

**Universidade Federal da Paraíba  
Centro de Informática  
Programa de Pós-Graduação em Computação, Comunicação e Artes**

**DILAINNE DANIEL DE ALBUQUERQUE**

**MODELAGEM E DESENVOLVIMENTO DE UMA BASE DE DADOS PARA  
TRADUÇÃO AUTOMÁTICA DE CONTEÚDOS DIGITAIS PARA LÍNGUA  
BRASILEIRA DE SINAIS**

**João Pessoa, PB  
2021**

**Dilainne Daniel de Albuquerque**

**Modelagem e Desenvolvimento de uma Base de Dados para Tradução Automática de Conteúdos Digitais para Língua Brasileira de Sinais**

Dissertação apresentada ao Programa de Pós-Graduação em Computação, Comunicação e Artes (PPGCCA) da Universidade Federal da Paraíba, como requisito parcial para a obtenção do título de Mestre em Computação, Comunicação e Artes, na linha de pesquisa Mídias em ambientes digitais.

Orientador: Prof. Dr. Tiago Maritan Ugulino de Araújo

João Pessoa, PB  
2021

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

A345m Albuquerque, Dilainne Daniel de.

Modelagem e Desenvolvimento de uma Base de Dados para Tradução Automática de Conteúdos Digitais para Língua Brasileira de Sinais / Dilainne Daniel de Albuquerque.

- João Pessoa, 2021.

102 f.: il.

Orientação: Tiago Maritan Ugulino de Araújo.

Dissertação (Mestrado) - UFPB/CI.

1. Base de dados. 2. Língua de sinais. 3. Tradução automática. 4. Corpus de tradução. 5. Glosa. I. Araújo, Tiago Maritan Ugulino de. II. Título.

UFPB/BC

CDU 004.65(043)

**MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DA PARAÍBA**

**ATA Nº 6/2021 - PPGCCA (11.00.64.17)**

**Nº do Protocolo: 23074.087311/2021-61**

**João Pessoa-PB, 27 de Agosto de 2021**

Aos vinte e sete dias do mês de agosto do ano de dois mil e vinte e um, às dezesseis horas, por meio de videoconferência, instalou-se a banca examinadora de dissertação de Mestrado da aluna DILAINNE DANIEL DE ALBUQUERQUE, matrícula 20191000817. A banca examinadora foi composta pelos professores doutores TIAGO MARITAN UGULINO DE ARAUJO, UFPB, presidente; GUIDO LEMOS DE SOUZA FILHO, UFPB, examinador interno; e TANYA AMARA FELIPE DE SOUZA, INES, examinador externo à instituição. Deu-se início a abertura dos trabalhos por parte do professor CARLOS EDUARDO COELHO FREIRE BATISTA, coordenador do Programa, que, após apresentar os membros da banca examinadora e esclarecer a tramitação da defesa, passou a presidência dos trabalhos ao professor TIAGO MARITAN UGULINO DE ARAUJO, que de imediato solicitou à candidata que iniciasse a apresentação da dissertação, intitulada “MODELAGEM E DESENVOLVIMENTO DE UMA BASE DE DADOS PARA TRADUÇÃO AUTOMÁTICA DE CONTEÚDOS DIGITAIS PARA LÍNGUA DE SINAIS”. Concluída a exposição, o Presidente, passou a palavra à professora TANYA AMARA FELIPE DE SOUZA, para arguir a candidata, e, em seguida, ao professor GUIDO LEMOS DE SOUZA FILHO para que fizessem o mesmo, após o que fez suas considerações sobre o trabalho em julgamento, tendo sido APROVADA a candidata, conforme as normas vigentes na Universidade Federal da Paraíba.

*(Assinado digitalmente em 31/08/2021 11:46)*  
CARLOS EDUARDO COELHO FREIRE BATISTA  
COORDENADOR DE CURSO  
Matrícula: 1545948

*(Assinado digitalmente em 31/08/2021 17:04)*  
TIAGO MARITAN UGULINO DE ARAUJO  
PROFESSOR DO MAGISTERIO SUPERIOR  
Matrícula: 2813926

*(Assinado digitalmente em 13/09/2021 16:02)*  
GUIDO LEMOS DE SOUZA FILHO  
PROFESSOR DO MAGISTERIO SUPERIOR  
Matrícula: 1149359

Para verificar a autenticidade deste documento entre em <https://sipac.ufpb.br/documentos/> informando seu número: **6**, ano: **2021**, documento (espécie): **ATA**, data de emissão: **27/08/2021** e o código de verificação: **42cd4fe90f**

*Com todo amor e admiração, dedico este  
trabalho à minha família.*

## AGRADECIMENTOS

Quero agradecer, em primeiro lugar, a Deus, pela vida, pela força e coragem que me foi concedida durante toda esta longa caminhada;

À minha família, especialmente aos meus pais, Valderi Santino de Albuquerque e Juveni Daniel de Albuquerque, aos meus irmãos, Daianne Daniel de Albuquerque, Dannubia Daniel de Albuquerque e Diorgenes Daniel de Albuquerque, aos meus cunhados Ronaldo Soares e Jacellyane Andrade, aos meus sobrinhos Nathan Henrique, João Gabriel, José Filipe e Pedro Miguel, que, com muito amor e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida;

À minha tia Maria Irma Nascimento de Lima e meu primo Marcos Antônio Nascimento de Lima, que me trouxeram a oportunidade de conhecer a Libras;

Ao meu orientador Tiago Maritan, pela confiança no meu trabalho, paciência, orientação e incentivo que tornaram possível a conclusão desta pesquisa.

À todos os meu colegas do LAVID - VLibras, pela disponibilidade e ajuda, em especial à minha coordenadora de projeto Manuella Ascoff pelo apoio;

Aos amigos, pelo incentivo e pelo apoio constante, em especial a Yurika Sato, Iris Regina Nascimento de Lima e Derek Warwick da Silva Tavares que deram suporte nos momentos que mais precisei durante a pesquisa;

À todos o meu muito obrigado!

*“O resultado mais sublime da educação é a tolerância.”*

(Helen Keller)

## RESUMO

Esta pesquisa apresenta uma proposta de modelagem e desenvolvimento de uma base de dados textual bilíngue para apoiar o desenvolvimento de tradutores automáticos de Português para Libras, utilizando uma notação textual em Libras, denominada sistema de transcrição em glosa. Para isso, inicialmente foram definidas algumas regras para especificar aspectos gramaticais que são importantes e característicos da Libras nesta notação. Com base nessa notação, uma base de dados textual bilíngue foi desenvolvida utilizando uma equipe de 10 intérpretes de Libras, resultando num corpus bilíngue contendo 57.206 exemplos de tradução em Português e sua correspondente representação glosa em Libras. Essa base de dados foi então utilizada e aplicada para o treinamento de um novo modelo do tradutor automático do VLibras, e os impactos produzidos no VLibras com relação ao tratamento desses aspectos gramaticais da Libras foram avaliados com usuários surdos. Os resultados indicaram que os usuários avaliaram que compreenderam melhor as sentenças traduzidas quando a base de dados foi incorporada ao VLibras.

**Palavras-chave:** Base de dados. Língua de sinais. Tradução automática. *Corpus* de tradução. Glosa.

## **ABSTRACT**

This research presents a proposal for modeling and developing a bilingual textual database to support the development of automatic translators from Portuguese to Libras, using a textual notation in Libras, called transcription system in gloss. For this, initially some rules were defined to specify grammatical aspects that are important and characteristic of Libras in this notation. Based on this notation, a bilingual textual database was developed using a team of 10 Libras interpreters, resulting in a bilingual corpus containing 57,206 examples of translations in Portuguese and their corresponding gloss representation in Libras. This database was then used and applied for training a new VLibras automatic translator model, and the impacts produced in VLibras in relation to the treatment of these grammatical aspects of Libras were evaluated with deaf users. The results indicated that users rated that they understood the translated sentences better when the database was incorporated into VLibras.

**Keywords:** Database. Sign language. Automatic translation. Translation Corpus. Gloss.

## LISTA DE ILUSTRAÇÕES

### Capítulo 2

- Figura 1** - Parâmetros principais da Libras..... 35
- Figura 2** - Configuração de mão (CM) ..... 35
- Figura 3** - Direcionalidade..... 36
- Figura 4** - Sinal de ‘sentar’ e ‘cadeira’..... 37
- Figura 5** - Verbo com concordância número-pessoal..... 38

### Capítulo 3

- Figura 6** - Representação esquemática..... 47

### Capítulo 4

- Figura 7** - Capturas de Tela do Aplicativo VLibras..... 56

### Capítulo 5

- Figura 8** - Processo de criação da base de dados..... 74
- Figura 9** - Traduções da palavra Adiantar..... 83

## LISTA DE TABELAS

|                    |  |    |
|--------------------|--|----|
| <b>Tabela 1</b> -  | Periodização da educação de surdos no Brasil.....                          | 27 |
| <b>Tabela 2</b> -  | Bases de Dados.....  | 45 |
| <b>Tabela 3</b> -  | Resultados da busca.....   | 46 |
| <b>Tabela 4</b> -  | Características dos cinco estudos selecionados para revisão sistemática... | 48 |
| <b>Tabela 5</b> -  | Pontuação.....   | 56 |
| <b>Tabela 6</b> -  | Aplicação da regra de pontuação.....                                       | 57 |
| <b>Tabela 7</b> -  | Verbos.....  | 57 |
| <b>Tabela 8</b> -  | Verbos de ligação e preposição.....  | 58 |
| <b>Tabela 9</b> -  | Conversão dos artigos.....   | 58 |
| <b>Tabela 10</b> - | Troca e fragmentação de palavras.....                                      | 58 |
| <b>Tabela 11</b> - | Fragmentação de palavras no plural.....                                    | 59 |
| <b>Tabela 12</b> - | Adjetivos.....   | 59 |
| <b>Tabela 13</b> - | Números cardinais.....   | 60 |
| <b>Tabela 14</b> - | Números ordinais.....  | 60 |
| <b>Tabela 15</b> - | Conversão dos números ordinais.....  | 61 |
| <b>Tabela 16</b> - | Número e quantidade.....   | 61 |
| <b>Tabela 17</b> - | Representação numérica.....  | 61 |
| <b>Tabela 18</b> - | Período de tempo (hora) .....  | 62 |
| <b>Tabela 19</b> - | Palavras compostas.....  | 62 |
| <b>Tabela 20</b> - | Verbos com concordância .....  | 63 |
| <b>Tabela 21</b> - | Verbos Direcionais relacionados aos pronomes pessoais no plural.....       | 63 |
| <b>Tabela 22</b> - | Verbos que incorporam negação.....   | 64 |
| <b>Tabela 23</b> - | Homônimos (palavra Colar) .....  | 65 |
| <b>Tabela 24</b> - | Aplicação de homônimos em frases (palavra colar).....                      | 65 |
| <b>Tabela 25</b> - | Homônimos (palavra Artigo) .....   | 65 |
| <b>Tabela 26</b> - | Aplicação de homônimos em frases (palavra Artigo).....                     | 66 |
| <b>Tabela 27</b> - | Concordância verbal (verbo Apagar) .....                                   | 66 |

|  |    |
|--|----|
| <b>Tabela 28</b> - Aplicação de homônimos em frases (verbo Apagar).....  | 66 |
| <b>Tabela 29</b> - Regra de representação de lugares.....                | 66 |
| <b>Tabela 30</b> - Classificação de lugar: contextualização.....         | 67 |
| <b>Tabela 31</b> - Decomposição de palavras: classificação de lugar..... | 67 |
| <b>Tabela 32</b> - Nomes pessoais.....                                   | 68 |
| <b>Tabela 33</b> - Amplificação e redução de intensidade.....            | 68 |
| <b>Tabela 34</b> - Expressões idiomáticas.....                           | 69 |
| <b>Tabela 35</b> - Substituição de frases por uma expressão.....         | 70 |
| <b>Tabela 36</b> - Ditados populares.....                                | 70 |
| <b>Tabela 37</b> - Substituição por sinônimos.....                       | 71 |
| <b>Tabela 38</b> - Ocorrência das regras no <i>Corpus</i> .....          | 75 |
| <b>Tabela 39</b> - Variação de sentenças.....                            | 75 |
| <b>Tabela 40</b> - Idade e identidade de gênero.....                     | 80 |
| <b>Tabela 41</b> - Uso de aplicativos de Tradução para Libras.....       | 81 |
| <b>Tabela 42</b> - Média Final do Grupo A – Regras.....                  | 87 |
| <b>Tabela 43</b> - Média Final do Grupo B- DL.....                       | 88 |
| <b>Tabela 44</b> - Média Final do Grupo C – Híbrido.....                 | 89 |
| <b>Tabela 45</b> - Média Final e desvio padrão dos Resultados.....       | 89 |

## LISTA DE ABREVIATURAS E SIGLAS

|                 |   |
|-----------------|---|
| <b>ASL</b>      | <i>American Sign Language</i>   |
| <b>BLEU</b>     | <i>Bilingual Evaluation Understudy</i>  |
| <b>CL</b>       | Classificadores   |
| <b>CM</b>       | Configuração de mão   |
| <b>CRSDs</b>    | <i>Continuous Vector-Space Representation of Sentences for Data Selection</i> |
| <b>CVR</b>      | <i>Continuous vector-space representation</i>                                 |
| <b>DA</b>       | Deficiente Auditivo   |
| <b>DGS</b>      | Língua de Sinais Alemã  |
| <b>DS</b>       | <i>Data selection</i>   |
| <b>Elis</b>     | Sistema de Escrita das Línguas de Sinais                                      |
| <b>FENEIDA</b>  | Federação Nacional de Educação e Integração do Deficiente Auditivo            |
| <b>GSL</b>      | Língua de Sinais Grega  |
| <b>HamNoSys</b> | <i>Hamburg Notation System</i>  |
| <b>IA</b>       | Inteligência Artificial   |
| <b>INES</b>     | Instituto de Educação dos Surdos  |
| <b>INSM</b>     | Instituto Nacional de Surdos-Mudos  |
| <b>ISL</b>      | Língua de Sinais Indiana  |
| <b>JIDM</b>     | <i>Journal of Information and Data Management</i>                             |
| <b>LAVID</b>    | Laboratório de Aplicações de Vídeo Digital                                    |
| <b>LIBRAS</b>   | Língua Brasileira de Sinais   |
| <b>LSF</b>      | Língua de Sinais Francesa   |
| <b>M</b>        | Movimento   |
| <b>MOSES</b>    | Sistema de tradução automática estatística                                    |
| <b>MPOG</b>     | Ministério do Planejamento, Orçamento e Gestão                                |
| <b>OD</b>       | Orientação ou direcionalidade   |
| <b>ODA</b>      | Objeto Digital de Aprendizagem  |
| <b>OMS</b>      | Organização Mundial de Saúde  |

|              |  |
|--------------|--|
| <b>P</b>     | Plural   |
| <b>PA</b>    | Locação o ponto de articulação                       |
| <b>PBT</b>   | <i>Phrase-based translation</i>                      |
| <b>PcD</b>   | Pessoa com Deficiência                               |
| <b>PLN</b>   | Processamento de Línguas Naturais                    |
| <b>pt-BR</b> | Português Brasileiro                                 |
| <b>RBMT</b>  | <i>Rule-Based Machine Translation</i>                |
| <b>S</b>     | Singular   |
| <b>SL</b>    | <i>Sign language</i>                                 |
| <b>SMT</b>   | <i>Statistical Machine Translation</i>               |
| <b>SSMT</b>  | <i>Structural statistical Machine translation</i>    |
| <b>SSMT</b>  | <i>Structural Statistical Machine Translation</i>    |
| <b>SW</b>    | <i>SignWriting</i>                                   |
| <b>TA</b>    | Tradução Automática                                  |
| <b>TILS</b>  | Tradutor e Intérprete de Língua Brasileira de Sinais |
| <b>TSL</b>   | Língua de Sinais Taiwanesa                           |
| <b>UFPB</b>  | Universidade Federal da Paraíba                      |

