

# BSL-Lex: Modelagem Morfossintático-Discursiva da Libras para o Processo de Tradução Automática Baseado em Contexto

Antônio Marcos C. Silva<sup>1</sup>

<sup>1</sup>Departamento de Engenharia – Universidade do Estado do Pará (UEPA)  
Belém – Pará – PA – Brazil

**Abstract.** *This paper presents itself as a doctoral proposal: The BSL-Lex, which will be a set of rules in the form of Context-Free Grammar (CFG), capable of describing the syntactic and discursive level from Libras (Brazilian Sign Language - BSB). These rules will compose the linguistic model of architecture to a semi-automatic Portuguese-Libras translation process through an avatar.*

**Resumo.** *Este projeto de doutorado apresenta como proposta o BSL-Lex, um conjunto de regras na forma de uma Gramática Livre de Contexto (GLC), capaz de descrever o nível sintático e discursivo da Libras. Essas regras irão compor a base de dados linguística da Arquitetura para o processo de tradução semiautomática Português-Libras por um avatar.*

## 1. Introdução

Os Surdos em muitos países possuem sua própria Língua de Sinais (LS), identidades e culturas. As LS são as línguas naturais e preferenciais das comunidades surdas em vários países [Felipe 2006] e são utilizadas também por comunidades indígenas em que há índios surdos. No entanto, embora as LS estejam inseridas também em outras culturas majoritariamente de ouvintes que utilizam línguas orais-auditivas, os Surdos ainda enfrentam dificuldades: com relação aos processos de ensino e aprendizagem e com ferramentas computacionais que poderiam ser alternativas de apoio à inclusão escolar e à cidadania.

Para [Ferreira and García 2018], um ambiente acessível aos estudantes surdos deve oferecer LS interpretada de forma semelhante à maneira de expressão no contexto real. Os serviços de tradução automática de Libras para o Português do Brasil (PT-BR) revisados [Volpe et al. 2003], [Araújo 2012], [Lima 2015], [De Martino et al. 2017], [Paiva 2019] ainda não apresentam um tratamento adequado para certos aspectos gramaticais da Libras o que leva à geração de sinais agramaticais. Na Libras existe um grupo de verbos que flexionam, no nível morfológico, como os verbos: "cair", "perguntar" e "colocar". Em minha pesquisa de mestrado [Silva 2020], [Silva et al. 2021], (No prelo) apresentei o Lex-Libras, um conjunto de regras que trataram alguns dos fenômenos de flexão no nível morfossintático da Libras. No entanto, ainda existem outros fenômenos morfossintáticos, sintáticos e discursivos, entre os quais as expressões não-manuais que marcam frases negativas, interrogativas ou afirmativas, que necessitam ser investigados e modelados para a implementação computacional.

A Arquitetura de Interação Humano-Computador em Língua de Sinais (HCI-SL) [Garcia et al. 2013] busca oferecer ferramentas que possibilitem a interação das

peças surdas com a tecnologia em LS. Além desta, o Modelo de Representação Computacional para Língua de Sinais (CORE-SL) [Antunes 2015] propõe um conjunto de requisitos linguístico-computacionais necessários ao funcionamento adequado da Arquitetura. Este projeto de doutorado, propõe estabelecer a ligação entre os níveis fono-morfo-sintático-discursivos, por meio de um conjunto de regras modeladas em uma Gramática Livre de Contexto (GLC). O modelo comporá o conjunto de bases conceituais e formais da HCI-SL, e irá proporcionar a geração de sinais e de frases na Libras corretos do ponto de vista gramatical e, assim, potencialmente passíveis de apropriação pelas comunidades surdas. Assim, este estudo tem como uma das principais motivações dar continuidade a minha pesquisa de mestrado [Silva 2020], [Silva et al. 2021], (No prelo).

### **1.1. Objetivos, Questões de Pesquisa e Passos Metodológicos**

O objetivo geral deste trabalho é formalizar, a partir do conhecimento construído para a modelagem formal e computacional do subconjunto de fenômenos de flexão tratados em [Silva 2020], [Silva et al. 2021] (No prelo), tratar os fenômenos morfo-sintáticos restantes, assim como os fenômenos sintático, semântico e discursivos da Libras para sua incorporação ao CORE-SL na HCI-SL. Cumprindo esse objetivo geral, esta pesquisa se propõe a responder às seguintes questões:

(I) É possível usar as unidades morfossintáticas do Lex-Libras como entrada para os níveis sintático e discursivo?

