntactical rules.

4.5 Formal Model

The conceptual vision of Lex-Libras was modeled in a hierarchical tree format to assist interdisciplinary research (education, linguistics, computing, among others). However, according to (Antunes, 2015), such a view can lead to ambiguous descriptions, as they do not present the rules or methods of how these rules should be created. Thus, after the previous definitions, the rules were formalized through a CFG (Context-Free Grammar), a formalism widely accepted and used for works of this nature (Amaral, 2012), (Antunes, 2015), (De Martino et al., 2017).

- $\beta \rightarrow \alpha$, where:
- α consists of an arbitrary sequence of terminal or non-terminal symbols;
- β consists of a singular non-terminal symbol;

Thus, any occurrence of β during the parsing phase, this non-terminal symbol can be replaced by a α regardless of context. In the context of Computing, the definition of a formal grammar is adopted through an EBNF (Extended Backus-Naur Form) (Wirth, 1977).

4.5.1 Initial Rules

The new CORE-SL structure includes the initial rules, which are defined by the models: phonetic, morphological, syntactic, and discursive The phonological model was the initial one proposed in (Antunes, 2015). Thus, the morphological model

[6] Figure is available in http://bit.ly/3pCkNto

[7] Figure is available in https://bit.ly/3dqFXbr

[8] Figure is available in http://bit.ly/3dvQW3w

was defined as the **”lex-libras”**, which in turn is a **”morpheme”**. MORPHEM: defines that this can be **inflection**, **no_inflection** or even a **NME**. The rule also describes that the three types of morphemes present suspension, categories, identification e case; NO_INFLECTION: defined by a sign (phonological model), the syntactic and discursive components are under construction. Figure 3 presents this proposal.

CORE-SL

core_sl = (phonetic_component, morphological_component, syntactic_component?, discursive_component?);
phonetic_component = identifier, sign

MORPHOLOGICAL RULES

morphological_component = lex_libras;
lex_libras = morpheme;

MORPHEME

morpheme = (infection | no_inflection | nme), suspension, categories, identification, case;
no_inflection = signal;

Figure 3: Initial rules.

4.5.2 Rules for Case and Inflected Verbs

The INFLECTION define that signs that inflect and are formed by the concatenation of a lexeme and a grammeme; the CATEGORIES concept defines the grammatical and morphological categories, and the SUSPENSION concept defines the relationship between the hands (Antunes, 2011) and sequence of a sign (when it is a compound sign). Figure 4 presents these concepts and attributes.

INFLECTION

inflection = (lexeme), (grammar);

CATEGORIES

categories = morfph_cat | gramm_cat;
morfph_cat = 'N (noun or name) | 'V (verb)' | 'Adv (adverb)' | 'Adj (adjective)' | 'Prep (preposition)' | 'num (numeral)' | 'conj (conjunction)' | 'int (interjection)' | 'pron (pronoun)';
gramm_cat = 'lexeme' | 'grammar';
suspension = relation, seq;
relation = 'symmetric' | 'asymmetric';

Figure 4: Rules for inflection e categories.

4.5.3 Grammatical Rules for the Formation of Lexeme and Its Case

The following attributes have been represented for determining the grammatical categories that specify the syntactic-semantic-discursive rules: **case**, **lexeme** and **grammar** CASE: defined by two syntactic and semantic rules, which reflect the thematic roles of a verbal network scheme (Felipe, 1998a), (Felipe, 1998b). LEXEME: defines that it is formed by the concatenation of four BSL parameters; MOVEMENT: defines that the movements can be local or displacement; GRAMMAR: defines the grammemes of verbal inflection addressed in this research. Figure 5 presents this set of rules.

CASE

case = syntactic | semantic
syntactic = 'ag (agent)' | 'Exp (experienced)' | 'Ben (beneficent)' | 'pac (patient)' | 'loc (locative)' | 'theme' | 'origin' | 'goal';
semantic = 'suj (subject)' | 'Obj-1 (direct object)' | 'Obj2 (indirect object)'

LEXEME

lexem = (hand_configuration, movement, location, orientation);
movement = local | displacement;

GRAMMEME

grammar = ps_grammar | gender_grammar | loc_grammar;

Figure 5: Case Rules, lexeme and grammeme.