## SUMÁRIO

|              |  |           |
|--------------|--|-----------|
| <b>1</b>     | <b>INTRODUÇÃO</b> .....  | <b>16</b> |
| <b>1.1</b>   | <b>Objetivos</b> .....   | <b>19</b> |
| <b>1.2</b>   | <b>Estrutura da Proposta de Dissertação</b> .....  | <b>19</b> |
| <b>2</b>     | <b>LÍNGUA DE SINAIS BRASILEIRA: INTERPRETAÇÃO E TRADUÇÃO AUTOMÁTICA</b> .....            | <b>21</b> |
| <b>2.1</b>   | <b>Definição da língua de sinais e cultura surda</b> .....                               | <b>24</b> |
| <b>2.2</b>   | <b>Interpretação e tradução da Libras</b> .....  | <b>32</b> |
| <b>2.3</b>   | <b>Linguística da Libras</b> .....   | <b>33</b> |
| <b>2.3.1</b> | <i>Aspectos fonológicos</i> .....  | <b>34</b> |
| <b>2.3.2</b> | <i>Aspectos Morfológicos</i> .....   | <b>37</b> |
| <b>2.3.3</b> | <i>Aspectos sintáticos</i> .....   | <b>38</b> |
| <b>2.3.4</b> | <i>Aspectos semânticos e pragmáticos</i> .....   | <b>39</b> |
| <b>2.4</b>   | <b>Tradução automática</b> .....   | <b>40</b> |
| <b>2.5</b>   | <b>Sistemas de notações escritas de línguas de sinais</b> .....                          | <b>42</b> |
| <b>2.6</b>   | <b>Considerações do Capítulo</b> .....   | <b>44</b> |
| <b>3</b>     | <b>REVISÃO SISTEMÁTICA DA LITERATURA</b> .....   | <b>45</b> |
| <b>3.1</b>   | <b>Protocolo de Revisão Sistemática</b> .....  | <b>45</b> |
| <b>3.2</b>   | <b>Busca, Seleção e Filtragem dos Trabalhos</b> .....                                    | <b>46</b> |
| <b>3.3</b>   | <b>Discussão e análise</b> .....   | <b>47</b> |
| <b>3.4</b>   | <b>Considerações do Capítulo</b> .....   | <b>52</b> |
| <b>4</b>     | <b>REGRAS APLICADAS NA CRIAÇÃO DE UM <i>CORPUS</i> PORTUGUÊS BRASILEIRO-LIBRAS</b> ..... | <b>54</b> |
| <b>4.1</b>   | <b>Pontuação: renomeação da pontuação para tags pré-definidas</b> .....                  | <b>56</b> |
| <b>4.2</b>   | <b>Verbos</b> .....  | <b>57</b> |
| <b>4.3</b>   | <b>Artigos</b> .....   | <b>58</b> |
| <b>4.4</b>   | <b>Troca de palavras</b> .....   | <b>58</b> |
| <b>4.5</b>   | <b>Adjetivos: Unificação do gênero dos adjetivos</b> .....                               | <b>59</b> |
| <b>4.6</b>   | <b>Números ordinais</b> .....  | <b>59</b> |

|              |  |           |
|--------------|--|-----------|
| <b>4.7</b>   | <b>Período de tempo (hora)</b> .....                                       | <b>62</b> |
| <b>4.8</b>   | <b>Palavras compostas</b> .....  | <b>62</b> |
| <b>4.9</b>   | <b>Verbos com concordância número-pessoal</b> .....                        | <b>63</b> |
| <b>4.10</b>  | <b>Frases negativas</b> .....  | <b>63</b> |
| <b>4.11</b>  | <b>Homônimos e palavras homógrafas</b> .....                               | <b>64</b> |
| <b>4.12</b>  | <b>Classificação de lugar - Cidade, Estado e País</b> .....                | <b>66</b> |
| <b>4.13</b>  | <b>Nomes pessoais</b> .....  | <b>67</b> |
| <b>4.14</b>  | <b>Amplificação e redução de intensidade</b> .....                         | <b>68</b> |
| <b>4.15</b>  | <b>Expressões idiomáticas</b> .....  | <b>69</b> |
| <b>4.16</b>  | <b>Contexto</b> .....  | <b>70</b> |
| <b>5</b>     | <b>RESULTADOS E DISCUSSÕES</b> .....                                       | <b>72</b> |
| <b>5.1</b>   | <b>Processo de criação da base de dados</b> .....                          | <b>72</b> |
| <b>5.2</b>   | <b>Instrumentação</b> .....  | <b>77</b> |
| <b>5.3</b>   | <b>Análise dos resultados</b> .....  | <b>80</b> |
| <b>5.3.1</b> | <i>Análise do Perfil dos Usuários</i> .....                                | <b>80</b> |
| <b>5.3.2</b> | <i>Análise das sentenças traduzidas automaticamente pelo VLibras</i> ..... | <b>81</b> |
| <b>6</b>     | <b>CONSIDERAÇÕES FINAIS</b> .....  | <b>90</b> |
|              | <b>REFERÊNCIAS</b> .....   | <b>92</b> |
|              | <b>ANEXO A</b> .....   | <b>99</b> |

## 1 INTRODUÇÃO

A tradução automática de língua de sinais tornou-se um meio de comunicação importante para pessoas surdas e ouvintes, tendo como principal finalidade a redução das barreiras de acesso à informação das pessoas surdas na sociedade, especialmente em cenários que possuem grande volume e dinamicidade de dados, como por exemplo, na Web. Devido ao fato de serem utilizados por usuários que não estão acompanhados de um intérprete da Língua Brasileira de Sinais (Libras), os tradutores automáticos tendem a apresentar uma sinalização mais focada na língua portuguesa, apresentando dificuldade para tratar determinados aspectos gramaticais da Libras. Esse processo de tradução é algumas vezes dito como resultado de uma Libras aportuguesada ou Português sinalizado, uma vez que emprega, de maneira geral, primordialmente as regras gramaticais da língua portuguesa, que pode ter como consequência uma sinalização agramatical da Libras.

Esse tradutor automático, sendo bilíngue (Português Brasileiro - Libras) consiste em uma ferramenta para usuários Surdos e ouvintes, cujo objetivo é estabelecer a comunicação entre esses usuários. Para os Surdos é justamente a compreensão da tradução de um item lexical ou uma frase em Português para a Libras, levando em consideração que a entrada dos caracteres é sempre na língua portuguesa. Isso significa dizer que em um cenário de uso, os Surdos utilizam essa ferramenta quando não há disponibilidade de intérprete que realize a tradução para a Libras ou estejam lendo algum tipo de texto em português. Por outro lado, os ouvintes a utilizam para consulta de itens lexicais, como fazem com o Dicionário da Libras (FELIPE et al., 2002/2005), ou quando estão no processo de aprendizagem, aquisição da língua de sinais, ou ainda na verificação de como seria uma frase específica em Libras.

Aprofundando estes estudos, percebeu-se que uma das dificuldades existentes no desenvolvimento desses tradutores automáticos para língua de sinais é a dificuldade de encontrar bases de dados contendo conteúdos descritos tanto nas línguas orais quanto nas línguas de sinais. Os *corpus* “[...] são registros de diferentes gêneros textuais escritos e falados de diferentes línguas” (QUADROS, 2016, p. 9). As produções de *corpus* concentram-se no “[...] registro de interações em Libras por meio de filmagens em vídeo” (QUADROS, 2016, p. 10), elaboração de glossários e simplificação textual<sup>1</sup>. No entanto, os *corpus* de tradução atuais ainda não atingiram um nível satisfatório, comparado às línguas de modalidade oral-auditiva. Um dos

---

<sup>1</sup> “Os arquivos gerados possuem informações gerais sobre o texto e sobre o corpus, sobre como estão organizadas as partes que compõem o texto (como parágrafos e sentenças), sobre tokens (palavras) e suas características gramaticais, sobre os alinhamentos entre sentenças de textos simplificados, sobre operações de simplificação ocorridas, bem como sobre o próprio texto em si.” (SANTOS et al., 2009, p.379)

problemas dessa modalidade é que “[...] a aquisição de dados para SLs<sup>2</sup>, é uma tarefa demorada e cara em comparação com a aquisição de dados falados ou escritos.” (PORTA et al., 2014, p. 794, tradução nossa)<sup>3</sup>.

Outra dificuldade é a ausência de uma forma de representação de glosa aplicada na estrutura da tradução da língua oral para língua de sinais que possa ser amplamente utilizada e reconhecida pela comunidade surda, de modo que as línguas de sinais possam ser descritas textualmente com regras definidas previamente.

Linguistas da área, tais como Felipe (1988, 1998, 2001, 2006), Quadros e Karnopp (2004), por exemplo, usam formas diferentes de descrever as glosas, ou seja, não existe um padrão de representação mesmo entre os principais linguistas da área. Como exemplo, pode-se observar o uso da arroba (@) em Libras indicando “[...] a ausência da marca de gênero [...] Exemplo: AMIG@, FRI@, MUIT@.” (PAIVA et al, 2016, p.23) Criada por Felipe (1998) e que vem sendo muito utilizada. Assim como a marcação das expressões faciais propostas por Felipe (1989), sendo ampliadas por essa pesquisadora em trabalhos posteriores e também por Quadros e Karnopp (2004), entre outras e a identificação de verbos direcionais numa sentença (JOÃO.a MARIA.b a.DAR.b LIVRO NÃO - a: João, b: Maria). (PAIVA et al, 2016) Tendo substituído a proposta de Felipe (200), com relação aos números para as pessoas do discurso (primeira, segunda e terceira – singular e plural: 1s, 2s, 3s; 1p, 2p, 3p)

Essa representação textual pode ser caracterizada como um sistema de transcrição com regras específicas, não sendo considerado um sistema de escrita. Sistema de transcrição e escrita de sinais possuem diferenças, estas determinadas de acordo com o uso da transcrição e objetivos específicos. Para a Libras, há duas propostas de escrita de sinais: a Elis (brasileira) e Signwriting (estadunidense).

A motivação da escolha desta temática, parte da observação de que, recentemente, os principais trabalhos na literatura científica que envolvem Aprendizagem de Máquina (*Machine Learning*) e Tradução Automática (TA) requerem o uso de grandes<sup>4</sup> bases de dados, contendo centenas de milhares de exemplos (de imagens, vídeos, exemplos de tradução, entre outros). Pois “[...] a precisão e qualidade da tradução dependem do tamanho e da cobertura do *corpus*

---

<sup>2</sup> Sign language.

<sup>3</sup> “[...] acquisition for SLs a time-consuming and expensive task compared to the acquisition of spoken or written data.”

<sup>4</sup> A qualidade da tradução pode ser beneficiada quando o número mínimo de aparições de cada palavra aumenta. Percebe-se, assim, “[...] a necessidade de usar o mais alto vocabulário para obter a melhor representação da frase.” (CHINEA-RIOS, 2019, p. 62, tradução nossa). Incluindo o maior quantitativo de léxico para uma melhor representação da sentença, em contextos específicos no discurso. (“[...] the necessity to use the highest vocabulary to obtain the better sentence representation”)

paralelo bilíngue, que precisa ser, em geral, bastante extenso para cobrir as diferenças sintáticas e semânticas<sup>5</sup> entre as línguas envolvidas.” (ARAÚJO, 2012, p. 26) De modo que, os modelos de tradução automática “[...] são construídos a partir de *corpus* paralelo alinhado sentencialmente, ou seja, um conjunto de sentenças que são traduções mútuas baseadas em frases” (CASELI; NUNES, 2009, p. 4), voltadas para a aprendizagem de máquina. Nesta perspectiva, pretendendo alcançar o maior número de pesquisas com a temática da tradução automática, foram realizadas pesquisas nas bases de dados arXiv, CAPES, Scielo, Scopus, utilizando as seguintes palavras-chave “*statistical machine translation*”, “*statistical machine translation*” AND “*sign language*”, “*statistical machine translation*” AND “*sign language*” AND “*Libras*”. Buscando exemplos de *corpus* com representação textual de sentença-sentença, apresentadas em glosas que possam ser usadas no processo de tradução automática das línguas orais para língua de sinais, isto é, exemplos de caracteres utilizados para transformar língua portuguesa em língua de sinais e assim aplicá-los no uso da lógica computacional dos tradutores automáticos. Não se trata do resultado do produto final (comunicados, poesias, aulas, palestras, piadas, histórias, DVDs), mas do processo de tradução realizado.

Diante da necessidade de melhorar o nível da tradução automática, visto que a Libras é “[...] uma língua genuína e independente da Língua Portuguesa, com estrutura gramatical própria, no processo automático de tradução entre a língua-fonte e a língua-alvo pode ocorrer perda de informação.” (REIS et al., 2018, p. 1), identificou-se que a modelagem e o desenvolvimento de uma base de dados de conteúdo de tradução voltada para a aprendizagem de máquina são importantes, como a criação de um padrão de tradução mais abrangente, com frases pré-definidas e elaboradas segundo as regras gramaticais específicas da língua de sinais.

De maneira geral, os tradutores automáticos não fazem uso de todas as características estruturais específicas da Língua de Sinais. Nesta pesquisa, através do desenvolvimento de um *corpus* de tradução, busca-se implementar algumas dessas características, incluindo, por exemplo, aspectos gramaticais importantes como direcionalidade, tratamento de advérbios de intensidade, tratamento de advérbios de negação, entre outros.

A base de dados utiliza uma linguagem formal (ou informal, uma vez que não segue a estrutura da língua portuguesa e sim da Libras, fazendo alterações quanto à contextualização) de descrição abrangendo a estrutura gramatical própria da Libras. Essa linguagem formal procura respeitar o estilo de tradução intersemiótica, definida como a tradução de um sistema de linguagem para outro, podendo utilizar signos, incluindo efeitos visuais e táteis, sendo um

---

<sup>5</sup> sintático-semântico-discursivas

processo de recriação, como uma elaboração de significados. A utilização da função semiótica é essencial para descrever o contexto e adotar as palavras que fariam mais sentido na Libras, a substituição de termos e frases tem como princípio apresentar os parâmetros da língua de sinais.

## 1.1 Objetivos

Para endereçar os problemas apresentados anteriormente, o objetivo geral deste estudo é modelar e desenvolver uma base de dados para auxiliar a tradução automática de conteúdos digitais para Língua Brasileira de Sinais.

Para operacionalizar este objetivo, foram definidos os seguintes objetivos específicos:

- Modelar uma linguagem formal de descrição para a tradução de uma língua oral escrita para língua de sinais representado textualmente;
- Elaborar um conjunto de regras de descrição formal que contemplem as adequações morfológicas, sintáticas, semânticas e pragmáticas da tradução de conteúdo da língua oral-auditiva na modalidade escrita para a língua de sinais representada textualmente, em glosas;
- Construir uma base de dados utilizando a linguagem formal desenvolvida com pares de exemplos da Língua Portuguesa do Brasil na modalidade oral-auditiva para a Língua Brasileira de Sinais - Libras.

## 1.2 Estrutura da Proposta de Dissertação

Esta dissertação está estruturada em seis capítulos. Após esta introdução, o segundo capítulo traz a fundamentação teórica da Libras, apresentando os principais aspectos que fundamentam este trabalho. As diferenças apresentadas pela tradução e interpretação da Libras, sua estrutura gramatical e características dos tradutores automáticos, como também a apresentação de outros sistemas de notação em escrita de sinais. No terceiro capítulo, registraram-se assuntos relativos à revisão sistemática da literatura, identificando pesquisas relacionadas ao tema proposto. Apresentando a definição de *strings* de busca, seleção e filtragem dos trabalhos relevantes para o desenvolvimento desta pesquisa, e a análise dos artigos selecionados.

O quarto capítulo apresenta, com base em pesquisas teóricas linguísticas, as regras desenvolvidas e aplicadas na criação de um *corpus* bilíngue. Com a utilização de métodos

descritivo-exploratórios para a seleção das informações fundamentais para a elaboração e desenvolvimento desta pesquisa.

No quinto capítulo, a metodologia adotada na análise dos resultados alcançados. Tais como, o processo de criação do *corpus* de tradução bilíngue, referente aos participantes e período de tempo empregado na modelagem e desenvolvimento da base de dados, o desenvolvimento e aplicação dos questionários, assim como a análise dos usuários e aspectos gramaticais das traduções. Por fim, as considerações sintetizam os resultados alcançados nesta pesquisa.

## 2 LÍNGUA BRASILEIRA DE SINAIS: INTERPRETAÇÃO E TRADUÇÃO AUTOMÁTICA

A Organização Mundial de Saúde - OMS, classifica os tipos de deficiência como: física, visual e auditiva, sendo a deficiência auditiva caracterizada pela perda ou incapacidade de ouvir. Os surdos, igual a diversos outros grupos sociais, possuem sua própria identidade cultural. Diante disso, “[...] têm assumido seu lugar enquanto cidadãos participando socialmente das decisões que englobam aspectos relacionados à inclusão social de pessoas surdas.” (NASCIMENTO, 2015, p. 79) Em suas comunidades<sup>6</sup> são realizados encontros, debates, unindo-se na busca por seus direitos, permitindo que suas identidades se formem a partir da convivência junto a seus iguais.

Um dos principais desafios das pessoas surdas está relacionado ao seu processo de comunicação, seja ela ouvinte-surdo e/ou surdo-surdo, isso quando “[...] o surdo não fica isolado do mundo das coisas, do mundo visual. O surdo fica isolado do mundo dos homens, do mundo da conversa, do mundo do diálogo verbal” (ROSA; CRUZ, 2001, p. 41). Isso acontece porque a maioria dos surdos pertence a famílias de ouvintes e constituem famílias (cônjuges e filhos que podem ser ouvintes), em alguns casos podem se sentir excluídos quando estão em ambientes de cultura oral-auditiva ou cultura ouvinte. Por essa razão, torna-se essencial esse acesso à informação, oportunizando aos indivíduos possibilidade de conhecimentos e desenvolvimento intelectual, “[...] enquanto os que não possuem informação tendem a ficar isolados desse meio.” (WILDAUER; INABA; SILVA, 2014, p. 10).