(II) Como são formadas as estruturas linguísticas da Libras para os níveis sintático e discursivo?

(III) Como formalizar estes níveis linguísticos de forma a serem passíveis de incorporação ao modelo CORE-SL?

(IV) Como atualizar o modelo do CORE-SL que inicialmente não previu o nível discursivo?

### **1.2. Passos Metodológicos**

Para a realização da pesquisa serão seguidos os passos:

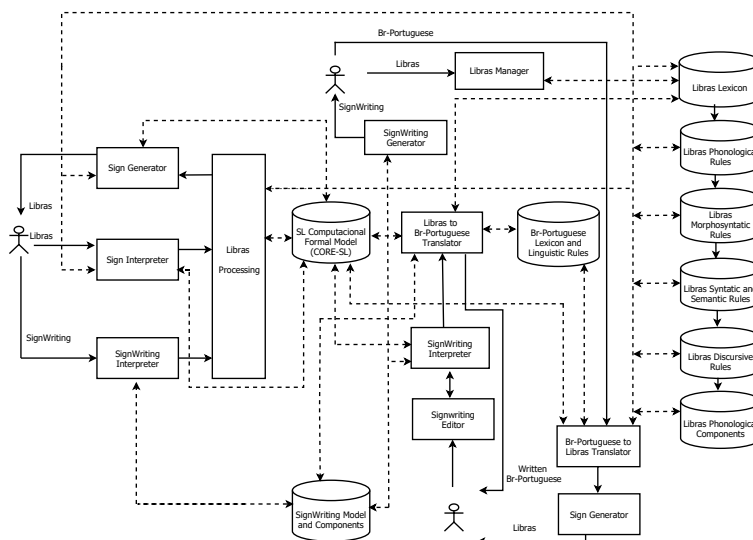
1. Extensão do conjunto de aspectos morfossintáticos apresentados em [Silva 2020], [Silva et al. 2021] (No prelo);
2. Investigação sobre os aspectos sintáticos e discursivos da Libras, com o intuito de criar uma base conceitual para o modelo computacional;
3. Estudo dos trabalhos correlatos, para compreender o contexto em que a pesquisa se realiza;
4. Avaliação linguística dos tradutores PT-BR - Libras, com o intuito de verificar como os programas geram as sentenças, a partir de uma entrada em PT-BR;
5. Modelagem em formato de árvore;
6. Reuniões com uma linguista computacional da Libras, para avaliação do modelo proposto;
7. Oficina com comunidade surda, para validação e avaliação do modelo;
8. Ajustes no modelo e definição em formato de uma GLC;

Para a apresentação desta proposta, o restante do texto traz as seguintes seções: apresentação do contexto da pesquisa (2), trabalhos correlatos (3), apresentação do modelo conceitual e formal (4), os resultados parciais (5) e as considerações finais (6).

## 2. CONTEXTO DE INSERÇÃO

### 2.1. A Arquitetura HCI-SL

A Arquitetura HCI-SL [Garcia et al. 2013] tem como objetivo proporcionar o desenvolvimento de um ambiente integrado, com estratégias metodológicas e serviços que sejam capazes de resolver de maneira correta tanto do ponto de vista linguístico como computacional, questões relacionadas ao tratamento computacional das LS e, consequentemente, auxiliar na eliminação da barreira social de acesso a informação e ao conhecimento sofrida pelos Surdos. As regras resultantes da pesquisa de doutorado integrarão o CORE-SL [Antunes 2015], [Antunes et al. 2015]. A Figura 1 apresenta a Arquitetura.

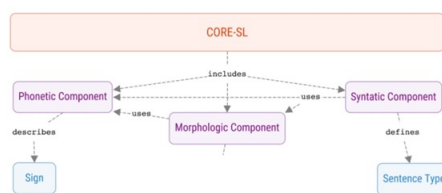


**Figura 1. Abstração HCI-SL**

Fonte:[Garcia et al. 2013]

### 2.2. O CORE-SL

Para o projeto do CORE-SL [Antunes 2015], foi proposta uma estrutura formal para representação computacional dos sinais. Além da abordagem computacional, foram projetados aspectos e requisitos linguísticos-computacionais para dar suporte adequado à Arquitetura. Por meio do CORE-SL foi possível criar um conjunto de sinais de amostra mais adequadas para treinar e avaliar os processos envolvidos na tradução, ou seja, o banco dados pode ser construído a partir dos parâmetros especificados no modelo, possibilitando um conjunto pequeno de sinais que consegue exemplificar vários casos da SL por meio de fonemas [Antunes et al. 2015]. A Figura 2 apresenta o modelo linguístico, que foi atualizado em [Silva 2020], [Silva et al. 2021] (No prelo) e será apresentado nas próximas seções.