4.5.4 Grammeme Rules for Verbs with Person of Speech Inflection

The PS_GRAMMAR rules define that the inflection for the person of speech, and it is carried out through directionality, proximity, and person: DIRECTIONALITY: defines the marking of the people involved in the speech (Felipe, 2006); PROXIMITY: defines the beginning and ending of the movement, and both present the same terminal values; PERSON: formed by the values which refer to people involved in the speech. Figure 6 presents these rules.

GRAMMEME RULES FOR VERBS WITH PERSON OF SPEECH INFLECTION

ps_grammar = (directionality, proximity, person)
directionality = 'towards the speaker in a straight line' | 'towards the speaker in an arc 45o to the left' | 'towards the speaker in an arc 45o to the right' | 'from the interlocutor - speaker' | ' another point 'negotiated by speakers' ';
proximity = begin | end;
begin = 'proximal' | 'medial' | 'distal' | 'longitudinal';
end = 'proximal' | 'medial' | 'distal' | 'longitudinal';
person = '1s (1st person singular)' | '2s (2nd person singular)' | | '3s (3rd person singular ')| | '1p (1st person plural)' | | '2p (1st person plural) '| | '3p (1st person plural) ';

Figure 6: Inflection rules for person of speech.

4.5.5 Grammeme Rules for Verbs with Gender Inflection

The rules: GENDER_GRAMMAR, defined by **handshape** and **classifier**; HANDSHAPE, defined by ”right-hand “and” left-hand”, present values which relate to hand configurations [9] that form the classifiers, and CLASSIFIER, which defines the classifier types. Figure 7 presents these rules.

GRAMMEME RULES FOR VERBS WITH GENDER INFLECTION

gender_grammar = (handshape, classifier)

handshape = righ_hand | left_hand;

right_hand = '1' | '2' | '8a' | '13' | '14' | '31' | '32' | '42' | '49' | '51a' | '59a' | '61' | '62';

left_hand = '1' | '2' | '8a' | '13' | '14' | '31' | '32' | '42' | '49' | '51a' | '59a' | '61' | '62';

classifier = animation| no_animation;

animation = 'person' | 'animal';

no_animation = 'object' | 'vehicle';

Figure 7: Inflection rules for gender.

4.5.6 Grammeme Rules for Verbs with Locative Inflection

The LOC_AGREE rules, defined by **localization** and **action**, formalize the location referring to the place in relation to the body where the signal will be articulated. The action represents the performances of a speaker when signaling; PROXIMITY defines the distance in relation to the speaker's body where the grammeme will be articulated; IPSILATERAL_DISPLACEMENT refers to the exact place with marked points in relation to the speaker's body; CENTRAL_LOCALIZATION defines the place in relation to the parts of the speaker's body where the grammeme will be articulated. Figure 8 presents this set of rules.

GRAMMEME RULES FOR VERBS WITH LOCATIVE INFLECTION

loc_grammar= loc_agree
loc_agree = localization | action
action = 'point to' | 'touch' | 'go around' | 'hold tight';

localization = (proximity, displacement-ipsilateral, central-location)
proximity = 'proximal' | 'medial' | 'distal' | 'extended'
ipsilateral_displacement = 'parallel to the medial line' | 'Parallel to the chest' | parallel to the shoulder;
central_localization = 'head (top)' | 'Test' | 'Forehead (lateral)' | 'Eyes' | 'Nose' | 'Mouth' | 'Chin' | 'Sternum' | 'Trunk' | 'Neck' | 'Abdomen' | 'Chest' | 'Shoulders' | 'Waist' | 'Stomach' | 'Legs' | 'Arms' | 'Arm muscles' ;

Figure 8: Inflection rules for the locative.