Conhecer as particularidades das comunidades surdas se torna necessário para realizar discussões a respeito de suas necessidades culturais e informacionais. Compreender seu modo de pensar e sua visão do mundo, pois os “[...] sujeitos surdos veem o mundo de maneira diferente em alguns aspectos, porque suas vidas são diferentes por terem mais experiência visual e por estarem longe da experiência auditiva.”(STROBEL, 2008, p.59)<sup>7</sup> Visto que nem os Surdos e nem as comunidades surdas são homogêneas, ou seja, elas apresentam um multiculturalismo<sup>8</sup>, identificar suas características é fundamental para reconhecer suas

---

<sup>6</sup> “A comunidade surda, na verdade não é só de surdos, já que tem sujeitos ouvintes junto, que são família, intérpretes, professores, amigos e outros que participam e compartilham os mesmos interesses em comuns em uma determinada localização que podem ser as associações de surdos, federações de surdos, igrejas e outros.” (STROBEL, 2009, p. 6)

<sup>7</sup> As autoras Strobel como a Perlin são pessoas surdas.

<sup>8</sup> “Perlin (1998) vai além, classificando a identidade surda em cinco grupos: 1) identidade surda: aquela que cria um espaço cultural visual dentro de um espaço cultural diverso, ou seja, recria a cultura visual, reivindicando à História a alteridade surda; 2) identidades surdas híbridas: aquelas de surdos pós-locutivos, que nasceram ouvintes e se tornaram surdos; 3) identidades surdas de transição e formadas por surdos que viveram sob o domínio da

necessidades como a forma de buscar e recuperar informações na sociedade e nos meios digitais.

Alguns Surdos se sentem excluídos ao não conseguirem se comunicar e interagir com outros grupos linguísticos e sociais. Esses grupos linguísticos podem ser classificados, inicialmente, pelas seguintes terminologias: surdos, Surdos e deficientes auditivos. O termo ‘surdo’, com letra minúscula, diz respeito aos sujeitos em sua totalidade, àqueles que não escutam ou não respondem a sons igual a um ouvinte (CARDOSO, 2016, p.5 apud LACERDA, 2020, p.3). O termo ‘Surdo’, com letra maiúscula, está relacionado à identificação de um grupo de pessoas com a cultura surda.

Ser Surdo (com “S” maiúsculo) é reconhecer-se por meio de uma identidade compartilhada por pessoas que utilizam língua de sinais e não veem a si mesmas como sendo marcadas por uma perda, mas como “membros de uma minoria linguística e cultural com normas, atitudes e valores distintos e uma constituição física distinta. (LANE, 2008. p. 284 apud BISOL & SPERB, 2010, p.8)

Essa utilização de “S” maiúsculo, foi utilizada pela primeira vez por Felipe (1988), a partir dessa utilização, em inglês, Deaf, pela surda estadunidense Carol Padden. Pode-se dizer que a “[...] cultura surda, a experiência visual e o uso da língua de sinais sustentam o encontro e a vida da comunidade surda.” (KARNOPP, 2010, p. 164) O termo ‘Deficiente Auditivo’ (D.A.), por outro lado, é mais utilizado na área da saúde, classificando os sujeitos surdos de acordo com os graus de surdez e perda sensorial. Também podem ser utilizados pelos sujeitos que “[...] não se identificam com a comunidade “Surda”.” (LACERDA, 2020, p. 3, grifo do autor)

Para alguns Surdos brasileiros, a língua portuguesa não deve ser definida como sua língua natural<sup>9</sup>. Eles consideram que sua língua natural de comunicação é a Libras. Pois, partindo da aquisição da língua de sinais, o sujeito desenvolverá sua identidade surda, como uso da língua, e contato com outros Surdos. Esse “[...] contato com outro surdo que também use a língua de sinais surgem novas possibilidades interativas, de compreensão, de diálogo, de aprendizagem, que não são possíveis apenas por meio da linguagem oral. (SANTANA; BERGAMO, 2005, p.567)

---

cultura ouvinte (em geral, os surdos oralizados) e que posteriormente são inseridos na comunidade surda (processo de “des-ouvintização” da representação da identidade); 4) identidade surda incompleta: aquela dos surdos que vivem sob o domínio da cultura ouvinte e negam a identidade surda; 5) identidades surdas flutuantes, formadas por sujeitos surdos que reconhecem ou não sua subjetividade, mas que desprezam a cultura surda, não se comprometendo com a comunidade.” (apud GESUELI, 2006, p.283)

<sup>9</sup> “[...] a criança surda adquire de forma espontânea sem que seja preciso um treinamento específico [...]” (DIZEU, CAPORALI; 2005, p. 584)

No entanto, no Brasil, a maioria dos ouvintes não utilizam a Libras, e muitas vezes nunca tiveram contato com pessoas surdas, podendo ser considerada uma questão de natureza político-educacional, como também, da indústria tecnológica nos momentos atuais (transplantes, implantes cocleares e aparelhos para audição). Há também benefícios para as comunidades ouvintes em aprender uma língua de outra modalidade, como a Libras, pois, além de favorecer a interação social, como o contato com uma identidade cultural específica, traz uma conscientização sobre a importância do aprendizado de uma segunda língua, sendo esta reconhecida oficialmente como uma língua brasileira. Esse aprendizado pode ser complementado através da utilização de tradutores automáticos ou Objetos Digitais de Aprendizagem (ODA's).

Nóbrega (2016) considera o ODA como um dos produtos que facilitam a interação social, utilizando técnicas e tecnologias de informação e comunicação, buscando estimular a produção de conhecimento. A partir dessas características, os objetos de aprendizagem podem ser inseridos em comunicação por redes de computação, especialmente utilizando o “[...] modelo de inteligência coletiva, caracterizado pela interatividade, participação, socialização e criatividade em redes virtuais de relacionamentos, quer seja pessoal, acadêmico, escolar ou científico.”(NÓBREGA, 2016, P. 33)

Strobel (2008, p.65), define a cultura surda que

[...], compreende a língua de sinais, as ideias, as crenças, os costumes e os hábitos de povo surdo, inclusive o artefato material que podem ser os dispositivos eletrônicos não auditivos, o telefone para surdos – um pouco maior que o telefone convencional, na parte de cima tem um encaixe de fone e embaixo dele tem um visor onde aparece escrito digitado e mais abaixo tem as teclas para digitar -, campainha com sinal luminoso, relógio com despertador vibratório e entre outros. (STROBEL, 2008, p.65)

Como exemplos de produções culturais do povo Surdo, pode-se apresentar a Literatura Surda, que a língua de sinais e cultura surda, podendo ser apresentada numa versão bilíngue (em português e escrita da língua de sinais). Para Karnopp (2006), a literatura surda “[...] é a produção de textos literários em sinais, que entende a surdez como presença de algo e não como falta, possibilitando outras representações de surdos, considerando-os como um grupo linguístico e cultural diferente.” (p. 102).

Pode-se, no contexto dos artefatos materiais, mencionar o desenvolvimento e constante aperfeiçoamento dos tradutores automáticos para Língua de sinais que serão descritos detalhadamente na Seção 2.4, relacionados ao Processamento de Linguagem Natural (PLN) que “[...] é o nome que se dá à área de pesquisa que se dedica a investigar, propor e desenvolver

formalismos, modelos, técnicas, métodos e sistemas computacionais que têm a língua natural como objeto primário.” (NUNES, 2005, p.3) Nesta pesquisa, está relacionada à tradução automática através de processamento eletrônico, pois a linguística computacional (PLN) lida com

[...] problemas relacionados à automação da interpretação e da geração humana em aplicações como Tradução automática, Sumarização Automática de Textos, Categorização Textual, Recuperação e Extração de Informação, Indexação Automática, entre outros. Além disso, também é seu objeto de pesquisa a criação de dicionários/léxicos e corpus linguísticos, análise morfossintática automática, análise semântico-discursiva automática etc. (LIMA, 2015, p.25)

Assim, neste capítulo abordaremos as definições de tradução e interpretação, as áreas centrais da linguística relacionada à língua de sinais, como a fonologia, morfologia, sintaxe e semântica da Libras. Como também, a tradução automática e sistemas de notações escritas de línguas de sinais.

## **2.1 Definição da língua de sinais e cultura surda**

Na história global sobre o surgimento das línguas de sinais, nos anos anteriores ao século XV, os surdos viviam à margem da sociedade. Somente depois foram vistos como potenciais educandos. No século XVI surgiram os primeiros educadores, sendo o monge beneditino Pedro Ponce de León o primeiro a desenvolver uma metodologia de educação de surdos, que incluía “[...] datilologia (representação manual das letras do alfabeto), escrita e oralização, e criou uma escola de professores surdos.” (GOLDFELD, 1997, p. 28) Sendo seguido por “[...] Juan Pablo Bonet (1573-1633), outro monge beneditino também da Espanha, que escreveu o livro *Reduccion de las letras y arte para enseñar á hablar á los mudos – alfabeto manual* (Madri, 1620).” (FELIPE, 2018, p. 194)

A partir dessas primeiras iniciativas, pode-se observar os períodos relacionados aos métodos de ensino, assim como as metodologias utilizadas, como o Oralismo (1485 - 1970), a Comunicação Total (1970 - 1980) e o Bilinguismo (a partir de 1980). Em 1815, um professor americano, Thomas Hopkins Gallaudet, foi para a Europa interessado em aprender sobre a educação de surdos. No ano de 1817, voltou aos Estados Unidos com um dos melhores alunos surdos da França. Fundou a primeira escola para surdos na América. Assim, a partir de 1821, surge a *American Sign Language* (ASL), principal idioma utilizado pelos surdos norte-americanos, começou a ser utilizada na escola. É provável que antes da ASL existisse uma

língua de sinais utilizada pelas comunidades surdas que deve ter sido complementada e influenciada pela Língua de Sinais Francesa (LSF), utilizada pelos professores que fundaram essa escola, sendo um deles francês (FELIPE, 2018).

Em 1860, o método oral passou a ganhar maior investimento, após o Congresso Internacional de Educadores de Surdos em Milão. Nesse evento decidiram, por votação e políticas públicas, que o Oralismo deveria ser a única forma utilizada na educação dos surdos, ficando proibida a prática da língua de sinais. O modelo oralista tem como base a ideologia de que a língua de sinais impedia que os surdos desenvolvessem a oralização, única forma de comunicação aceita na época.

Surgiram, no século XVII, várias metodologias oralistas com publicações trazendo orientações sobre a educação de surdos-mudos, como as de Ramiro de Carrion (1622) na Espanha, e do Juan Bulwer na Inglaterra. Outros professores, a partir dessas pesquisas, desenvolveram métodos próprios de ensino. (FELIPE, 2018)

Estes métodos objetivavam “[...] levar o surdo a falar e a desenvolver competência linguística oral, o que lhe permitiria desenvolver-se emocional, social e cognitivamente do modo mais normal possível [...]”. (CAPOVILLA, 2000, p. 102) Para que o surdo fosse integrado à comunidade ouvinte deveria ser oralizado, como uma reabilitação de fala para a normalidade característica da sociedade da época.

O Oralismo vê a surdez como uma deficiência a ser minimizada. Órgãos governamentais enviavam verbas generosas à aquisição de equipamentos e para projetos de formação de professores amadores, que muitas vezes atuavam como fonoaudiólogos, para os sujeitos surdos falarem e ouvirem. Essa forma de avaliação clínica perdurou até recentemente. As técnicas utilizadas no modelo oral eram o treinamento auditivo, estimulação auditiva para reconhecer e distinguir sons e ruídos, o desenvolvimento da fala, exercícios que flexibilizaram os órgãos envolvidos na fonação, e a leitura labial, treinamento para identificar palavras faladas com decodificação de movimentos orais. (PERLIN; STROBEL, 2008) As crianças começaram a receber apenas treinamento oral, passando a ser a prioridade na educação, deixando as outras disciplinas em segundo plano. Essa prática fez com que ocorresse uma queda no nível de escolarização dos surdos.

Esse método foi muito criticado devido à limitação, como a descontextualização das palavras, fazendo com que a linguagem fosse difícil e artificial.

Na década de 1960, surgiram estudos relacionados às línguas de sinais utilizadas pelas comunidades surdas. Ainda que fosse proibido o uso de sinais e gestos “[...] raramente se encontrava uma escola ou instituição para surdos que não tivesse desenvolvido, às margens do

sistema, um modo próprio de comunicação através dos sinais.” (LACERDA, 1998, p. 5) Surgindo assim, novas propostas pedagógico-educacionais relacionadas à educação dos surdos, a Comunicação Total.

A comunicação total “[...] é a prática de usar sinais, leitura orofacial, amplificação e alfabeto digital para fornecer inputs linguísticos para estudantes surdos, ao passo que eles podem expressar-se nas modalidades preferidas” (Stewart 1993, p. 118 apud LACERDA, 1998, p 6). Tendo como objetivo permitir que a criança desenvolva uma comunicação efetiva em seu convívio diário. Podendo utilizar sinais utilizados pela comunidade surda, sinais gramaticais e elementos visuais. Comparado ao Oralismo, a Comunicação Total apresentou melhores resultados em processos escolares, como compreensão e comunicação. No entanto, foram observados problemas relacionados à escrita, fazendo com que seus usuários não alcançassem níveis acadêmicos adequados a sua idade. Pois, “[...] como a Libras não era ensinada nem utilizada como língua de instrução, essa proposta educacional resultou em monolinguismo, uma vez que a Língua Portuguesa era a língua de instrução e a língua ensinada para os surdos.” (FELIPE, 2018, p. 203) Apesar disso, a comunicação total permitiu um contato maior com os sinais, garantindo “[...] que os surdos se dispusessem à aprendizagem das línguas de sinais, externamente ao trabalho escolar.” (LACERDA, 1998, p. 6)

Simultaneamente, surgiram estudos relacionados à educação bilíngue. A proposta de uma educação bilíngue teve início na década de 1980, com o desenvolvimento de políticas educacionais que representavam uma pedagogia mais progressista. Para Nunes et al (2015), a “[...] escola bilíngue seria o espaço de socialização, de construção de uma identidade positivada, de acesso ao conhecimento e a uma comunicação significativa para os que costumeiramente são “sem-lugar”” (p. 542). Kubaski e Moraes (2009) classificam o bilinguismo como uma abordagem educacional que procura a capacitação do surdo em duas línguas, a “[...] língua de sinais e a língua da comunidade ouvinte” (p. 3414) O modelo inclusivo de educação do surdo traz o bilinguismo que “[...] envolve a competência e o desempenho em duas línguas, podendo ser individual ou grupal” (QUADROS, 2009). O bilinguismo concebe a Língua de Sinais como primeira língua - L1 para os surdos (considerados para as pessoas surdas) e o português em sua forma escrita como segunda língua - L2 também para os Surdos. Dessa forma, havia a presença (ainda que mínima) da Língua de Sinais nas escolas regulares, e certamente havia nas escolas especiais para surdos.

Com a evolução dos estudos da área surgiram educadores que desenvolveram escolas especializadas para os surdos. Felipe (2018) traz, em sua pesquisa sobre a periodização da educação de surdos, a delimitação dos “[...] pressupostos filosóficos e pedagógicos, as

propostas educacionais, os métodos, os tipos de escolas, os resultados esperados e alcançados pelas políticas para educandos surdos ao longo de vários séculos.” (p. 191) Na Tabela 1, pode-se observar os períodos político-educacionais apresentados por Felipe (2018).