**Figura 2. Modelo linguístico**

Fonte:[Antunes 2015]

### 3. TRABALHOS CORRELATOS

A proposta do Projeto TLibras [Felipe 2003] foi desenvolver um tradutor do PT-BR para a Libras e a criação das regras fonológicas para a geração de sinais, no entanto, a equipe responsável pela criação do avatar e da programação do nível fonológico para a inclusão de sinais, não conseguiu finalizar essa etapa e, por isso, a equipe que trabalhou com a elaboração do processo de tradução criou o PULO – (Portuguese UNL-LIST deOralizer) [Volpe et al. 2003], [Felipe 2003] um sistema de tradução automática unidirecional de uma língua oral-auditiva, o PT-BR, para a representação linear (Libras Script for Translation – LIST) de uma língua gestual-visual, contudo a equipe não conseguiu testar o programa criado, cujo objetivo seria converter uma sentença originalmente produzida em PT-BR para uma transcrição especializada de Libras que seria sinalizada através de um “avatar inteligente” que sinalizaria um sinal a partir de suas configurações fonológicas, ou seja, dos cinco parâmetros: configuração de mão (CM), ponto de articulação (PA), direcionalidade (DIR), movimento (MOV) e expressões faciais (ENM)

Uma outra iniciativa foi o Prodeaf<sup>1</sup>, uma proposta de tradução de texto e voz do PT-BR para a Libras. A tradução é feita por duas ferramentas: um aplicativo móvel e um plugin para sites. Atualmente, esse projeto foi comprado pela equipe do HandTalk<sup>2</sup> que foi criado a partir de um projeto chamado Falibras, sistema computacional que converte textos e áudio em Libras. O HandTalk realiza tradução digital e automática para Libras, por meio de um tradutor de sites e de um aplicativo.

Outra proposta de tradutor é o VLibras [Araújo 2012], consiste em um conjunto de ferramentas, que tem por finalidade traduzir conteúdos digitais (texto, áudio e vídeo). Segundo o Grupo de Pesquisa que trabalha com este aplicativo, o sistema apresentava deficiências sintática e semântica e, por isso, foi elaborado um componente de tradução automática, que traz uma linguagem formal de descrição de regras de tradução sintático-semântica e uma gramática [Lima 2015]. As pesquisas com relação às questões morfosintáticas e discursivas desse tradutor através de avatar ainda estão em processo de desenvolvimento.

As pesquisas de [De Martino et al. 2017] e [Paiva 2019] apresentaram um sistema de tradução automática baseado em regras, que possui um módulo tradutor que analisa as entradas e então as converte em uma representação intermediária, chamada de “Linguagem Intermediária”, que serve de entrada para o módulo de animação.

Uma avaliação morfosintática desses tradutores foi apresentada por [Silva 2020], onde dada uma entrada em PT-BR, é observada se a saída é gerada de acordo com

<sup>1</sup><http://prodeaf.net/>

<sup>2</sup><https://www.handtalk.me/br>

a gramática da Libras. A partir da análise desses tradutores foi possível constatar que, embora tenham alcançado os seus objetivos, ainda geram sinais ou frases que são consideradas agramaticais pelos Surdos. Por isso, nossa hipótese é que tal incongruência ocorra devido a não inclusão de regras fono-morfossintático-semântico-discursivas para a equivalência das unidades em ambas as línguas e, por isso, esta pesquisa se propõe a abordar essas questões.

## **4. O LEX-LIBRAS**

### **4.1. Fenômenos morfossintáticos tratados na dissertação**

Na Libras, como em toda língua natural, se relacionam ou projetam os aspectos formais (níveis fonológico, morfológico e sintático) que inter-relacionam-se aos aspectos do significado (semântico e pragmático). Nas LS, o sinal corresponde ao item lexical em línguas orais-auditivas e em suas regras para a criação de novos sinais acontecem também a partir unidades mínimas com significado (morfemas) que são combinadas entre si, dando origem a novos sinais.