4.5.7 Rules for Morphological Functions of NME

According (Felipe, 2013) NME can have phonological, morphological, syntactic, semantic, and discursive functions. To represent these expressions, there are the following categories: TYPE, which defines the morphological functions of the NME, and IDENTIFICATION, which defines how the morphemes will be written. Figure 9 shows these categories.

NON-MANUAL EXPRESSION

nme = type, specific (face, head_specific) | trunk_specif;

type = 'morpheme-adjective' | morpheme-adverbial '| 'deixis' | 'morpheme degree of adjective'

specific = 'sadness' | 'Joy' | 'Happiness' | 'Smiling' | 'Angry' | 'Doubt' | 'Indifference' | 'Discomfort' | 'Affective (cry with emotion)' | 'Affective (crying with sadness)' | 'Affective' | 'Cry' | 'Anger' | 'Interrogative' | 'Confirmation (yes)' | 'Confirmation (no)' | 'Affirmative' | 'Exclamatory' | 'Negative' | 'Nod' | 'Disappointment' | 'Concern' | ' fear'

IDENTIFICATION

identification = (character | integer) +

character = [a-Z];

integer = [0-9];

Figure 9: Rules for NME.

5 CONCLUSIONS AND FUTURE WORK

The Lex-Libras just presented is part of the CORE-SL and one of the bases for the HCI-SL architecture. Its main contribution consists of the specification of a model that can be computerized for an automatic or semi-automatic translation. Such a model aims to serve as input to the "intelligent avatar". In addition, it also works as an input to the next linguistic level, the syntactic one.

Lex-Libras has been theoretically validated by a Libras computational linguist, as it includes the Morphology Theory of BSL, thus adding linguistic knowledge to the construction of the model. Lex-Libras will allow to computational vision (recognition and transcription of signals for the morphosyntactic structure), processing of natural sign languages, and synthesis (3D avatars).

Additionally, Lex-Libras proved the hypothesis raised by (Antunes, 2015), that it is possible to describe the morphosyntactic level by means of the CORE-SL gestural-spatial parameters. In this way, Lex-Libras is considered to be partially integrated into the CORE-SL formal model. Never-the-less, the full integration will only take place when the remaining Libras linguistic levels are specified and an experimental framework enables the intelligent avatar together with the overall testing.

To synthesize the proposal, an abstraction of the use of the Lex-Libras in HCI-SL will involve (1) The user, who interacts with the Architecture

[9]The values of the hand settings were consulted (Felipe, 2002) BSL Dictionary Version 2.0. Available at: http://www.ines.gov.br/dicionario-delibras/main site / libras.htm

through the application layer, selects a data unit, such as the sign generator module (Avatar); (2) The user enters a particular sign; (3) The CORE-SL receives the request and recognizes that this signal is composed by an inflection process; (4) The CORE-SL, then, forwards the request to Lex-Libras, which is responsible for the morphological module, and looks for the sign; (5) The Lex-Libras supplies the adequate descriptions in JSON format for the application use, as shown in 10 below.

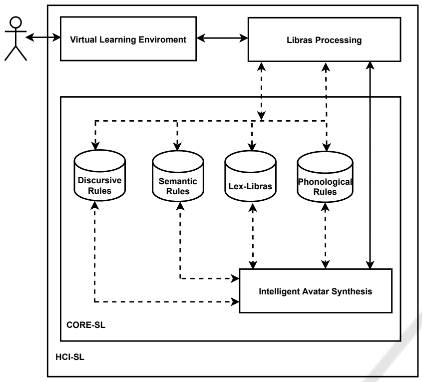

Figure 10: Abstract of a usage of Lex-Libras in HCI-SL.

The phonological model of the CORE-SL is, hypothetically, universal regarding the SL, although they may vary regarding hand configurations and the specificity of the phonological rules for each language, as it also happens with the oral-auditory languages. These languages are described from the international phonetic alphabet, although each language uses specific phonemes within these possibilities.

Regarding the hypothesis of the CORE-SL formal model being universal within sign languages, this paper has brought, as an additional contribution, the methodological steps to build the morphosyntactic rules for other sign languages. The research described in this paper is innovative as a multi, inter, and transdisciplinary work. This dialogue among disciplines happens in the evaluation of related results and in the construction of the rules of the different but interconnected linguistic levels.