**Tabela 1** - Periodização da educação de surdos no Brasil

| Períodos                        | 1485 - 1880  | 1880 - 1970   | 1970 - 1980  | 1970 - 1980  | 1980 - 1996   | 1996 - 2018  |
|---------------------------------|--|---|--|--|---|--|
| <b>Pressupostos filosóficos</b> | Teoria da Potência e Ato   | Rousseau, Pestalozzi e Froebel  | Teoria liberal   | Teoria liberal   | Teorias pós-modernas<br>Teorias críticas  | Teoria liberal   |
| <b>Pressupostos pedagógicos</b> | <b>Pedagogia cristã</b><br>Processo de atualização da potência humana a partir de critérios éticos                         | <b>Pedagogia da existência</b><br>Adequação às leis naturais que regulam a vida;<br>Critérios de avaliação técnicos | <b>Pedagogia liberal: Tradicional</b><br>Adaptar o aluno à escola, aos conteúdos, aos métodos, à relação professor-aluno   | <b>Pedagogia liberal: Renovadora</b><br>Adequar às necessidades individuais ao meio social; integração pela experiência e vivência dos educandos | <b>Pedagógica progressista: Libertária</b><br>Propostas crítico-sociais dos conteúdos – realidades sociais; apropriação do saber, perspectivas linguística, sociolinguística e antropolinguística | <b>Pedagogia liberal: Tecnicista</b><br>Modelar o comportamento humano e integrar os alunos no sistema social global por meio de recursos tecnológicos |
| <b>Propostas educacionais</b>   | <b>Oralismo</b>  |   | <b>Ensino integrado</b>  | <b>Comunicação total</b>   | <b>Educação Bilíngue para educandos surdos</b>  | <b>Educação Bilíngue inclusiva</b>   |
| <b>Métodos</b>                  | <b>Método combinado e utilização da comunicação dos surdos</b><br>Bonet, L’Epée, Sicard, Bebian, Ballesteros e Villabrilie | <b>Métodos oralistas</b><br>Perdocini, Verbo-tonal, etc.  | <b>Método padronizado</b><br>Crianças ouvintes e surdas em salas regulares e Métodos oralistas – em classe/escola especial | <b>Método bimodal</b> (Cicone, Anete)<br>Método bimodal em classe/escola especial  | <b>Educação Bilíngue</b><br>Língua patrimonial – L1 e língua majoritária – L2;<br>Bilinguismo e biculturalismo em escolas para educandos surdos   | <b>Educação Bilíngue transitória</b><br>Alunos com deficiência auditiva na Educação Regular e na Educação Especial no contraturno                      |

| <b>Tipos de escola</b>          | <b>Primeira escola para surdos no Rio de Janeiro</b>   | <b>Escola Especial</b>   | <b>Escolas Regular, Especial e Classe Especial</b>                                  | <b>Escola Especial e Classe Especial</b>  | <b>Escala Bilíngue ou Classe Bilíngue somente para educandos surdos</b>   | <b>Escola Regular e AEE em contraturno para alunos com deficiência auditiva</b>  |
|---------------------------------|--|--|---|---|---|--|
| <b>Implicações linguísticas</b> | 1º momento: Desenvolvimento da Libras e ênfase na escrita da LP<br><br>2º momento: Ênfase na Língua Portuguesa – Modalidades oral e escrita. Difusão da Libras | Educação monolíngue<br><br>Língua Portuguesa;<br><br>Ênfase: modalidade oral | Educação monolíngue;<br><br>ênfase na Língua Portuguesa – Modalidade oral e escrita | Educação monolíngue;<br><br>possibilidades de várias formas de comunicação, mas ênfase na Língua Portuguesa – Modalidade oral e escrita | Educação Bilíngue;<br><br>Aquisição/ensino da Libras – L1, também língua de instrução;<br><br>Ensino da Língua Portuguesa – L2,<br><br>Modalidade escrita | Libras como acessibilidade – língua de instrução, por meio de intérpretes;<br><br>Ensino de Português – L1<br><br>Modalidade escrita;<br><br>Libras no AEE |
| <b>Resultados esperados</b>     | Aprendizagem da LP – modalidade escrita e da Libras utilizada por repetidores  | Aprendizagem da LP – modalidades falada e escrita                            | Aprendizagem da Língua Portuguesa   | Aprendizagem da Língua Portuguesa   | Aquisição/Aprendizagem de duas línguas: Pessoas bilíngues e biculturais   | Aprendizagem da Língua Portuguesa – modalidade escrita   |
| <b>Resultados alcançados</b>    | Sem registro   | Maioria dos educandos surdos não foi bem-sucedida                            | Maioria dos educandos surdos não foi bem-sucedida                                   | Maioria dos educandos surdos não foi bem-sucedida   | Não houve tempo hábil para se obter resultados  | Maioria dos educandos surdos tem protestado contra a inclusão sem condições adequadas  |

Fonte: Felipe (2018, p. 192)

A educação dos surdos passou por várias etapas e processos, como observado na Tabela 1 (Periodização da educação de surdos no Brasil). E, no Brasil, com as leis atuais sobre a inclusão da pessoa com deficiência e língua de sinais, as escolas de Surdos que “[...] adotaram o Bilinguismo como metodologia de ensino puderam oferecer aos seus alunos melhores condições de acesso ao conhecimento do que as escolas inclusivas.” (MONTEIRO, 2006, p. 301)

No Brasil, a educação de surdos teve início em 1855. Quando D. Pedro II trouxe para o país um professor surdo francês, chamado H Ernest Huet, para educar duas crianças surdas. Em

26 de setembro de 1857, fundou-se no Rio de Janeiro a primeira escola para surdos no Brasil, o Instituto Nacional de Surdos-Mudos (INSM), hoje, Instituto Nacional de Educação dos Surdos (INES) e “[...] nesse mesmo dia comemora-se o Dia Nacional dos Surdos no Brasil. (ALMEIDA; ALMEIDA, 2012, p.316) Em 1873, foi editado o livro de Flausino José de Gama. Esse livro, esquecido por séculos, foi reencontrado por Felipe (1998) e descrito no apêndice de sua tese de doutorado. Uma nova pesquisa sobre esse livro (SOFIATO; PAGNEZ, 2014) analisando também esse livro, o comparou com outro da Língua de Sinais Francesa que é o original desse livro editado aqui no Brasil.

Por esse livro é possível também concluir que nesse instituto se utilizava uma língua de sinais, seguindo o padrão mundial, mas a Língua de Sinais passa a ser proibida no Brasil, como também no INES. Em 1911,

[...] seguindo os passos internacionais que em 1880 no Congresso de Milão proibira o uso da Língua de Sinais na educação de surdos, estabelece-se que o INSM passaria a adotar o método oralista puro em todas as disciplinas. Mesmo assim, muitos professores e funcionários surdos e os ex-alunos que sempre mantiveram o hábito de frequentar a escola, propiciaram a formação de um foco de resistência e manutenção da Língua de Sinais. (RAMOS, 2004, p. 9)

Em 1930, foi fundada a primeira Associação Brasileira de Surdos-Mudos, por um pequeno grupo de estudantes do INES, no entanto não possuía estatuto. Em 1953, foi fundada outra Associação, com ajuda da professora Dona Ivete Vasconcelos. Essa associação era composta por Surdos da Congregação de Surdos do Rio de Janeiro. Como os ex-estudantes voltavam para suas cidades no Brasil, foi fundada em 1954 a segunda Associação de Surdos-Mudos de São Paulo. Em 1956, foi fundada a Associação de Surdos de Belo Horizonte em Minas Gerais, sendo esta a terceira no país. Hoje há diversas federações e associações de Surdos no país, com competições e festas comemorativas, esses encontros contribuem “[...] para a preservação da Língua de Sinais da Identidade Cultural Surda e conseqüentemente para o fortalecimento da luta pelos direitos dos Surdos.” (MONTEIRO, 2006, p. 297)

Na educação, a língua de sinais continuou sendo utilizada em sala de aula. No entanto, em 1957 a língua de sinais foi proibida, oficialmente, em sala de aula. E medidas foram tomadas, “[...] como o impedimento do contato de alunos mais velhos com os novatos foram tomadas, mas nunca o êxito foi pleno e a LIBRAS sobreviveu durante esses anos dentro do atual INES.” (RAMOS, 2004, p. 9) Assim, no ano de 1969, foi realizada a primeira tentativa em registrar a Língua de sinais utilizada no país, pelo missionário americano Eugênio Oates, com a publicação de um dicionário de sinais - Linguagem das mãos.

Nos anos 70 foi fundada a Federação Nacional de Educação e Integração do Deficiente Auditivo (FENEIDA), por profissionais ouvintes relacionados à surdez. No entanto, o conhecimento acerca da existência dessa federação não era conhecido pelos surdos. Já,

[...] em meados da década de oitenta, foi introduzido no Brasil, o paradigma “Educação Bilíngue para Surdos”, quando começamos a mostrar, linguisticamente, que a Língua de Sinais dos Centros Urbanos Brasileiros – LSCB – era, de fato, uma língua de outra modalidade, ou seja, gestual-visual, diferenciando-a das línguas de modalidade oral-auditiva. (FELIPE, 2018, p. 203)

Na Constituição de 1988, há artigos específicos sobre a inclusão da Pessoa com Deficiência - PcD. Desse modo, a proposta de uma educação inclusiva teve início na década de 90, sustentada pela Declaração de Salamanca de 1994, que, segundo ela

[...] as escolas regulares inclusivas seriam o meio eficaz de combater a discriminação, já que a escola consistiria em um ambiente que acolheria e faria a socialização dessas pessoas que há tanto estavam esquecidas e segregadas do meio social. Sendo a escola espaço de socialização e de discussão para acabar com a discriminação, as pessoas poderiam mostrar que apesar de suas limitações físicas, sensoriais, intelectuais ou motoras, são capazes de aprender tanto quanto pessoas comuns sem deficiência. (SILVA, 2014, p. 14)

Assim, a formação de escolas bilíngues para surdos foi substituída pela criação de escolas bilíngues inclusivas, ocasionando o fechamento de escolas e classes para Surdos com a proposta bilíngue mesmo antes de sua implementação. (FELIPE, 2018) No entanto,

[...] para a implementação da proposta de Educação Bilíngue Inclusiva nas décadas passadas, o Ministério da Educação – MEC –, por meio da antiga Secretaria de Educação Especial – SEESP –, atendendo ‘a reivindicação dos Surdos, financiou o Programa Nacional para apoio à Educação dos Surdos e o Programa Nacional Interiorizando a Libras. (FELIPE, 2018, p. 206)

É preciso visualizar alguns pontos da historicidade da educação para entender melhor a trajetória da Língua de Sinais e seu reconhecimento. Dentre os artigos, Leis e Decretos, tem-se o Decreto 7.612/11 relacionado aos direitos referentes à acessibilidade, prioridade e benefícios das PcD nos setores da sociedade. Pode-se notar que, ao longo dos tempos, os direitos dos surdos, foram sendo reconhecidos pela sociedade e a legislação em seu favor da garantia desses direitos foi se ampliando.

Atualmente, os direitos linguísticos dos surdos estão protegidos por Lei, mas nem sempre foi assim. Pode-se dizer que um dos primeiros benefícios foi a possibilidade de matricular os estudantes surdos em escolas próximas às suas casas. No entanto, essa estratégia

política não foi uma proposta bem recebida pela comunidade surda, visto que as escolas de Surdos acabaram sendo fechadas. Pois, os surdos foram encaminhados para as escolas próximas e inseridos em classes consideradas “inclusivas”, mesmo sem nenhum apoio a uma educação bilíngue. Para Zuffo (2010), a

[...] inclusão social faz parte de políticas públicas que buscam a transformação, a interação, o crescimento mútuo, o desenvolvimento pessoal dos portadores de necessidades especiais, por meio da integração ao convívio na vida social, econômica, política, assegurando-lhes os direitos como cidadãos. (ZUFFO, 2010, p.3)

No entanto para essa inclusão social, os surdos tem o direito de ter uma escola específica que atendam às suas especificidades linguísticas e, por isso, a simples inclusão em uma classe com discentes ouvintes desconsideram as lutas das comunidades surdas há décadas, exigido escolas bilíngues.

Debater sobre os direitos do surdo nos compele a apresentar as Leis, presentes em nossa Constituição Federal (BRASIL, 1988), artigos específicos sobre a inclusão de pessoas com deficiência e relacionados à língua de sinais. Como a Lei nº 10.436/02 que dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências, resultante das lutas protagonizadas pelos Surdos no Brasil, sendo essas conquistas positivas para a comunidade surda. (CASSIANO, 2017) E a Lei nº 10.845/04 que institui o programa de complementação ao atendimento educacional especializado às pessoas portadoras de deficiência<sup>10</sup>, e dá outras providências.

A Lei nº 12.319/10 regulamenta a profissão do Tradutor e Intérprete Língua Brasileira de Sinais - LIBRAS (TILS), assim como suas atribuições, mas Santos (2018) salienta que foram necessários “[...] 22 anos de mobilizações na luta pela regulamentação da profissão TILS, no Brasil.” (p. 71) A Lei estabelece atribuições no exercício das competências do tradutor e intérprete, além de garantir e preservar os valores éticos e humanos no que envolve o exercício da profissão, ressaltando que o tradutor e intérprete de Libras precisa interpretar a língua de sinais e o português de forma simultânea ou consecutiva, além de ser proficiente em tradução e interpretação tanto de Libras quanto da Língua Portuguesa.

Essa Lei surgiu com a necessidade de possibilitar o acesso das pessoas surdas, surdo-cegas, com deficiência auditiva e usuárias da Libras aos espaços sociais, educação, saúde, lazer,

---

<sup>10</sup> Pessoa com deficiência (PcD), este conceito foi “[...] constitucionalmente adotado pelo Brasil por força da ratificação da Convenção Internacional da ONU sobre os Direitos das Pessoas com Deficiência, transcende o aspecto meramente clínico e assistencialista que pautava a legislação anterior. Ressalta o fator político para que se reconheça a necessidade de superarem-se as barreiras sociais, políticas, tecnológicas e culturais.” (FONSECA, 2012, p. 53)

entre muitos outros, promovendo informação e comunicação. Somente a Lei por si só não oferece garantias trabalhistas aos profissionais tradutores e intérpretes da Libras, apenas reconhece e regulamenta, frente a isso, muitos grupos se uniram e buscaram associação aos sindicatos de Intérpretes de Línguas<sup>11</sup>, a fim de estabelecer diretrizes para os profissionais, sempre de acordo com a Lei nº 12.319/10.

As dificuldades em acesso a conteúdo online e em instituições de atendimento ao público, que, apesar da existência de leis específicas que obriga lugares públicos a presença de um intérprete de Libras, alguns não consideram a disponibilidade de um profissional como fundamental. Entre as instituições de ensino que buscam o acompanhamento de intérpretes, estão as escolas de ensino fundamental I e II, ensino médio, assim como as de ensino técnico e superior. E instituições com atendimento ao público, como museus, hospitais, Organizações Não Governamentais (ONGs), entre outros. Em cinemas e teatro não há a presença de um apoio linguístico, apesar de alguns artistas buscarem a integração deste público com shows inclusivos, essa prática ainda é minoria, são alguns casos específicos.

## **2.2 Interpretação e tradução da Libras**

Para as línguas de sinais há duas formas de tradução, a escrita e a oral, utilizadas pelos tradutores/intérpretes de Libras. Diferente da interpretação, na tradução textual, o tempo não deve influenciar. Nesse sentido, entende-se que a cultura (visão do mundo) não deve interferir no resultado da tradução, deve-se buscar uma isenção de preconceitos. Os surdos encontram dificuldades em se comunicar, esse problema pode ocorrer pela “não divulgação” ou falta de incentivo em aprender língua de sinais. Apesar de que, o Ensino de Libras é Lei para a os cursos de formação de Professores, entre outros (Decreto 5.626/05). No entanto, infelizmente, não são todos os ambientes sociais que dispõem de uma pessoa qualificada para a interpretação. Em momentos em que o surdo se encontra só, sua dificuldade na compreensão da língua oral é evidenciada.

Assim, a tradução automática realiza uma quebra de barreira na comunicação, podendo ser empregada como uma ferramenta acessível para o sujeito surdo. Em termos de comparação, podemos citar o Tradutor do Google como uma ferramenta de tradução automática baseada em exemplo com um desempenho, na maioria dos casos, satisfatório.

---

<sup>11</sup> <http://sintra.org.br>

Assim como qualquer língua ou forma de comunicação, sendo elas oral, visual e corporal. Pode-se afirmar que “a língua de sinais brasileira apresenta uma estrutura gramatical rica e é usada pelos surdos brasileiros para expressar ideias, pensamentos, sonhos, arte e histórias e reproduzem discursos [...]” (QUADROS, 2003, p. 92), dessa forma, a descrição da língua de sinais deve seguir determinadas regras. Prezando por “[...] uma tradução que não seja palavra-por-palavra e, sim, sentido-por-sentido” (HUMBLÉ, 2005, p. 236), englobando todas as informações contidas no texto, ou no caso do intérprete de Libras, na sinalização.

Podemos diferenciar tradução e interpretação, sendo que a tradução é realizada a partir de pesquisas documentadas. O profissional desenvolve sua tradução tendo acesso ao texto original, utilizando de normas padronizadas, adaptações homogêneas e mais formais. A tradução “[...] pode ser realizada fora do contexto a que se destina, já que o texto fonte está pronto e o público terá acesso ao texto alvo após sua conclusão.” (RODRIGUES; SANTOS, 2018, p. 2) Na interpretação simultânea, por outro lado, não há padronização, pois, o texto é apresentado apenas uma vez, fazendo com que o intérprete tenha pouco tempo para receber, processar e emitir a informação. Dessa forma, muitas vezes não é possível fazer uma revisão prévia da tradução. No entanto, para a interpretação em eventos internacionais que implicam participantes, cujas línguas são distintas, a organização desses eventos contrata tradutores-intérpretes profissionais e solicitam as palestras escritas previamente para esses profissionais poderem estudar previamente o texto a ser interpretado simultaneamente às palestras, o que não tem acontecido com relação à interpretação (tradução simultânea) para a Libras ou de Libras para o português.

No entanto, em palestras e eventos, em respeito ao intérprete, deve-se enviar o texto previamente, para análise e adaptações. Com relação ao ambiente escolar, o intérprete deve se reunir com os professores no momento do planejamento pedagógico, para antecipadamente discutirem sobre os conteúdos das aulas.

Durante a interpretação, “[...] o intérprete realiza “explicações”, como se fossem notas de rodapé no ato da interpretação, para contextualizar um fato [...]” (MARCON, 2012, p. 248). Na tradução, por outro lado, essas referências não são apresentadas, sendo necessário a criação de um *corpus* com regras elaboradas de acordo com a variedade contextual da língua portuguesa e da Libras, isso também se aplica à interpretação *in loco*.

### 2.3 Linguística da Libras

Nos estudos linguísticos “o exercício da linguagem repousa numa faculdade que nos é dada pela Natureza, ao passo que a língua constitui algo adquirido e convencional” (SAUSSURE, 1969, p. 17). Na concepção de Saussure, a tomada da língua ocorre sempre por processos de linguagem que quando articulados e envolvidos em um sistema sógnico adquire significado e convenção, passando a atuar na cultura comunicacional do homem através de atos individuais e exteriores de fala. Chomsky (1980), afirma que “os princípios que governam a estrutura e o emprego da linguagem são universais, de acordo com uma necessidade biológica e não histórica”. Dessa forma, “certas realizações intelectuais como a aprendizagem da linguagem revelam estritamente uma capacidade biológica determinada” (CHOMSKY, 1980, p. 9).