Para [Felipe 2006] os cinco parâmetros fonológicos da Libras: configuração de mão (CM), ponto de articulação (PA), direcionalidade (Dir), movimento (MOV) e expressão não manual(ENM) podem ser morfemas no nível morfossintático-discursivo, podem ser de dois tipos, como nas línguas orais-auditiva:

- Lexema: são as raízes. Em Português nos verbos: comprar, andar, amar (to buy, to walk, to love) - os lexemas são: {compr-} {and-} {am-}
- gramema: são as desinências, prefixos, sufixos e infixos, ou seja: desinência de gênero: {-o}, {-a}; desinências verbais: vogais temáticas: {-a}, {-e}, {-i}; desinência número-pessoal: {-o}, {-s}, { $\emptyset$ }, {-mos}, {-is}, {-m}.

Identificar esses componentes mínimos é útil para o processamento computacional, tanto em línguas orais-auditivas quanto em gestuais- visuais.

### **4.2. Flexão verbal na Libras**

A flexão verbal pode ocorrer por meio de 5 processos: flexão para a pessoa do discurso, flexão para o aspecto verbal [Finau 2004], flexão para gênero, a flexão para o locativo e os verbos com flexão de frequência ou de casos modais, [Felipe 2002], [Felipe 2006], [Felipe 2007]. Para esta pesquisa selecionou-se três processos de flexão.

#### **4.2.1. Flexão para a pessoa do discurso**

Por meio do deslocamento, o parâmetro direção do movimento pode fazer trajetórias retilíneas ou em semi-círculo, fazendo assim a flexão com as pessoas do discurso. O início e fim do movimento de deslocamento marcam o sujeito e o objeto que concordam com o verbo. A sinalização desse verbo é realizada através do movimento de deslocamento da mão considerada ativa, em relação a pessoa do discurso, sendo o sinalizador a primeira pessoa do discurso (sujeito). O locutor, olhando para o interlocutor, realiza um movimento direcionado a esse interlocutor (objeto), que será transcrito "1sPERGUNTAR2s"<sup>3</sup>, traduzido: Eu pergunto para você.

---

<sup>3</sup>Figura disponível em: <http://bit.ly/3pCkNto>

### 4.2.2. Flexão para gênero

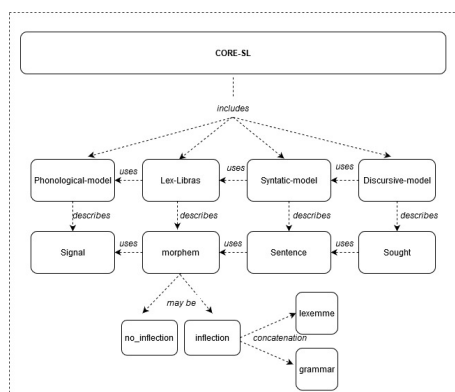
A flexão de gênero não se refere aos valores masculino e feminino na Libras mas, como em algumas línguas orais-auditivas, é uma classificação para pessoas, animais, objetos, plantas ou veículos [Felipe 2002]. Neste caso, há a marcação de concordância de gênero animado (pessoa ou animal) ou inanimado (objeto ou veículo). Nos verbos classificadores, o parâmetro CM vai se realizar como um gramema, enquanto marca de concordância verbal com o sujeito ou o objeto da frase, dependendo verbo. Nos três exemplos com o verbo CAIR<sup>4</sup>, o parâmetro CM muda de acordo com o sujeito da frase.

### 4.2.3. Flexão para o locativo:

O parâmetro PA pode indicar um tipo de flexão. O verbo "COLOCAR"<sup>5</sup>, que além de ser um verbo classificador, também possui concordância para o locativo. A marca de concordância de locativo indica que o objeto direto da frase terá seu ponto final de deslocamento no locativo apresentado anteriormente ao objeto direto, por isso, a importância da ordem para esse tipo de verbo. Ou seja: primeiro o sinal do locativo, depois o sinal do objeto e, por último o sinal do verbo que finalizará sua articulação no ponto de articulação do locativo.

## 4.3. Modelo Conceitual

A proposta arquitetônica do CORE-SL [Antunes 2015] foi a inserção do sinal a partir de nível fonológico, não apresentando uma proposta para a ligação com os níveis morfológico e sintático-discursivo, por isso, a modelagem do Lex-Libras complementa essa proposta. Assim, o modelo morfológico recebe como entrada o modelo fonológico, definindo um morfema, este por sua vez fornecerá entradas para modelo sintático e, finalmente, o sintático fornecerá as entradas ao modelo fono-morfossintático-discursivo. A Figura 3 apresenta a estrutura:



**Figura 3. Atualização do Modelo Linguístico.**

Fonte: [Silva et al. 2021], no prelo.