The research reported in this article took a bottom-up approach, starting from the phonological description to the morphosyntactic one, through the Lex-Libras, which is a formalism to generate lexical items and organize them in the form of lexemes and grammemes. This was based on the 5 phonological parameters and used the phonological model as an input to the morphological level.

The construction of the formal model with all the internal and external properties and initiated by the phonological level allowed the cataloging of signs in a textual standard representation. Thus, it provides for the extension of the rules of the successive linguistic levels, where the previous level functions as input to the immediately subsequent level (Antunes, 2011), (Antunes et al., 2015), (Silva, 2020).

The architecture of hypothesis, HCI-SL, considers the future generation of an “intelligent avatar” to interpret the registered signal. Deaf communities have an intuitive knowledge of their language. As in any translation endeavor, in the construction of rules of equivalence, it is necessary to count on the contribution and the perspective of experts from specific areas. It is also crucial to consider the participation of deaf bilinguals who master such specific areas of knowledge. Information on sentence generation problems by existing translators is already available (Silva, 2020).

The referred knowledge was produced by the perception of their own deaf communities that, yet, cannot determine which rule and at which specific level the problems occur. Nonetheless, deaf communities will benefit from this work as an input to future participatory workshops. On the other hand, listeners who use the translate engines, when signaling the searched sentences, will be committing the same grammatical errors of the avatars, since they are unaware of the grammatical rules of BSL. In this way, Lex-Libras can add more quality and assertiveness to translation in the future and, thus, provide better accessibility and communicability resources to the Deafs by providing access to information and knowledge in BSL.

In this research, similar works that focus on other SL of other countries were also raised. Since they have their own grammatical-discursive rules, we did not carry any analysis of generated sentences in different SL. However, even in different SL, the HCI-SL Architecture can still be used, modifying the rules of its levels that, in this paper, are specific to BSL. The extension of this work based on a *corpus*[10] , like those presented in (Iatskiu, 2019), (Silva, 2020), built in a formal computerized way, is part of future work, as is the module of generation by an avatar. Additionally, the results of the research reported here will be appropriate by games and educational applications already developed (Canteri et al., 2019)

[10]For a more in-depth study involving PLN, it is necessary to use a linguistic corpus, which is the set of written texts and oral records in a given language and that serves as a basis for analysis. This approach is currently known as Corpus Linguistics, since the interpretation of the linguistic phenomenon was based on the observation of data from corpus and not on intuitions of the analyst (Felipe, 2007).

and under development for deaf communities.

REFERENCES

Amaral, W. (2012). *Sistema de Transcrição da Língua Brasileira de Sinais Voltado a Produção de Conteúdo Sinalizado por Avatares 3D*. PhD thesis, Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação, Campinas, SP.

Antunes, D. (2011). Um modelo de descricao computacional da fonologia da lingua de sinais brasileira. Master's thesis, Programa de Pos-Gradução em Informática, Curitiba- PR.

Antunes, D. (2015). *Proposta de um Modelo Computacional para Representação de Sinais em uma Arquitetura de Serviços HCI-SL para Língua de Sinais*. PhD thesis, Programa de Pos-Graduação em Informática do Setor de Ciencias Exatas da Universidade Federal do Paraná, Curitiba - PR.

Antunes, D. R., Guedes, A. L. P., and García, L. S. (2015). A context-based collaborative framework to build sign language databases by real users. In Antona, M. and Stephanidis, C., editors, *Universal Access in Human-Computer Interaction. Access to Interaction*, pages 327–338, Cham. Springer International Publishing.

Araújo, T. (2012). *Uma Solução para Geração Automática de Trilhas em Língua Brasileira de Sinais em Conteúdos Multimídia*. PhD thesis, Tese de doutorado. Universidade Federal do Rio Grande do Norte, Natal, Brasil, Rio Grande do Norte, Natal.