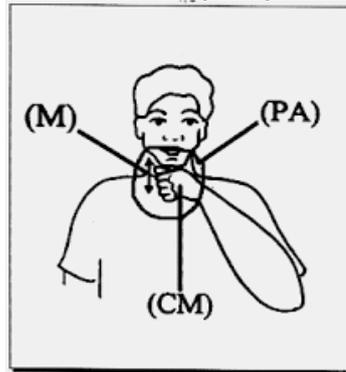
A Libras, assim como em qualquer outra língua, possui uma estrutura gramatical específica e própria. Se distinguindo das línguas oral-auditivas por se tratar de uma língua gestual-visual. Sua gramática é

[...] constituída a partir de elementos constitutivos das palavras ou itens lexicais e de um léxico (o conjunto das palavras da língua) que se estruturam a partir de mecanismos morfológicos, sintáticos e semânticos que apresentam especificidade, mas seguem também princípios básicos gerais. (BRITO, 1997, p. 23)

Assim, a linguística pode ser dividida em áreas que estudam aspectos da linguagem, essas áreas são: fonologia, morfologia, sintaxe e semântica e pragmática. A Libras, com suas regras gramaticais específicas, apresenta variações de acordo com os aspectos apresentados a seguir.

### ***2.3.1 Aspectos fonológicos***

Nas línguas orais, como o Português Brasileiro (pt-BR), a fonologia tem como objeto de estudo a “[...] a forma sistemática como cada língua organiza os sons [...]” (BISOL, 2005, p. 11), esta organização apresenta o fonema como unidade mínima sonora. Na Libras, uma língua gesto-visual, a fonologia está relacionada ao uso das mãos e do corpo, como também do espaço da sinalização em concordância com as propriedades visuais e manuais no contexto da comunicação. (SILVA, 2020) As línguas de sinais são produzidas pelas mãos, utilizando movimentos do corpo e da face. A configuração de mão (CM), movimento (M) e locação o ponto de articulação (PA) são considerados os fundamentos (parâmetros fonológicos) da Libras, exemplificado na figura abaixo.

**Figura 1 - Parâmetros principais da Libras**

Fonte: FERNANDES; STROBEL, 1998, p. 7

A Libras possui cinco parâmetros fonológicos, produzidos pelas mãos, movimento do corpo e expressões faciais. São a Configuração de mão, o Movimento, a Locação ou Ponto de articulação, a Orientação ou direcionalidade e a Expressão facial e corporal.

#### - **Configuração de mão**

A configuração de mão (CM) “[...] são as formas em que a mão ou as mãos são posicionadas para execução do sinal em Libras.” (FARJADO et al., 2015, p. 189). Antunes (2015) conceitua a CM como uma “[...] forma assumida pela mão enquanto um sinal específico é articulado, isto é, na disposição que os dedos assumem na mão dominante ou em ambas as mãos na execução do sinal.” (P. 38) Felipe (1997) caracteriza a CM como as “[...] formas das mãos, que podem ser da datilologia (alfabeto manual) ou outras formas feitas pela mão predominante (mão direita para os destros), ou pelas duas mãos do emissor ou sinalizador.” (FELIPE, 1997, p. 82) Na Libras, segundo Felipe et al. (2005), há 64 configurações de mão, como pode-se observar na figura 2.

**Figura 2 - Configuração de mão (CM)**

Felipe, Tanya. Dicionário da Libras. Versão 2.0 - 2005.

Fonte: FELIPE, MONTEIRO, 2007, p.28

### - Movimento

O movimento (M) é o “[...] deslocamento no espaço durante a realização do sinal. [...] Apresentando as seguintes características: tipo, modo, direção, intensidade e frequência.” (FARJADO et al., 2015, p. 190) Os sinais podem ou não apresentar movimento, esses movimentos direcionais podem ser categorizados como unidirecionais, bidirecionais e multidirecionais. Há também o modo do movimento, como a velocidade e frequência. (KARNOPP, 2006)

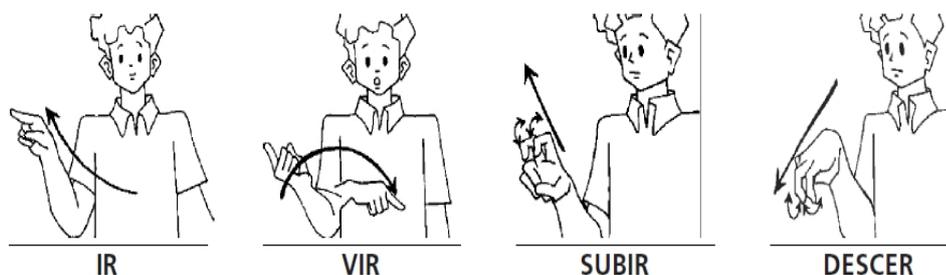
### - Locação ou Ponto de articulação

Com relação aos parâmetros de um sinal, o ponto de articulação (PA) é o local “[...] onde incide a mão predominante configurada, podendo esta tocar alguma parte do corpo ou estar em um espaço neutro vertical (do meio do corpo até à cabeça) e horizontal (à frente do emissor).” (RAMOS, 2004, p. 11) Os pontos são os locais, próximos ou não, ao corpo onde os sinais são realizados, compreendendo todos os “[...] pontos articulatorios dentro do raio de alcance das mãos, no qual os sinais são articulados (cabeça, mão, tronco, braço e espaço neutro)” (PASSOS, 2014, p. 45)

### - Orientação/ Direcionalidade

A orientação ou direcionalidade (OD) está relacionado aos outros parâmetros da Libras, já mencionados. É necessário observar a “[...] direção para a qual a palma da mão aponta quando o sinal é produzido.” (FARJADO et al., 2015, p. 189) Pois alguns sinais podem apresentar a configuração de mão, ponto de articulação e movimento iguais e direcionalidade diferente. Para Felipe (1997) “[...] os sinais podem ter uma direção e a inversão desta pode significar ideia de oposição, contrário ou concordância número-pessoal, [...]” (FELIPE, 1997, p.82) Assim, essa alteração pode interferir no significado do sinal. Pode-se observar essa alteração nos verbos IR-VIR e SUBIR-DESCER, apresentados por Felipe e Monteiro (2007)

**Figura 3 - Direcionalidade**



Fonte: FELIPE, MONTEIRO, 2007, p.23

### - Expressão facial e corporal

As expressões não-manuais podem ser caracterizadas como expressão facial (EF) e expressão corporal (EC). As corporais são os movimentos feitos com o corpo,

[...] as expressões faciais na Libras podem ser gramaticais, relacionadas à estrutura, morfologia (por exemplo, elas marcam grau de intensidade, tamanho) e sintaxe (negação, interrogação, ênfase). [...] usadas para expressar sentimentos (alegria, angústia, ansiedade, sofrimento). (GUIMARÃES et al., 2018, p.14)

As expressões são fundamentais para a comunicação através dos sinais, pois há “[...] palavras que só podem ser compreendidas pela expressão facial, porque não usam as mãos para representação.” (FARJADO et al, 2015, p. 190)

### 2.3.2 Aspectos Morfológicos

A morfologia é a ciência que estuda a formação das palavras, a estrutura interna das palavras ou sinais, suas unidades mínimas são os morfemas (QUADROS, 2004). Este estudo apresenta a flexão e a derivação como áreas de investigação. No processo de derivação de sinais, há a mudança de um substantivo para um verbo, essas variações podem ser observadas no padrão de movimento. Na figura 4, que representa os sinais SENTAR e CADEIRA, podemos observar que no primeiro sinal o movimento é único e prolongado, já no segundo sinal o movimento é repetido.

**Figura 4** - Sinal de ‘sentar’ e ‘cadeira’



Fonte: QUADROS & KARNOPP, 2004, p. 97

No processo de composição, a junção de dois sinais pode criar um novo significado. Como a união dos sinais CASA + ESTUDAR, que sinalizados juntos em sequência formam o sinal de ESCOLA.

Na Libras, os sinais sofrem modificações da forma do sinal para incorporar informação gramatical, como a incorporação de numeral, negação, localização, formato e intensidade. Essa

incorporação de numeral está relacionada aos numerais referentes a ordens sequenciais (números ordinais) e tempo (minuto, hora, dia, mês e frequência), podendo ser aplicada aos pronomes pessoais. A incorporação da negação é formada pela junção de sinais não-manuais, como a negação pela movimentação da cabeça.

Sobre a incorporação da localização, pode-se citar os verbos direcionais. Assim, a “[...] realização desses verbos consiste em mover a(s) mão(s) de um ponto no espaço de sinalização associado a um de seus argumentos em direção ao ponto associado a outro.” (XAVIER; NEVES, 2016, p. 135) E há a concordância com a pessoa do discurso (pronomes pessoais), como exemplo temos a sinalização dos verbos PERGUNTAR e AVISAR, que podem variar sua localização de acordo com a orientação inicial e final do sinal. Há também, os sinais locativos que “[...] incorporam a localização associada à(s) entidade(s) a que se referem, encontram-se sinais que variam sua forma por poderem ser totalmente realizados no lugar associado a seu referente.” (XAVIER; NEVES, 2016, p. 138)

A incorporação do formato pode ocorrer através de configuração de mão, para um tipo de verbo se adequar à sinalização de um objeto como seu tamanho e formato. A incorporação de intensidade pode ser expressa por meio de mudanças na configuração de mão, movimento e expressões não-manuais. Como exemplos de advérbios de intensidade, pode-se citar a palavra “rapidamente” com movimento repetido e acelerado, e do intensificador “muito” que apresenta um movimento lento e alongado para a frente do emissor. (FELIPE, 2006)

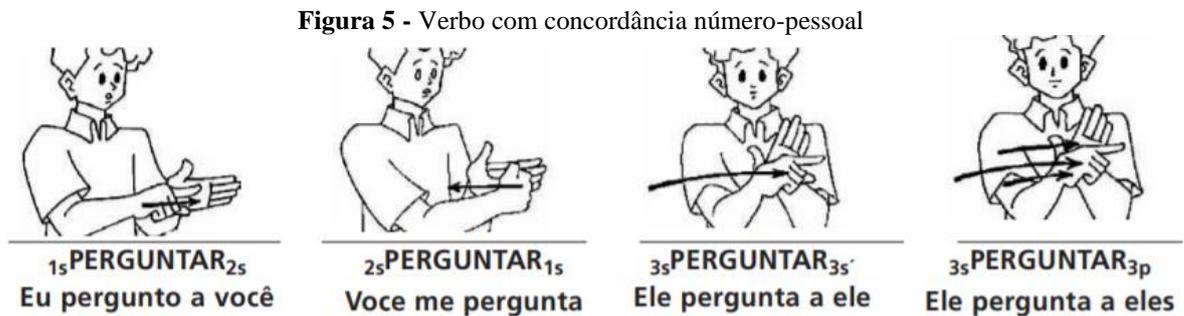
### **2.3.3 Aspectos sintáticos**

Na Libras, ao analisar sua sintaxe deve-se considerar o espaço onde o sinal está sendo executado, pois as relações sintáticas utilizam os sistemas pronominal e nominal para este propósito. (LIMA, 2015) Sua sintática apresenta aspectos estruturais específicos, como a concordância verbal “[...] o estabelecimento nominal e o uso do sistema pronominal [...]”. (QUADROS; KARNOPP, 2004, p. 127) Em relação aos tipos de verbos na Libras, Quadros e Karnopp (2004) os destacam como verbos com concordância e sem concordância (simples). Assim, os verbos “[...] sem concordância são os que não flexionam em pessoa e número, sendo que alguns podem flexionar em aspecto<sup>12</sup>.” (SOUSA, 2010, p. 99) Os verbos com concordância “[...] são os que flexionam em pessoa, número e aspecto.” (QUADROS; KARNOPP, 2004, p. 201) Já Felipe (1998) classifica os verbos em gênero, número-pessoal e locativo, a partir do seu

---

<sup>12</sup> “[...] a flexão de aspecto verbal se dá pela mudança na frequência ou na velocidade dos movimentos, ocorrendo diferenças entre marcas para os aspectos durativo, distributivo e contínuo.” (FINAU, 2004, p. 42)

sistema de flexão. A figura 5, pode-se observar o verbo PERGUNTAR em concordância número-pessoal de acordo com a execução da ação



Fonte: FELIPE, MONTEIRO, 2007, p.168

A respeito da estrutura sequencial das palavras nas frases, na Libras podem ser expressas de diversas formas, dependendo do uso da ênfase que o usuário queira dar. Na construção do *corpus* bilíngue em questão, estamos utilizando como base para a estrutura frasal a ordem básica gramatical o SVO de acordo com Quadros (1999), sendo S para Sujeito, V para Verbo e O para Objeto. Como podemos observar no exemplo:

**JOÃO GOSTA DE FRUTA**  
 ↙        ↓        ↘  
**S            V            O**

Para Cotovicz et al (2018) os “verbos direcionais possuem uma ordem fixa de organização das sentenças (SVO). Verbos não-direcionais possuem uma maior liberdade na ordenação das sentenças.” (p.32) Assim, dependendo da situação - contexto, existem as formas Objeto, Sujeito e Verbo (OSV) e Sujeito, Objeto e Verbo (SOV), derivadas da ordem básica gramatical, há o fenômeno da topicalização que se refere a uma “[...] forma de produção [...] topicalização ou tópico-comentário.” (SOUSA, 2010, p. 99) Sendo marcada por expressões não manuais, como franzir ou elevar as sobrancelhas.

#### **2.3.4 Aspectos semânticos e pragmáticos**

Na Libras, também acontece o fenômeno da homonímia, podendo um sinal representar dois conceitos diferentes e o sentido do sinal pode ser determinado por intermédio do contexto, assim duas palavras diferentes, como, por exemplo, os sinais SÁBADO e LARANJA que apresentam configuração de mão, movimento e locação idênticas. (MÁXIMO; SOUZA, 2014)

Assim, a semântica procura esclarecer as propriedades do significado individual ou sentencial da palavra, e a pragmática compreende a relação entre linguagem e contexto, procurando descobrir as estratégias, intuições, formas e estruturas utilizadas pelo usuário da língua. (QUADROS; KARNOPP, 2004)

A pragmática estuda os fenômenos resultantes do uso da linguagem. dessa forma, pode-se entender que a semântica e a pragmática “[...] estudam os conceitos que construímos em nossas mentes quando estamos diante de um signo linguístico, seja ele uma palavra, uma sentença, [...] ou um texto.” (MCCLEARY; VIOTTI, 2009, p.4)

Na Libras, o significado linguístico é influenciado pelos aspectos sintáticos do verbo, que podem ser subdivididos em verbos com “[...] direcionalidade implícita, verbos com raiz “de”, verbos com raiz “para” e verbos “multidirecionais”. (SILVA, 2006, p. 265) Que numa perspectiva semântico-pragmática, a orientação espacial que pode apresentar movimentos direcionais relacionados ao enunciador do ato de fala. Esses verbos podem ser classificadores, pois possuem flexão para concordar com o objeto e locativo.

Neste sentido, os classificadores (CL) são um tipo de morfema que incorporam a ação verbal, sendo “[...] certas configurações de mãos que funcionam como morfemas que marcam certas características de um objeto nas línguas de sinais.” (FELIPE, 2002, p. 6) Os classificadores “[...] são afixos incorporados ao radical verbal ou nominal.” (BRITO, 2002, p. 20) São marcadores de concordância de gênero (pessoa, animal e coisas), podem estar no plural quando representam duas ou mais pessoas e animais, quando representam coisas apresentam características do objeto durante a ação verbal. De acordo com Felipe (1998) os verbos, classificados de acordo com as características morfológicas e semânticas, estão divididos em flexão número-pessoal, flexão para locativo e concordância de gênero.

## 2.4 Tradução automática

O desenvolvimento dos tradutores automáticos Português Brasileiro-Libras teve como objetivo facilitar a comunicação do surdo, proporcionando uma melhoria na comunicação e independência social. Mas, apesar de todos os avanços na área, não há um suporte idealizado para o aperfeiçoamento dos resultados alcançados através de uma tradução automática.

Na “década de 90 [...] começaram as pesquisas com as línguas de modalidade gesto-visual” (FELIPE, 2006, p. 155), possibilitando que fossem desenvolvidos Banco de Dados específicos da língua de sinais, com *corpus* de tradução específicos. Embora as bases tenham sido essencialmente compostas por vídeos e suas respectivas transcrições que foram analisadas

e incluídas suas regras fono-morfo-sintático-discursivas, a produção de um avatar que pudesse realizar frases gramaticais não foi possível. Neste estudo, é proposta uma base de dados com *corpus* contendo sentenças transcritas normalizadas.

Para o aperfeiçoamento dos resultados através da criação de um *corpus* de tradução, é necessário um sistema de combinação onde não “[...] considera uma única frase, o contexto de todo um *corpus* é levado em conta para alcançar uma alta qualidade de alinhamento.” (STEIN; SCHMIDT; NEY, p. 329, tradução nossa)<sup>13</sup> Sendo essa opção considerada uma das melhores formas de tradução, pois mantém a integralidade da informação inicial.