<sup>4</sup>Figura disponível em: <https://bit.ly/3dqFXbr>

<sup>5</sup>Disponível em: <http://bit.ly/3dvQW3w>

#### 4.4. O modelo formal

A visão conceitual do Lex-Libras foi modelada em formato de árvore hierárquica para auxiliar as pesquisas interdisciplinares (educação, linguística, computação, entre outros). Contudo, de acordo com [Antunes 2015], tal visão pode acarretar em descrições ambíguas, pois não apresentam as regras ou os métodos de como essas regras devem ser criadas. Assim, após as definições anteriores, as regras foram formalizadas por meio de uma GLC (Gramática Livre de Contexto), formalismo amplamente aceito e utilizado para trabalhos desta natureza [Amaral 2012], [Antunes 2015], [De Martino et al. 2017], [Paiva 2019]. A gramática é formada por um símbolo inicial, um conjunto de variáveis (símbolos não-terminais), de palavras (símbolos terminais) e de regras de produção que possuem a forma:

- $\beta \rightarrow \alpha$ , onde:
- $\alpha$  consiste de uma sequência arbitrária de símbolos terminais ou não terminais;
- $\beta$  consiste de um símbolo não terminal;

Dessa forma, qualquer ocorrência de  $\beta$  durante a fase de *parsing*, este símbolo não-terminal pode ser substituído por um símbolo  $\alpha$  independente do contexto. No contexto da Computação, adota-se a definição de uma gramática formal por meio de uma EBNF (*Extended Backus-Naur Form*) [Wirth 1977].

##### 4.4.1. Regras iniciais

O novo organograma para o CORE-SL inclui as regras iniciais, onde é definido pelos modelos: fonético, morfológico, sintático e discursivo. O modelo fonológico, foi o inicial proposto em [Antunes 2015]. Assim, o modelo morfológico foi definido como o **”lex-libras”**, que por sua vez é um **”morphem”**. MORFHEM: define que este pode ser **inflection**, **no\_inflection** ou ainda uma **nme**. A regra também descreve que os três tipos de morfemas possuem **suspension**, **categories**, **identification** e **case**; NO\_INFLECTION: definido por um sinal (modelo fonológico), os componentes sintático e discursivo estão em elaboração. A Figura 4 apresenta essa proposta.

```

CORE-SL

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

MORPHOLOGICAL RULES

morphological_component = lex_libras;
lex_libras = morpheme;

MORPHEME

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

**Figura 4. Regras iniciais.**

Fonte: [Silva et al. 2021], no prelo.

### 4.4.2. Regras para caso e verbos com flexão

Os conceitos INFLECTION e CATEGORIES-INFLECTION definem os sinais que flexionam e são formados pela concatenação de um lexema e um gramema; e o conceito CATEGORIES define as categorias gramatical e morfológica e o conceito SUSPENSION define a relação entre as mãos [Antunes 2011] e sequência de um sinal (quando este é composto). A Figura 5, apresenta esses conceitos e atributos.

<b>INFLECTION</b>
<code>inflection = (lexeme), (grammar);</code>
<b>CATEGORIES</b>
<code>categories = morph_cat   gramm_cat;</code>
<code>morph_cat = 'N (noun or name)'   'V (verb)'   'Adv (adverb)'  </code>
<code>'Adj (adjective)'   'Prep (preposition)'   'num (numeral)'   'conj (conjunction)'   'int</code>
<code>(interjection)'   'pron (pronoun)';</code>
<code>gramm_cat = 'lexeme'   'grammar';</code>
<code>suspension = relation, seq;</code>
<code>relation = 'symmetric'   'asymmetric';</code>

Figura 5. Regras para inflection e categories.

Fonte: [Silva et al. 2021], no prelo.

### 4.4.3. Regras gramaticais para a formação do lexema e seu caso

Para a especificação das categorias gramaticais que especificam as regras sintático-semântico-discursiva foram representados os seguintes atributos: **case**, **lexema** e **grammar**. CASE: definido por duas regras **syntactic** e **semantic**, que refletem os papéis temáticos de um esquema de rede verbal. LEXEMA: define que este é formado pela concatenação de quatro parâmetros da Libras; MOVIMENT: define que os movimentos podem ser locais ou de deslocamento; GRAMMAR: define os gramemas de flexão verbal abordados nesta pesquisa. A Figura 6 apresenta esse conjunto de regras.

<b>CASE</b>
<code>case = syntactic   semantic</code>
<code>syntactic = 'ag (agent)'   'Exp (experienced)'   'Ben (beneficent)'   'pac (patient)'   'loc</code>
<code>(locative)'   'theme'   'origin'   'goal';</code>
<code>semantic = 'suj (subject)'   'Obj-1 (direct object)'   'Obj2 (indirect object)'</code>
<b>LEXEME</b>
<code>lexem = (hand_configuration, movement, location, orientation);</code>
<code>movement = local   displacement;</code>
<b>GRAMMAR</b>
<code>grammar = ps_grammar   gender_grammar   loc_grammar;</code>

Figura 6. Regras para case, lexeme e grammars.

Fonte: [Silva et al. 2021], no prelo.



#### 4.4.4. Regras do gramema dos verbos com flexão para a pessoa do discurso

As regras PS\_GRAMMAR definem a flexão para pessoa do discurso é realizada através da directionality, proximity e person: DIRECTIONALITY: define a marcação das pessoas envolvidas no discurso [Felipe 2006]; PROXIMITY: define o início e fim do movimento, em ambas os valores terminais são os mesmos; PERSON: formado pelos valores referentes às pessoas envolvidas no discurso. A Figura 7 apresenta essas regras.

---

**GRAMMAR INFLECTED VERBS FOR THE SPEECH PERSON**

---