Canteri, R., Sánchez, L., Felipe, T., Galvão, L., and Antunes, D. (2019). Conceptual framework to support a web authoring tool of educational games for deaf children.

De Martino, M., Silva, I.and Bolognini, C., Costa, P., Kumada, K., Coradine, L., Brito, P., do Amaral, W., Benetti, A., Poeta, E., Angare, L., Ferreira, M., and De Conti, D. (2017). Signing avatars: making education more inclusive. *Universal Access in the Information Society*, 16(3):793–808.

Felipe, T. (1998a). *A Relação Sintático-Semântica dos Verbos Língua Brasileira de Sinais (LIBRAS. Vol I.* PhD thesis, Programa de Pós-Graduação da Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ.

Felipe, T. (1998b). *A Relação Sintático-Semântica dos Verbos Língua Brasileira de Sinais (LIBRAS). Vol II.* PhD thesis, Programa de Pós-Graduação da Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro - RJ.

Felipe, T. (2002). Sistema de flexão verbal na libras: Os classificadores enquanto marcadores de flexão de gênero. *Anais do Congresso Surdez e Pós-Modernidade: Novos rumos para a educação brasileira*, pages 37–58.

Felipe, T. (2003). Projeto Tradutor Libras. Technical report, FENEIS - Federação Nacional de Educação e Integração dos Surdos.

Felipe, T. (2006). O processo de formação de palavras na libras. *ETD – Educação Temática Digital, Campinas,*, pages 199–216.

Felipe, T. (2007). Descrição da língua de sinais: desafios teóricos e práticos 1.

Felipe, T. (2013). O discurso verbo-visual na língua brasileira de sinais – Libras / The verbal-visual discourse in Brazilian Sign Language – Libras. *Bakhtiniana. Revista de Estudos do Discurso, url=http://ref.scielo.org/cjh56s*, 8(2):67–89.

Felipe, T. (2014). *Banco de Dados e Sistemas de Transcrição para as Línguas de Sinais, Instrumentos Linguísticos: usos e atualizações, Orgs: Balbas and A. Caldas*, pages 155–188. Cartolina Editora.

Ferreira, M. and García, L. (2018). Requirements for avatar in virtual environment of support learning in the literacy of deaf people in portuguese mediated by libras. In *Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems*, IHC 2018, New York, NY, USA. Association for Computing Machinery.

Finau, R. (2004). *Os sinais de tempo e aspecto na Libras*. PhD thesis, Programa de Pós-Graduação em Letras da Universidade Federal do Paraná, Curitiba-PR.

Garcia, L., Guimarães, C., Antunes, D., and Fernandes, S. (2013). HCI architecture for deaf communities cultural inclusion and citizenship. *ICEIS 2013 - Proceedings of the 15th International Conference on Enterprise Information Systems*, 3(October 2017):126–133.

Iatskiu, C. (2019). *Core-SL-SW-generator: gerador automático da escrita da libras a partir de um modelo de especificação formal dos sinais*. PhD thesis, Programa de Pos-Graduação em Informática do Setor de Ciencias Exatas da Universidade Federal do Paraná, Curitiba - PR.

Lima, M. (2015). Tradução automática com adequação sintático-semântica para libras. Master's thesis, Centro de Informática da Universidade Federal da Paraíba, João Pessoa - PB.

Paiva, F. (2019). Faculdade de Engenharia Elétrica e de Computação Tradução Automática de Português Brasileiro para Libras e Análise de Processos de Intensificação Tradução Automática de Português Brasileiro para Libras e Análise de Processos de Intensificação.

Silva, A. (2020). Descrição das unidades morfossintáticas da libras através dos parâmetros fonológicos para o processo de tradução automática por um avatar. Master's thesis, Programa de Pós Graduação em Informática - Universidade Federal do Paraná, Curitiba - PR.

Stokoe, W. (1960). *Sign Language Structure. An Outline of the Visual Communication Systems of the American Deaf.* SpringLinstok.

Wirth, N. (1977). What can we do about the unnecessary diversity of notation for syntactic definitions? *Communications of the ACM*, 20(11):822–823.