Esse armazenamento de dados deve ser pensado e escolhido para proporcionar a forma mais fácil de recuperação de informação. Diante disso, a base de dados foi empregada como um suporte para o desenvolvimento de um sistema de tradução baseado em exemplos, com representações em sentenças, abrangendo os métodos linguísticos característicos das linguagens de descrição.

Outro aspecto importante é que a base de dados desenvolvida contenha, em seu repertório, uma grande quantidade de sentenças. Elaboradas de acordo com as regras da língua de sinais, as proposições devem apresentar uma qualidade análoga, permitindo ao sistema a extração do conhecimento sem obstrução ou ruídos. Essa decodificação sem interferências permitirá que o tratamento de dados seja realizado de forma dinâmica e satisfatória.

As estratégias de tradução automática estatística (*Statistical Machine Translation - SMT*) geralmente requerem o uso de um *corpus* de tradução de tamanho significativo, contendo centenas de milhares de exemplos, podendo ser de imagens, vídeos e textos (frases). Desse modo, Horst (1999) afirma que “[...] em um sistema de aprendizado é necessário utilizar linguagens para representar exemplos ou instâncias, conceitos induzidos e conhecimento prévio sobre o domínio, denominado conhecimento de fundo ou teoria de domínio.” (HORST, 1999, p.10)

Chiu et al. (2006) apresenta sua preocupação com o alinhamento e inserção correta da gramática do chinês para sua língua de sinais<sup>14</sup> respectiva, pois o “[...] significado de um sinal é mais rico que o de uma palavra.” (CHIU et al., 2006, p. 28, tradução nossa)<sup>15</sup>. Comparada a tradução automática, o intérprete de Libras utiliza de conhecimentos próprios para realizar a

---

<sup>13</sup> “[...] considering a single sentence, the context of a whole corpus is taken into account in order to achieve high alignment quality.”

<sup>14</sup> *Chinese and Taiwanese Sign Language (TSL)*.

<sup>15</sup> “[...] meaning of a sign is richer than that of a word.”

tradução e troca de significados ou palavras quando necessário, sendo essa umas das diferenças mais marcantes entre a tradução automática e a interpretação.

Além do repertório de frases, a tradução deve seguir determinadas regras, sendo de extrema importância conhecer a estrutura gramatical das línguas envolvidas. Assim,

[...] nas línguas de sinais, essas línguas, como as línguas oral auditivas, também possuem estruturas fonológicas, que se constituem a partir da configuração de unidades discretas, feixes de traços distintivos, que são os quatro parâmetros: configuração de mão (CM), movimento (M), direcionalidade (Dir) e ponto de articulação (PA). (FELIPE, 2006, p. 202)

Além da expressão facial, considerado o quinto parâmetro da Libras, nos verbos direcionais, a concordância entre emissor e receptor deve ser respeitada, pois na “[...] língua de sinais brasileira, esses verbos têm que concordar como o sujeito e/ou com o objeto indireto/direto da frase.” (QUADROS; KARNOPP, 2004, p. 130)

Assim, determinando um padrão de descrição, retratadas em sentenças simples com regras linguísticas, permitirá à indução<sup>16</sup> computacional, resultados semânticos e estruturalmente semelhantes aos exemplos disponíveis na base de dados de tradução. A normalização padronizada, linguagem de descrição, deve seguir determinados requisitos específicos das línguas portuguesa e de sinais, além da interpretação contextual, um conjunto de regras aplicadas na criação de um *corpus* de Português Brasileiro-Libras, possibilitando o desenvolvimento e aperfeiçoamento da tradução do português para língua de sinais, empregando um treinamento bilíngue.

## 2.5 Sistemas de notações escritas de línguas de sinais

A escrita nem sempre foi como conhecemos. Suas características foram desenvolvidas de acordo com a necessidade, ferramentas e conhecimentos existentes. Dessa forma, surgiram formas escritas de registrar as línguas de sinais. Visto que a escrita da língua de sinais torna possível “[...] publicações na língua de sinais: Livros, revistas, dicionários e literatura.” (SUTTON, 2000, p.6)

A cultura, incluindo todos os meios de manifestação artística ou científica, faz parte do desenvolvimento social de uma comunidade, tornando necessária a criação de um método de

---

<sup>16</sup>“A indução é a forma de inferência lógica que permite obter conclusões genéricas sobre um conjunto particular de exemplos.” (MONARD; BARANAUSKAS, 2003, p. 40)

transcrição e registro desse processo intelectual. Dessa forma, a escrita de sinais surgiu como uma alternativa para proteger o desenvolvimento cultural e intelectual da comunidade surda.

Publicado em 1822, por Roch Ambroise Auguste Bébien, a notação *Mimographie* se trata de um método de transcrição fonética da língua de sinais. O Sistema de notação de Stokoe, foi o primeiro estudo sobre a linguística da língua de sinais publicado em 1965, direcionado apenas aos lingüísticos relacionados à língua de sinais americana - ASL.

Em seguida, baseado no sistema de Stokoe, o *Hamburg Notation System* (HamNoSys), publicado em 1984, distingue configuração de mão, orientação, locação, ação e componentes não manuais, além de possuir um sistema computacional. Próximo aos anos 90, Paul Jouison criou o sistema D'Sign. No entanto, após sua morte, a Dr.<sup>a</sup> Brigitte Garcia em sua dissertação mostrou que não se tratava de uma simples notação, mas sim um sistema muito elaborado capaz de trabalhar longas sequências, com frases sinalizadas e transcritas em D'Sign. (STUMPF, 2005)

Desenvolvida em 1996, a notação de François Neve, utiliza códigos que tornam “[...] possível uma numeração e um tratamento informático dos signos. A escritura é feita em colunas verticalmente de cima para baixo, em uma só coluna quando a mão dominante sinaliza.” (STUMPF, 2005, p. 48)

Criado em 1974, por Valerie Sutton, o sistema *SignWriting* (SW) é uma das formas de escrita de sinais e que passou por diversas mudanças para chegar à forma como é conhecida atualmente. Quadros (1999) classifica *SignWriting*<sup>17</sup> como “um sistema de escrita para escrever língua de sinais”. Assim, *SignWriting* é a representação escrita da língua gestual. Sua estrutura apresenta diversos parâmetros, entre eles a orientação de mãos, tipos de contatos, expressão facial e corporal.

No Brasil, a escrita de sinais se tornou conhecida a partir de 1996, através da criação de um grupo de trabalho dedicado ao sistema de escrita da língua de sinais.

A escrita de língua de sinais como ferramenta de comunicação, por sua natureza, permite construir um modelo teórico do real e expressar a coerência desse modelo inventando as relações entre os elementos e possibilitando elaborar um ponto de vista sobre o mundo, passando do conjuntural expresso pela língua de sinais para o estruturado, expresso pelo texto [...] (STUMPF, 2005, p. 42)

O Sistema de Escrita das Línguas de Sinais (ELiS) criado oficialmente em 2008 pela Dr.<sup>a</sup> Mariângela Estelita Barros, apresenta base alfabética, linear, que utiliza “[...] os 3

---

<sup>17</sup><https://www.signwriting.org/>

parâmetros propostos por Stokoe (Configuração de Mãos, Locação e Movimento) [...]” (AGUIAR; CHAIBUE, 2015, p.21), sendo posteriormente acrescentado a orientação e expressões não manuais.

Assim, pode-se perceber que os sistemas de escrita buscam abranger todas as características de línguas de sinais variadas, garantindo a consistência entre o pensamento e a transcrição do conhecimento, e o acesso à informação gerada e compartilhada na comunidade surda.

## **2.6 Considerações do Capítulo**

Neste capítulo, foi apresentada uma fundamentação teórica sobre a Libras, sua definição, aspectos linguísticos da tradução da língua de sinais. Sendo estes conceitos essenciais para o desenvolvimento deste trabalho.

Na seção 2.1, foram apresentadas as principais definições relacionadas à língua de sinais. A evolução das pesquisas na área da educação e o surgimento das primeiras instituições de ensino do surdo, mundial e nacional. Como também a criação de movimentos sociais, o reconhecimento e consolidação da língua de sinais no Brasil.

A seção 2.2 apresenta a definição de interpretação e tradução de Libras e suas diferenças, como a tradução textual. A seção 2.3, inclui os estudos linguísticos apresentando os aspectos fonológico, morfológico, sintático, semântico e pragmático. Na seção 2.4, foram apresentados os conceitos de tradução automática, características e estratégias relacionadas à tradução automática estatística e criação de *corpus* de tradução. A seção 2.5 apresentou os diferentes sistemas de notações escritas de línguas de sinais.

No próximo capítulo, será apresentada uma revisão sistemática da literatura sobre a tradução automática estatística, e importância do desenvolvimento de um *corpus* de tradução de línguas orais para línguas de sinais, sendo este o foco desta pesquisa. Esta revisão apresenta um mapeamento em relação ao tema proposto, identificando as limitações e contribuições presentes na literatura científica.

### 3 REVISÃO SISTEMÁTICA DA LITERATURA

A revisão sistemática da literatura busca identificar, no campo científico da pesquisa, um determinado conhecimento específico. Permitindo “[...] maximizar o potencial de uma busca, encontrando o maior número possível de resultados de uma maneira organizada.” (COSTA; ZOLTOWSKI, 2014, p. 54). Para Vosgerau e Romanowski os estudos que buscam realizar uma revisão de literatura permitem uma melhor “[...] compreensão do movimento da área, sua configuração, propensões teóricas metodológicas, análise crítica indicando tendências, recorrências e lacunas” (VOSGERAU; ROMANOWSKI, 2014, p.167).

Para Marconi e Lakatos a “[...] pesquisa bibliográfica não é mera repetição do que já foi dito ou escrito sobre certo assunto, mas propicia o exame de um tema sob um novo enfoque ou abordagem, chegando a conclusões inovadoras.” (MARCONI; LAKATOS, 2003, p. 183). Assim, designar *strings* de busca farão com que os artigos utilizados sejam o mais próximo possível do tema proposto.

#### 3.1 Protocolo de Revisão Sistemática

Para iniciarmos o refinamento da bibliografia existente sobre determinado assunto devemos definir o protocolo de revisão sistemática, buscando identificar as palavras-chaves essenciais para a pesquisa. Para esse estudo, visando alcançar o maior número de pesquisas que abordam a tradução automática como tema principal e, sequencialmente, utilizando bases de dados estatísticos com o desenvolvimento de um *corpus* de sentenças traduzidas, foram utilizadas três *strings* de busca para pesquisa dos trabalhos. Foram elas: “*statistical machine translation*”, “*statistical machine translation*” AND “*sign language*”, “*statistical machine translation*” AND “*sign language*” AND “*Libras*”.

As bases de dados utilizadas para as buscas por trabalhos relacionados estão apresentadas na Tabela 2.

**Tabela 2 – Bases de Dados**

| <b>BASE DE DADOS</b> | <b>ACESSO</b>    | <b>ENDEREÇO</b>   |
|----------------------|------------------|---|
| ARXIV                | <i>Gratuito</i>  | <a href="https://arxiv.org/">https://arxiv.org/</a>   |
| CAPEL                | <i>Gratuito*</i> | <a href="https://www.periodicos.capes.gov.br/">https://www.periodicos.capes.gov.br/</a>                       |
| JIDM                 | <i>Gratuito</i>  | <a href="https://periodicos.ufmg.br/index.php/jidm/index">https://periodicos.ufmg.br/index.php/jidm/index</a> |
| SCIELO               | <i>Gratuito</i>  | <a href="https://www.scielo.org/">https://www.scielo.org/</a>   |
| SCOPUS               | <i>Gratuito*</i> | <a href="https://www.scopus.com/home.uri">https://www.scopus.com/home.uri</a>                                 |

Fonte: Dados da pesquisa, 2019.

\* O conteúdo assinado está disponível para os acessos com IP identificado das instituições participantes, com acesso limitado ao conteúdo.

O resultado obtido da busca nas bases de dados, relacionados a quantidade e localização dos periódicos são apresentados na Tabela 3. É importante mencionar que nenhum trabalho foi encontrado nas bases de dados da Biblioteca Nacional<sup>18</sup> e Biblioteca Digital FGV<sup>19</sup>.

**Tabela 3** – Resultados da busca

| BASES DE DADOS | PROTOCOLOS DE BUSCA                      |  |   |
|----------------|--|--|---|
|                | <i>“statistical machine translation”</i> | <i>“statistical machine translation” AND “sign language”</i> | <i>“statistical machine translation” AND “sign language” AND “Libras”</i> |
| ARXIV          | 256                                      | 1  | 0   |
| CAPEL          | 2.761                                    | 58   | 2   |
| JIDM           | 10                                       | 0  | 0   |
| SCIELO         | 7  | 0  | 0   |
| SCOPUS         | 109*                                     | 30   | 0   |

Fonte: Dados da pesquisa, 2019.

\* *“Open Access”*, conteúdo gratuito com acesso limitado.

A escolha das *strings* de busca possibilitou a limitação da pesquisa a temas atuais, com resultados comprovados e possíveis de comparação. Foi possível observar que o número de artigos foi menor quando a *string* de busca *“statistical machine translation” AND “sign language” AND “Libras* foi utilizada. Diante disso, foi necessária uma busca manual em acervos particulares para, enfim, selecionar os estudos mais relevantes.

Na definição dos critérios de inclusão e exclusão, os estudos recuperados a partir da *string* de busca (*statistical machine translation AND sign language*), tiveram seus resumos analisados. Como critérios de inclusão foram selecionados artigos publicados entre os anos de 2009 a 2019, escritos em Português e Inglês. Como critérios de exclusão, foram excluídas teses e dissertações, e pesquisas que não estão relacionadas ao desenvolvimento de um *corpus* de tradução.

A partir das informações coletadas, optou-se pelas pesquisas mais semelhantes, com delineamento e critérios metodológicos condizentes à pesquisa atual.

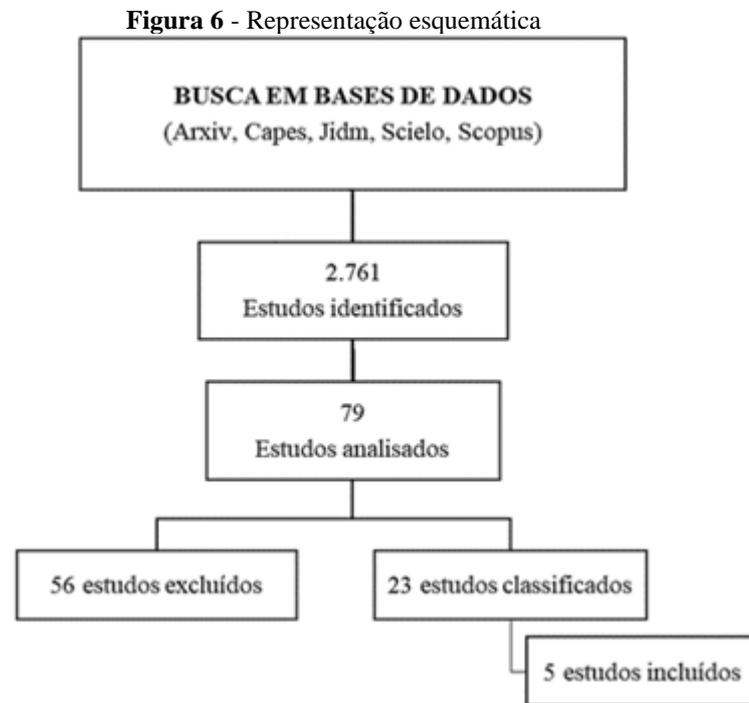
### 3.2 Busca, Seleção e Filtragem dos Trabalhos

Conforme mencionado na Seção 3.1, foram identificados 2.761 estudos envolvendo a *string* de busca *“statistical machine translation”*, no entanto, apenas 5 estudos passaram nos critérios de inclusão e exclusão e foram selecionados. Os estudos selecionados respondem às

<sup>18</sup> <https://www.bn.gov.br/>

<sup>19</sup> <https://sistema.bibliotecas-bdigital.fgv.br/>

principais questões da pesquisa: a importância do desenvolvimento de bases de dados empregando um *corpus* de tradução; e tratamento da tradução de uma língua oral para gestual referente à necessidade da construção de significados através da aplicação correta da estrutura gramatical. A Figura 6 apresenta uma representação esquemática do processo de seleção dos artigos.



Fonte: Dados da pesquisa, 2019.

Neste sentido, o primeiro desafio para chegarmos ao cerne do problema, que é a ausência de base de dados textuais em língua de sinais, é a necessidade de uma formalização de regras que sejam aplicadas à representação escrita da língua de sinais como um sistema de transcrição denominada GLOSA.

### 3.3 Discussão e análise

Os artigos selecionados a partir dos critérios de inclusão apresentam, em sua maioria, mecanismos de tradução automática de diversas línguas. Seja sobre pesquisas iniciais de desenvolvimento de tradução, ou, o aperfeiçoamento dos resultados por meio da criação de um *corpus* de tradução.