```
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)';
```

---

**Figura 7. Regras flexão para pessoa do discurso.**

Fonte: [Silva et al. 2021], no prelo.

#### 4.4.5. Regras do gramema dos verbos com flexão de gênero

As regras: GRAMGEN, definido por **handshape** e **classifier**; HANDSHAPE, definido por “right hand” e “left hand”, possuem os valores relativos às configurações de mão <sup>6</sup> que formam os classificadores e CLASSIFIER define os tipos de classificadores. A Figura 8 apresenta essas regras.

---

**GRAMMAR INFLECTED VERBS FOR THE GENDER**

---

```
gender_grammar = (handshape, classifier)
handshape = right_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';
```

---

**Figura 8. Regras flexão para gênero.**

Fonte: [Silva et al. 2021], no prelo.

#### 4.4.6. Regras do gramema dos verbos com flexão para o locativo

As regras LOC\_AGREE, definidas por **localization** e **action**, formaliza a localização referente ao local em relação ao corpo onde o sinal será articulado. A ação representa as execuções de um falante ao sinalizar; PROXIMITY define a distância em relação ao

<sup>6</sup>Os valores das configurações de mão foram consultados [Felipe 2002] Dicionário da Libras Versão 2.0. Disponível em: [http://www.ines.gov.br/dicionario-delibras/main site/libras.htm](http://www.ines.gov.br/dicionario-delibras/main%20site/libras.htm)

corpo do falante onde o gramema será articulado; IPSILATERAL\_ DISPLACEMENT refere-se ao local exato com pontos marcados em relação ao corpo do falante; CENTRAL\_ LOCALIZATION define o local em relação às partes do corpo do falante onde o gramema será articulado. A Figura 9 apresenta esse conjunto de regras.

---

**GRAMMAR INFLECTED VERBS FOR THE LOCATIVE**

**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' ;

---

**Figura 9. Regras flexão para o locativo.**

Fonte: [Silva et al. 2021], no prelo.

#### 4.4.7. Regras para funções morfológicas das ENM

Segundo [Felipe 2013] as nme podem ter funções fonológicas, morfológicas, sintático, semânticas e discursiva. para representar essas expressões, serão utilizadas as categorias TYPE, que define as funções morfológicas das ENM e IDENTIFICATION, que define como serão escritos os morfemas. A Figura 10 apresenta estas categorias.

---

**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];

---

**Figura 10. Regras para ENM.**

Fonte: [Silva et al. 2021], no prelo.

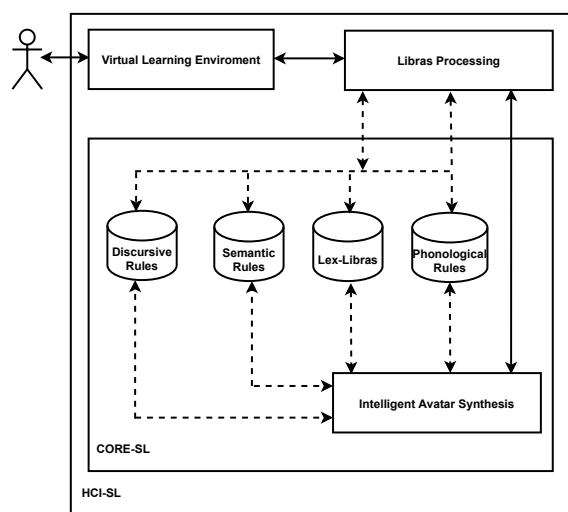
## 5. RESULTADOS PARCIAIS

O Lex-Libras [Silva 2020], [Silva et al. 2021] (No prelo) apresenta-se como uma das bases conceituais para o arcabouço linguístico-computacional do CORE-SL e da HCI-SL. A principal contribuição da minha dissertação consistiu na especificação de um modelo passível de tratamento computacional para a tradução automática e semi-automática, de entrada ao próximo nível linguístico, o Sintático.

O Lex-Libras dará suporte para as ferramentas que operam na camada de Aplicação da HCI-SL, como: Visão Computacional (reconhecimento e transcrição

de sinais para estrutura morfossintática), Processamento de Línguas Naturais e de Módulos de Síntese (Avatares 3D). O Lex-Libras provou a hipótese levantada por [Antunes 2015], de que é possível descrever o nível morfo-sintático por meio dos parâmetros gestuais-espaciais do CORE-SL. O modelo foi validado de forma teórica, por uma linguista computacional, pois inclui a Teoria da Morfologia da Libras, dessa forma agregamos conhecimento linguístico à construção do modelo.

Para uma síntese da proposta, uma abstração de uso do Lex-Libras na HCI-SL implicará: 1 - O usuário que interage com a Arquitetura via Camada de Aplicação e seleciona um dado serviço, como o modulo gerador de sinais (Avatar); 2- O usuário digita determinado sinal; 3- O CORE-SL recebe a solicitação e reconhece que aquele sinal é composto por um processo de flexão; 4- O CORE-SL então encaminha a solicitação para o Lex-libras, responsável pelo módulo morfológico, este faz a consulta do sinal; 5- O Lex-Libras fornece para aplicação as descrições adequadas em formato JSON, conforme Figura 11.



**Figura 11. Abstração de Funcionamento do Lex-Libras na HCI-SL.**

Fonte: [Garcia et al. 2021], no prelo.

O modelo fonológico do CORE-SL é, por hipótese universal no que se refere às LS. Embora possam variar com relação às configurações de mãos às especificidade das regras fonológicas para cada língua, como acontece também com as línguas orais-auditivas, descritas a partir do alfabeto fonético internacional, apesar de cada língua fazer uso de fonemas específicos dentro dessa possibilidade. Um das principais metas, a partir dessa pesquisa, se refere à necessidade de desenvolver um trabalho multi, inter e transdisciplinar na construção das regras dos diferentes níveis linguísticos.

## 6. CONSIDERAÇÕES FINAIS

O Lex-Libras [Silva 2020], [Silva et al. 2021] (No prelo) realizou uma abordagem *bottom-up*<sup>7</sup>, usando a o modelo fonológico para a descrição morfossintática, assim o formalismo gera itens lexicais e os organiza na forma de lexemas e gramemas. A

<sup>7</sup>Nesta abordagem, começamos a modelagem pelo nível mais baixo - o fonológico, para então subir até os níveis mais altos da língua.

construção do modelo formal com todas as propriedades internas e externas e iniciada pelo nível fonológico permitiu a catalogação de sinais numa representação padrão textual, prevendo a extensão das regras dos sucessivos níveis linguísticos, onde o nível anterior funciona como entrada para o o nível imediatamente posterior [Antunes 2011], [Antunes 2015], [Silva 2020], [Silva et al. 2021] (No prelo).

As comunidades surdas têm um conhecimento intuitivo de sua língua. Para a construção de regras de equivalência é necessário, como em qualquer elaboração de tradutores, uma perspectiva de especialistas de áreas específicas e também considerar a necessidade da participação de surdos bilíngues que dominem tais áreas de conhecimento específicas. Já estão disponíveis informações sobre problemas de geração de frases pelos tradutores existentes [Silva 2020], pela percepção das próprias comunidades surdas que, contudo, não conseguem determinar qual regra e em qual nível específico os problemas ocorrem, no entanto as comunidades surdas deverão referendar este trabalho em futuras oficinas.

Por outro lado, os ouvintes que utilizam os tradutores, quando sinalizarem as frases pesquisadas [Silva 2020], estão cometendo os mesmos equívocos gramaticais dos avatares, uma vez que desconhecem as regras gramaticais da Libras. Dessa forma, esta pesquisa deve trazer mais qualidade e assertividade para a tradução no futuro e, assim, proporcionar melhores recursos de acessibilidade e comunicabilidade.

A extensão deste trabalho com base num *corpus*<sup>8</sup> como os apresentados em [Herbig 2018], [Iatskiu 2019], [Silva 2020], construído de maneira computacional formal, faz parte dos trabalhos futuros, assim como o módulo de geração por um avatar e ainda uma revisão para adequação de outros fenômenos morfológicos que não foram feitos na dissertação de mestrado, além da formalização dos níveis sintático-discursivo. Adicionalmente, os resultados desta pesquisa serão apropriados por jogos e aplicativos educacionais já desenvolvidos, [Canteri et al. 2019], e em desenvolvimento para as comunidades surdas.