Othman e Jemni (2011) classificam o *corpus* como

[...] uma coleção cientificamente preparada de exemplos de como uma linguagem é comumente usada. Um corpus pode conter um grande número de textos escritos ou conversas gravadas ou filmadas. Essas coleções de dados são usadas para explorar o uso de uma linguagem ou para descobrir o vocabulário e a gramática dessa linguagem. Língua de Sinais é caracterizada por sua interatividade e multimodalidade, o que causa dificuldades na coleta e anotações de dados. (OTHMAN; JEMNI, 2011, p. 67, tradução nossa)<sup>20</sup>

Dentre os artigos analisados apenas um abordava a Libras relacionada a tradução automática. No entanto, a pesquisa se direcionava à implantação da língua de sinais em tvs, e não abordando tema do desenvolvimento de um *corpus* de tradução Português-Libras. (ARAÚJO, 2013)

Dentre os 23 artigos, classificados, 6 pesquisas abordam como tema principal a tradução automática baseada em regras aplicadas na elaboração de um dicionário/glossário de palavras na língua de sinais. Os outros 17 trabalhos abordam *corpus* de sentenças de tradução, utilizando a palavra-chave *Sign language* (SL). Diante dessa comparação é possível verificar que, apesar de muitos procurarem formas de desenvolver a tradução automática, não são todas as pesquisas que utilizam a elaboração de um *corpus* de tradução com o intuito de melhorar a qualidade do resultado da tradução. Isso pode ocorrer devido a fases iniciais de elaboração de tradutores automáticos, visto que algumas línguas ainda estão em fase de pesquisa e desenvolvimento de glossário.

Os artigos selecionados, classificados de acordo com a proposta deste trabalho, serão analisados considerando suas contribuições sobre o tema. Assim, os 5 trabalhos incluídos podem ser observados na Tabela 4.

**Tabela 4** – Características dos cinco estudos selecionados para revisão sistemática

| REFERÊNCIAS   | ANO DE PUBLICAÇÃO | TÍTULO  | PALAVRAS-CHAVE   |
|---|-------------------|---|--|
| CHINEA-RIOS, Mara;<br>SANCHIS-TRILLES,<br>Germán;<br>CASACUBERTA,<br>Francisco. | 2019              | Vector sentences representation for data selection in statistical machine translation | <i>Statistical machine translation. Datas election. Continuous vector -space representation. Cross-entropy. Infrequent ngrams recovery</i> |
| MISHRA, Gouri Sankar; SAHOO, Ashok Kumar;<br>RAVULAKOLLU, Kiran Kumar.          | 2017              | Word based statistical machine translation from English text to Indian sign language. | <i>Indian sign language. Statistical machine translation. Word alignment. Gloss. Corpus</i>  |

<sup>20</sup> “[...] a scientifically prepared collection of examples of how a language is commonly used. A corpus can contain a large number of written texts, or recorded or filmed conversations. Such data collections are used to explore the usage of a language or to find out about the vocabulary and grammar of this language. Sign Language is characterized by its interactivity and multimodality, which cause difficulties in data collection and annotation.”

|  |      |   |   |
|--|------|---|---|
| KOUREMENOS,<br>Dimitrios et al.                        | 2018 | Statistical Machine Translation for Greek to Greek Sign Language Using Parallel Corpora Produced via Rule-Based Machine Translation.          | <i>Machine translation. Greek. Greek Sign Language. GSL. Deaf people communication. SMT. Moses. Phrase model.</i> |
| STEIN, Daniel;<br>SCHMIDT, Christoph;<br>NEY, Hermann. | 2012 | Analysis, preparation, and optimization of statistical sign language machine translation.   | <i>Sign languages. Scarce resources. Syntactic methods. Parallel input</i>  |
| SU, Hung-Yu; WU,<br>Chung-Hsien.                       | 2009 | Improving structural statistical machine translation for sign language with small corpus using thematic role templates as translation memory. | <i>Sign language. Small corpus. Structural statistical. Machine translation (SSMT). Thematic relation.</i>        |

Fonte: Dados da pesquisa, 2019.

China-Rios, Sanchis-Trilles e Casacuberta (2019), classificam os *corpus* bilíngues como instrumentos essenciais para a linguística computacional, principalmente para a tradução automática estatística. Tendo como objetivo a criação de “[...] modelos matemáticos que podem descrever o processo de tradução com precisão, [...] aproveitando os dados de treinamento bilíngue.” (p. 1, tradução nossa)<sup>21</sup> Classificam o método manual como a forma mais confiável de avaliação, sendo mais dispendioso. Apesar de haver um debate constante no estudo básico de avaliação bilíngue (*Bilingual Evaluation Understudy* - BLEU), o uso de métricas automáticas continua sendo o método mais popular. Utilizam o *Europarl corpus*<sup>22</sup> para a avaliação de seu método de seleção de dados (*Data selection* - DS), chamado Representação Contínua de Vetor-Espaço de Sentenças para Seleção de Dados (*Continuous Vector-Space Representation of Sentences for Data Selection* - CRSDs). Esse método “[...] de seleção de dados baseado em CVR<sup>23</sup> de sentenças ou documentos, que produzem representações semelhantes para sentenças semanticamente próximas [...]” (CHINEA-RIOS; SANCHIS-TRILLES; CASACUBERTA, 2019, p. 14, tradução nossa)<sup>24</sup>.

<sup>21</sup> “[...] mathematical models that can describe the translation process accurately, [...] mathematical models that leverage bilingual training data.”

<sup>22</sup> <https://www.statmt.org/europarl/>

<sup>23</sup> *Continuous vector-space representation.*

<sup>24</sup> “[...] data selection method based on CVR of sentences or documents, which yield similar representations for semantically close sentences.”

Mishra, Sahoo e Ravulakollu (2017), propõem uma técnica de tradução automática estatística baseada em palavras, no entanto na Libras não funcionaria devido às suas regras morfossintático-discursivas, o *corpus* foi criado a partir da coleta de frases utilizadas pelo *Indian Railway Inquiry* (Departamento do Governo da Índia, responsável pela rede ferroviária do país), e dos anúncios feitos pela *Ramakrishna Mission Vivekananda University*. A ferramenta MOSES (Sistema de tradução automática estatística) foi utilizada para o desenvolvimento do *corpus*, com o inglês como língua de entrada e saída do sistema em glosas na Língua de Sinais Indiana (ISL). Este estudo contribuiu com o “[...] modelo de tradução lexical no qual as palavras das frases de origem são traduzidas para as palavras do idioma de destino e a ordem dessas palavras é alterada para formar uma frase de destino válida.” (p. 487, tradução nossa)<sup>25</sup> No entanto, ressaltam que esse método funciona com frases curtas, por se tratar de uma tradução baseada em palavras, visto que nesta técnica “a dependência das regras e gramática de ambas as línguas são ignoradas.” (MISHRA; SAHOO; RAVULAKOLLU, 2017, p. 487, tradução nossa)<sup>26</sup>

A pesquisa de Kouremenos et al. (2018) enfoca na Língua de Sinais Grega (GSL), propondo “[...] um sistema RBMT<sup>27</sup> inovador, que rapidamente produz um *corpus* GSL com glosas de alta qualidade. [...] o foco está principalmente na sintaxe, então são usadas glosas em vez de notação fonológica.” (p. 3, tradução nossa)<sup>28</sup> Esse sistema de tradução, apesar de usar um sistema de tradução automática, é supervisionado por um tradutor profissional. Permitindo que, os “[...] os textos de saída sejam corrigidos e novas regras de transferência e dados de mapeamento de léxico [...] adicionados ao RBMT, de modo que quaisquer casos novos (fenômenos linguísticos) sejam cobertos.” (p. 6, tradução nossa)<sup>29</sup> Com sentenças paralelas, e avaliação humana realizada por um tradutor profissional comprovando a precisão do processo de tradução, o *corpus* (Grego/GSL) possui um total de 1.015 sentenças e 20.287 *tokens*<sup>30</sup>. Relacionado aos recursos lexicais bilíngues, os autores afirmam que eles “[...] têm um efeito positivo nas abordagens baseadas em dados.” (KOUREMENOS et al., 2018, p. 11, tradução nossa)<sup>31</sup>

<sup>25</sup> “[...] lexical translation model in which words of source sentences are translated to the words of target language and the order of these words are changed to make a valid target sentence.”

<sup>26</sup> “[...] dependency on the rules and grammar of both the languages are ignored.”

<sup>27</sup> Rule-Based Machine Translation.

<sup>28</sup> “[...] innovative RBMT system is proposed, which quickly produces high quality large glossed GSL corpus. [...] the focus is primarily on syntax, so glosses are used instead of phonological notation.”

<sup>29</sup> “[...] output texts are corrected and new transfer rules and lexicon mapping data [...] added to the RBMT, so that any newly appearing cases (linguistic phenomena) are covered.”

<sup>30</sup> Palavras.

<sup>31</sup> “[...] has a positive effect in data-based approaches.”

Stein, Schmidt e Ney (2012) apresentam um sistema de tradução baseado em frases, chamado *Phrase-based translation* (PBT), do alemão para a Língua de Sinais Alemã (*Die Deutsche Gebärdensprachen* - DGS). Realizam um estudo sobre a língua de sinais e sua estrutura gramatical, como a variação fonética de sinais com o mesmo significado, e apesar de haver a possibilidade de treinamento de um sistema de tradução automática que apresente essas variantes, a quantidade insuficiente de dados de treinamento pode ser agravada pela escolha de um método de notação que permita uma inconstância de sinais. Deste modo, é importante

[...] usar um sistema de notação no qual as variantes de sinais possam ser facilmente normalizadas. Ao traduzir da língua de sinais, a normalização é feita como uma etapa de pré-processamento e, portanto, não é visível para o usuário. Ao traduzir para a língua de sinais, o computador escolheria uma das variantes de sinais em detrimento das outras. Pode-se, portanto, personalizar o avatar para produzir certos dialetos regionais ou preferências individuais em relação às variantes de sinais. (STEIN; SCHMIDT; NEY, 2012, p. 330, tradução nossa)<sup>32</sup>

No entanto, há algumas desvantagens e limites neste sistema de notação. Como a utilização de *classifier predicates* (predicados classificadores), empregues na descrição de ocorrências espaciais mais complexas (como o movimento, localização e tamanho) da sinalização. Na Libras, os classificadores (CLs), “[...] são configurações de mãos que, relacionadas à coisa, pessoa, animal e veículo, funcionam como marcadores de concordância.” (FELIPE, 2007, p.172) De acordo com Stein, Schmidt e Ney (2012), a dificuldade encontra-se no mapeamento dos movimentos tridimensionais num determinado conjunto de glosas. Visto que, nas línguas de sinais, as informações podem ser transmitidas pelas mãos, como “[...] também pela expressão facial, pelo olhar, pelos movimentos do tronco ou dos ombros e pelos movimentos da cabeça, como balançar a cabeça.” (p. 331, tradução nossa)<sup>33</sup> Apresentam como sugestão o pré-processamento, com marcação de final de frase, categorização e heurísticas de extração. Para os autores, as línguas de orais são mais complexas no nível morfológico, portanto “[...] devem ser pré-processadas com ferramentas morfossintáticas adequadas para facilitar o procedimento de tradução.” (STEIN; SCHMIDT; NEY, 2012, p. 354, tradução nossa)<sup>34</sup>

---

<sup>32</sup> “[...] use a notation system in which sign variants can easily be normalized. When translating from the sign language, the normalization is done as a preprocessing step and is therefore not visible to the user. When translating to the sign language, the computer would choose one of the sign variants over the others. One could therefore customize an avatar to produce certain regional dialects or individual preferences with regard to sign variants.”

<sup>33</sup> “[...] also by facial expression, eye-gaze, torso or shoulder movements and by head movements such as head shakes.”

<sup>34</sup> “[...] should be preprocessed with proper morpho-syntactic tools to facilitate the translation procedure.”

Su e Wu (2009) apresentam um modelo de tradução automática de estatística estrutural (*Structural Statistical Machine Translation - SSMT*), utilizando um pequeno *corpus* bilíngue para traduzir do chinês para a Língua de Sinais Taiwanesa (TSL).

Neste estudo, 1.651 sinais foram coletados do livro "Shou Neng Sheng Chyau" que é o material de ensino mais usado para TSL. Em "Shou Neng Sheng Chyau", Wayne et al. atribuído a cada sinal TSL um significado principal que é representado por uma palavra chinesa com o mesmo significado (posteriormente traduzido para uma palavra em inglês com letras maiúsculas) os sinais em TSL podem ser facilmente compreendido por Alunos TSL, e os significados sinônimos ou estendidos são considerados como os significados alternativos. (SU; WU, 2009 p. 1307, tradução nossa)<sup>35</sup>

Sua estratégia de tradução é baseada em regras, com o *corpus* bilíngue respeitando as variações estruturais da língua chinesa, contendo pares sentenciais limitados devido à necessidade da transcrição manual realizada por especialistas.

Concluindo que "a dispersão de dados é um grande problema na tradução automática, especialmente para línguas minoritárias, como SLs." (p. 1313, tradução nossa)<sup>36</sup>

Diante disso, a conversão de "[...] uma frase na estrutura sintática correspondente pode reduzir o problema de dispersão de dados utilizando regras gramaticais aplicadas na geração de idiomas." (p. 1313, tradução nossa)<sup>37</sup> No processo de desenvolvimento do *corpus* bilíngue foram observadas algumas limitações, devido às restrições gramaticais presentes na estrutura das frases. Estas restrições surgem porque, às vezes, as frases substantivas "[...] contêm vários rótulos com a mesma função temática e, portanto, resultam em erros de tradução no processo de correspondência." (SU; WU, 2009, p. 1313, tradução nossa)<sup>38</sup>

### 3.4 Considerações do Capítulo

Este estudo permitiu a análise dos artigos científicos, apresentados anteriormente, relacionados à tradução automática para línguas de sinais. Cada um deles com uma abordagem

---

<sup>35</sup> "[...] In this study, 1651 signs were collected from the text book "Shou Neng Sheng Chyau" [28] which is the most frequently used teaching material for TSL. In "Shou Neng Sheng Chyau," Wayne et al. assigned each TSL sign a principle meaning which is represented by a Chinese word with the same meaning (further translated into an English word with capital letters) making TSL signs can be easily understood by TSL learners, and the synonymous or extended meanings are regarded as the alternative meanings."

<sup>36</sup> "Data sparseness is a major problem in machine translation, especially for minority languages, such as SLs."

<sup>37</sup> "[...] sentence into the corresponding syntactic structure can reduce the data sparseness problem by using grammar rules which are used to generate languages."

<sup>38</sup> "[...] contain multiple labels of the same thematic roles and therefore result in translation errors in the matching process."

e finalidade específicas, sendo de grande importância para o delineamento e questionamento da pesquisa a ser desenvolvida.

É importante ressaltar que não foi encontrado nenhum trabalho sobre esse tema específico – o desenvolvimento de uma base de dados de tradução Português Brasileiro-Libras voltada para a aprendizagem de máquina. Assim, haverá a necessidade de inserir na lista de trabalhos acadêmicos importantes, materiais pessoais e de áreas diversas, como Educação e Linguística.

É possível observar nesse estudo, que o tema ainda não foi suficientemente explorado, visto que a maioria das pesquisas observadas abrange apenas os tradutores automáticos, não investigando o resultado e aprimoramento dessa tradução, e a possibilidade de desenvolvimento de uma base de dados de tradução para aprendizagem de máquina. Os tradutores automáticos Português/Libras a disposição atualmente na literatura (HandTalk<sup>39</sup>, ProDeaf<sup>40</sup> e Suite VLibras<sup>41</sup>) utilizam tradução baseada em regras e não direcionadas à criação de um *corpus* de tradução para treinamento de máquinas. Diante disso, a elaboração de regras de tradução que, diferente daquelas aplicadas às palavras que compõem um dicionário de Libras, considerem a gramática e o contexto da frase completa, trará mais qualidade à tradução.

No entanto, por ser uma produção morosa devido à extensão numérica da base de tradução, e desenvolvida manualmente por intérpretes de Libras, que possuem conhecimento nas duas línguas envolvidas, podem ocorrer erros simples durante a escrita da tradução, resultantes da participação de indivíduos. Sendo essa a principal limitação a ser considerada. Apesar disso, esse método é mais indicado para utilização de técnicas de aprendizagem de máquina.

---

<sup>39</sup> <https://www.handtalk.me/br>

<sup>40</sup> <http://www.prodeaf.net/>

<sup>41</sup> <https://www.vlibras.gov.br/>

#### **4 REGRAS APLICADAS NA CRIAÇÃO DE UM *CORPUS* PORTUGUÊS BRASILEIRO-LIBRAS**

Conforme mencionado no Capítulo 1, um dos desafios relacionados à criação de um *corpus* Português Brasileiro-Libras é a necessidade de definir um conjunto de regras para escrita de sinais em Libras, uma vez que as estratégias de tradução automática estatística geralmente demandam um *corpus* escrito na língua fonte e na língua alvo. Para endereçar esse problema, um conjunto de regras foi modelado e criado a partir de pesquisas teóricas linguísticas das áreas relacionadas – Português e Libras, para representar a escrita em Libras, através de uma notação de glosa.

A pesquisa realizada teve como base a utilização de métodos do tipo exploratório e descritivo, com pesquisa bibliográfica, documental, para a seleção das informações necessárias para a produção e desenvolvimento do estudo.