## Referências

- Amaral, W. M. (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. R. (2011). Um modelo de descrição computacional da fonologia da língua de sinais brasileira. Master's thesis, Programa de Pós-Graduação em Informática, Curitiba-PR.
- Antunes, D. R. (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 Pós-Graduação em Informática do Setor de Ciências Exatas da Universidade Federal do Paraná, Curitiba - PR.

---

<sup>8</sup>Para um estudo mais aprofundado envolvendo PLN, é necessário o uso de um corpus linguístico, que é o conjunto de textos escritos e registros orais em uma determinada língua e que serve como base de análise. Essa abordagem é atualmente conhecida como Linguística de Corpus, uma vez que a interpretação do fenômeno linguístico foi baseada na observação dos dados do *corpus* e não em intuições da analista [Felipe 2007].

- 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. M. U. d. (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. R. (2019). Conceptual framework to support a web authoring tool of educational games for deaf children.
- De Martino, J. M., Silva, I. R., Bolognini, C. Z., Costa, P. D. P., Kumada, K. M. O., Coradine, L. C., Brito, P. H. S., do Amaral, W. M., Benetti, Â. B., Poeta, E. T., Angare, L. M. G., Ferreira, C. M., and De Conti, D. F. (2017). Signing avatars: making education more inclusive. *Universal Access in the Information Society*, 16(3):793–808.
- 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. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/etd/article/view/803>.
- Felipe, T. (2007). *Descrição da língua de sinais: desafios teóricos e práticos 1*. WalPrint Gráfica e Editora.
- 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.
- 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. S., Felipe, T. A., Guedes, A. L. P., Antunes, D. R., Iatskiu, C. E., Todt, E., Bueno, J., Trindade, D. d. F. G., Gonçalves, D. A., Canteri, R., Canal, M., Ferreira, M. A. M., Silva, A. M. C., Galvão, L., and Rodrigues, L. (2021). Deaf inclusion through brazilian sign language: A computational architecture supporting artifacts and interactive applications and tools. In *Universal Access in Human-Computer Interaction. Access to Interaction*. Springer International Publishing. No prelo.
- Garcia, L. S., Guimarães, C., Antunes, D. R., 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.

- Herbig, G. R. (2018). Core-sl-sign register: Ferramenta computacional para a especificação de sinais da libras. Master's thesis, Programa de Pós-Graduação em Informática, setor de Ciências Exatas, da Universidade Federal do Paraná, Curitiba - PR.
- 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 Pós-Graduação em Informática do Setor de Ciências Exatas da Universidade Federal do Paraná, Curitiba - PR.
- Lima, M. A. C. B. (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. M. (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.
- Silva, A. M., Felipe, T. A., Garcia, L., Antunes, D. R., and Guedes, A. P. (2021). Lex-Libras: Morphosyntactic Model of the Brazilian Sign Language to Support a Context-Based Machine Translation Process. *ICEIS 2021 - Proceedings of the 23th International Conference on Enterprise Information Systems*. no prelo.
- Volpe, G., Hasegawa, R., and Carlos, S. (2003). Projeto pul . Technical report, Núcleo Interinstitucional de Linguística Computacional - NILC.
- Wirth, N. (1977). What can we do about the unnecessary diversity of notation for syntactic definitions? *Communications of the ACM*, 20(11):822–823.