A aprendizagem de máquina, uma subárea de pesquisa em Inteligência Artificial (IA), tem como objetivo de estudo os procedimentos computacionais que permitem aquisição de habilidades, conhecimento e métodos de organizar conteúdos efetivos (MITCHELL, 1997). Dentre as abordagens de aprendizado existentes a realizada por indução engloba a temática escolhida neste estudo. Visto que, o aprendizado indutivo com exemplos distribuídos em sentenças de texto-texto, traz o desenvolvimento de descrições com base em definições pré-estabelecidas como exemplos, como principal atribuição. Para a língua de sinais, particularmente, os exemplos foram elaborados seguindo um conjunto de regras linguísticas, além da necessidade de contextualizar as informações quando observadas ambiguidades.

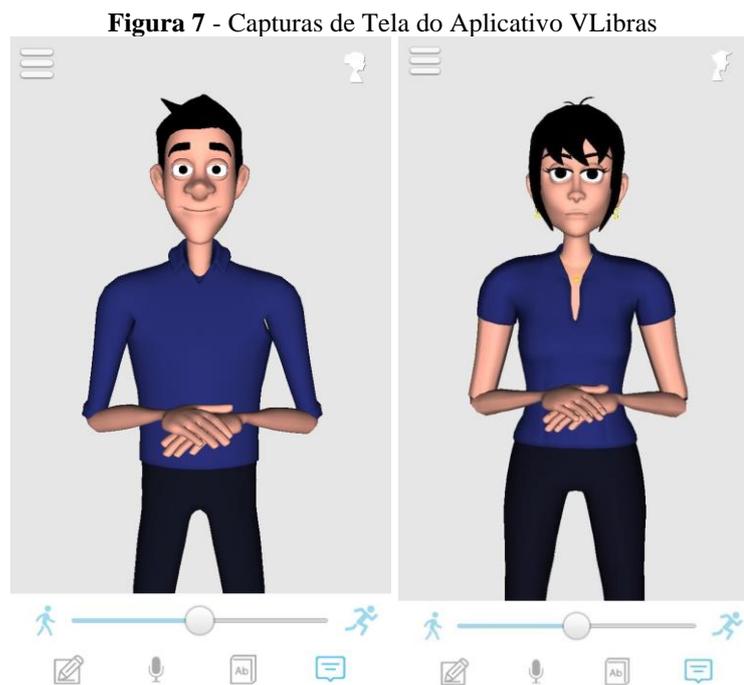
Aprendizado indutivo pode ser dividido em aprendizado supervisionado e não supervisionado. Neste trabalho, foi utilizado o aprendizado supervisionado, caracterizado pela construção de um modelo de exemplos (base de dados de tradução de língua de sinais) estabelecendo um conjunto de aglomerados padronizados, apresentando sentenças de entrada e saída, implementando uma sequência de resultados esperados. Para Batista (2003), o “aprendizado indutivo é um dos mais úteis pois permite obter novos conhecimentos a partir de exemplos, ou casos, particulares previamente observados.” (BATISTA, 2003, p. 12). Para Lima (2015), a linguagem de descrição tem como um de seus objetivos “[...] agregar os conhecimentos linguísticos da área para que estes produzam novas regras formais de tradução e adequação morfossintático-semântico.” (p. 73) Como também dos aspectos fonológicos da língua de sinais.

Desse modo, as regras desenvolvidas para tradução de conteúdos devem estar de acordo com as particularidades das línguas apresentadas, tornando os resultados obtidos melhores e mais compreensíveis ao usuário. Nessa perspectiva, é importante o constante aperfeiçoamento de softwares de tradução automática.

Para a validação da pesquisa, foi utilizado o tradutor automático VLibras, desenvolvido pelo Laboratório de Aplicações de Vídeo Digital (LAVID) da Universidade Federal da Paraíba (UFPB) em parceria com o Ministério do Planejamento, Orçamento e Gestão (MPOG). O projeto é

[...] um software que contém um conjunto de ferramentas utilizadas para a tradução automática do Português para a Língua Brasileira de Sinais. A utilização dessas ferramentas está disponível em várias plataformas, Desktop, Android, IOS, e em distintos navegadores de internet. (Freitas; Medeiros; Sousa, 2017, p. 573)

Na Figura 7, são apresentadas duas capturas de tela da aplicação VLibras, com seu avatar<sup>42</sup> nas versões masculina e feminina, assim como as opções de tradução (escrita, voz, dicionário).



Fonte: Captura de tela do App VLibras, 2020.

A proposta é que o *corpus* de tradução desenvolvido neste trabalho, formado por sentenças retiradas da web e traduzidas a partir de regras definidas de acordo com a estrutura

---

<sup>42</sup> Animação em 3D.

gramatical da língua de sinais, seja utilizado para treinamento de máquina e o aprimoramento do resultado da tradução automática do VLibras.

#### 4.1 Pontuação: renomeação da pontuação para tags pré-definidas

Sabe-se que, na Libras, a pontuação é um parâmetro paralelo, incorporado em outros sinais durante a sinalização através de expressões não-manuais. Na notação utilizada, os caracteres identificadores de pontuação (“.”, “!” e “?”) serão substituídos pela grafia do seu nome por extenso entre colchetes (ver Tabela 5). Sabendo que esse processo de tradução é realizado a partir de uma língua oral-auditiva na modalidade escrita para uma língua de sinais, utilizando um avatar, correspondente a sua modalidade gesto-visual.

Neste sistema de transcrição, a utilização de colchetes que identificam a pontuação não afetará a tradução da sentença, servindo apenas para marcar o final da sentença. Essa regra é necessária para que o tradutor identifique a marcação e finalização do contexto, visto que “sinais de pontuação, como ponto final, ponto e vírgula, parênteses, colchetes, pontos de interrogação e exclamação, servem para sinalizar ao sistema o início e o fim de um segmento”. (STUPIELLO, 2014, p.82) Além disso, é importante ressaltar que não há sinalização das vírgulas, que nas línguas de sinais pode ser realizada uma pausa, um intervalo entre as frases ou elementos do enunciado. Sendo esta, a representação de um tipo de frase expressa pela entonação/expressão corpo-facial, como também nas línguas de modalidade oral-auditiva. Exceto daquelas empregadas em números como separadores da parte inteira e decimal e valores monetários. Assim, neste caso, o sinal “,” muda para sua palavra por extenso “VÍRGULA”.

Na Libras, a Vírgula pode ser considerada um sinal icônico pois “[...] reproduz as características mais salientes de uma ação ou de um objeto.” (PIZZIO et al, 2008, .18) Neste sentido, esta regra foi construída para que o tradutor automático identifique esta informação e permita o tratamento.

**Tabela 5 - Pontuação**

| <b>SENTENÇA</b> | <b>REGRA APLICADA</b> |
|-----------------|-----------------------|
| Ponto final: .  | [PONTO]               |
| Exclamação: !   | [EXCLAMAÇÃO]          |
| Interrogação: ? | [INTERROGAÇÃO]        |
| Vírgula: ,      | VÍRGULA               |

Fonte: Dados da Pesquisa, 2019.

Na Tabela 6, são apresentadas algumas regras aplicadas em sentenças.

**Tabela 6** - Aplicação da regra de pontuação

| <b>SENTENÇA</b>        | <b>REGRA APLICADA</b>                     |
|------------------------|---|
| Gosto de comer.        | GOSTAR COMER [PONTO]                      |
| Ele é lindo!           | ELE LINDO [EXCLAMAÇÃO]                    |
| Pronto?                | PRONTO [INTERROGAÇÃO]                     |
| Consegui 7,9 na prova. | CONSEGUIR SETE VÍRGULA NOVE PROVA [PONTO] |

Fonte: Dados da Pesquisa, 2019.

## 4.2 Verbos

Quadros e Karnopp (2004) classificam os verbos na Libras como: simples, com concordância e espacial. Nesta seção, serão abordados os verbos simples, que não apresentam flexão verbal em número e pessoa, como também não apresentam afixos locativos, visto anteriormente na seção 2.3.3 sobre os aspectos sintáticos da língua de sinais.

Assim, na tradução do português brasileiro para a Libras, os verbos sem concordância devem estar em seu estado natural, no infinitivo, pois “não se alteram mesmo mudando-se as pessoas do discurso, mas eles admitem modificadores com intensificador ou advérbios de modo” (FELIPE, 2007, p. 398). O tempo pode ser marcado, sintaticamente, através de itens lexicais ou advérbios de tempo, como ontem, hoje e amanhã. Visto que na Libras não apresenta

[...] em suas formas verbais a marca de tempo como o português, [...] quando o verbo refere-se a um tempo passado, futuro ou presente, o que vai marcar o tempo da ação ou do evento serão itens lexicais ou sinais adverbiais como ONTEM, AMANHÃ, HOJE, SEMANA-PASSADA, SEMANA-QUE-VEM. (BRITO, 2002, p. 18)

No entanto, quando “não há, na frase, um advérbio de tempo específico, geralmente a frase, no presente, não é marcada, ou seja, não há nenhuma especificação temporal [...]” (FELIPE, 2007, p. 155). A Tabela 7 apresenta um exemplo de uso dessa regra.

**Tabela 7** - Verbos

| <b>SENTENÇA</b>  | <b>REGRA APLICADA</b>    |
|------------------|--------------------------|
| Vendo meu carro. | VENDER MEU CARRO [PONTO] |

Fonte: Dados da Pesquisa, 2019.

Observando a estrutura linguística presente nas sentenças em Libras, em alguns casos é possível perceber a ausência de preposição, de conjunções e de verbos de ligação (MIRANDA, 2014). Assim, os verbos de ligação (SER, ESTAR, FICAR) e preposição (PARA) devem ser removidos, pois na Libras seu uso não é frequente e no PB eles nem sempre carregam carga semântica, sendo apenas auxiliares na construção das sentenças (ver Tabela 8).

**Tabela 8** - Verbos de ligação e preposição

| SENTENÇA               | REGRA APLICADA           |
|------------------------|--------------------------|
| Meu caderno está sujo. | MEU CADERNO SUJO [PONTO] |
| Estou indo para casa.  | EU IR CASA [PONTO]       |

Fonte: Dados da Pesquisa, 2019.

### 4.3 Artigos

Os artigos definidos (a, o, as, os) e indefinidos (um, uma, uns, uma) devem ser excluídos. É importante ressaltar que a exclusão deve ser apenas do artigo indefinido e não dos numerais, ou seja, quando expressam quantidade. Estes elementos são, na maioria das vezes, elementos de ligação ou itens lexicais, não sendo sinalizados na Libras. Na Tabela 9, são apresentados alguns exemplos de uso dessa regra.

**Tabela 9** - Conversão dos artigos

| SENTENÇA                   | REGRA APLICADA                 |
|----------------------------|--------------------------------|
| O cachorro latiu.          | CACHORRO LATIR [PONTO]         |
| Procurei um bom argumento. | PROCURAR ARGUMENTO BOM [PONTO] |

Fonte: Dados da Pesquisa, 2019.

### 4.4 Troca de palavras

Na Libras, nos processos de formação de palavras

[...] deve-se destacar os inputs, que são as diferenças básicas entre as **regras de modificação de raiz** - alterações sistemáticas de uma base através da adição ou supressão de afixos ou modificações internas, e **as regras de composição** - conjunto de duas ou mais bases, que se combinam em uma outra forma, a partir de outro elemento ou modificações concomitantes. (FELIPE, 2006, p.201)

Neste sentido, a troca de palavras deve ser considerada em casos onde uma palavra ou expressão no PB seja formado pelo conjunto de 2 ou mais sinais em Libras (ver Tabela 10).

**Tabela 10** - Troca e fragmentação de palavras

| SENTENÇA                     | REGRA APLICADA                              |
|------------------------------|---|
| Vamos sair no fim de semana? | VAMOS PASSEAR SÁBADO DOMINGO [INTERROGAÇÃO] |
| Eu fui à igreja ontem.       | ONTEM IR IGREJA [PONTO]                     |
| Eu esquiu com frequência.    | EU ESQUIAR FREQUÊNCIA [PONTO]               |

Fonte: Dados da Pesquisa, 2019.

As palavras no plural são modificadas para o singular após a tradução. Assim, deve ser inserida uma palavra que caracterize quantidade quando houver a necessidade de manter essa característica (ver Tabela 11).

**Tabela 11** - Fragmentação de palavras no plural

| <b>SENTENÇA</b>          | <b>REGRA APLICADA</b>              |
|--------------------------|------------------------------------|
| Ele tem 2 casas.         | ELE TER DOIS CASA [PONTO]          |
| Ele tem casas na cidade. | ELE TER VÁRIOS CASA CIDADE [PONTO] |

Fonte: Dados da Pesquisa, 2019.

Nesta regra, em específico, não serão utilizados sinais de repetição para a palavra CASA quando esta estiver no plural. Pois, o tradutor automático retira a pluralidade das palavras, havendo a necessidade de inserir um sinal para expressar quantidade. No entanto, em pesquisas futuras pode-se apresentar como uma opção mais viável de tradução.

#### 4.5 Adjetivos: Unificação do gênero dos adjetivos

Os adjetivos são “[...] sinais que formam uma classe específica na LIBRAS e sempre estão na forma neutra, não havendo, portanto, nem marca para gênero (masculino e feminino), nem para número (singular e plural)” (BRITO, 2013, p.63). Diversos autores que trabalham com a língua de sinais adotam o padrão de utilização do “@” quando não há desinência de gênero na Libras. Este símbolo foi utilizado, pela primeira vez, por Felipe (1996, 1998) e, desde então os pesquisadores aderiram a essa marca para representar o gênero neutro na Libras. Considerando que, na Libras, não há marca de gênero nos adjetivos e como forma de simplificar o processo computacional nas regras propostas, optamos por manter a flexão neutra, mantendo no masculino. Essa decisão foi tomada uma vez que a sinalização dos adjetivos não sofre mudança com a flexão de gênero, bem como, por facilitar o processo de aprendizagem de máquina da tradução automática.

Na Tabela 12, são apresentados alguns exemplos de uso das regras envolvendo adjetivos.

**Tabela 12** - Adjetivos

| <b>SENTENÇA</b>       | <b>REGRA APLICADA</b>           |
|-----------------------|---------------------------------|
| Essa flor é cheirosa! | ESSE FLOR CHEIROSO [EXCLAMAÇÃO] |

Fonte: Dados da Pesquisa, 2019.

#### 4.6 Números ordinais

Na Libras há “[...] formas diferentes para apresentar os numerais quando utilizados como cardinais, ordinais, quantidade, medida, idade, dias da semana ou mês, horas e valores

monetários.” (FELIPE, 2007, p. 43) Em relação aos números cardinais, adota-se como regra que a notação sempre será por extenso, não devendo referenciar o algarismo (ver Tabela 13). Vale ressaltar que a sinalização da sentença apresentada na Tabela 13 será feita normalmente com o sinal referente ao número 4 e não sua datilologia.

**Tabela 13** - Números cardinais

| <b>SENTENÇA</b>             | <b>REGRA APLICADA</b>                   |
|-----------------------------|---|
| Preciso comprar 4 cadernos. | PRECISAR COMPRAR QUATRO CADERNO [PONTO] |

Fonte: Dados da Pesquisa, 2020.

No que se refere aos números ordinais, há a necessidade de expressá-los em sua notação por extenso. Acresce-se ainda a necessidade de inserir um identificador de desambiguação<sup>43</sup> aos numerais ordinais (do primeiro ao quinto), utilizando como marcador de contexto o caractere especial “&”, pois são considerados palavras homônimas na língua portuguesa como pode-se observar na Tabela de Palavras Homônimas (ver Seção 4.11). Na Tabela 14, pode ser observada a regra aplicada para números ordinais<sup>44</sup>.

**Tabela 14** - Números ordinais

| <b>SÍMBOLO</b> | <b>LEITURA</b>          |
|----------------|-------------------------|
| 1º             | PRIMEIRO&ORDINAL        |
| 2º             | SEGUNDO&ORDINAL         |
| 3º             | TERCEIRO&ORDINAL        |
| 4º             | QUARTO&ORDINAL          |
| 5º             | QUINTO&ORDINAL          |
| 6º             | SEXTO                   |
| 7º             | SÉTIMO                  |
| 8º             | OITAVO                  |
| 9º             | NONO                    |
| 10º            | DÉCIMO                  |
| 11º            | DÉCIMO PRIMEIRO&ORDINAL |
| 12º            | DÉCIMO SEGUNDO&ORDINAL  |
| 20º            | VIGÉSIMO                |

Fonte: Dados da Pesquisa, 2020.

Nos exemplos apresentadas na Tabela 15, é possível visualizar as diferenças aplicadas no *corpus* da base de dados.

<sup>43</sup> “Desambiguação lexical é um dos mais importantes e mais desafiadores pontos da tradução, considerando que os sinais devem ser sinalizados não de forma independente, mas de acordo com o contexto que atribui um valor semântico que consequentemente deve ser utilizado para a apresentação do sinal correto.”(COLLING; BOSCARIOLI, 2014, p. 9)

<sup>44</sup> Até esta etapa da criação do *corpus* foram utilizadas essas regras, no entanto, foi observado que a partir do número “10º” não há diferença entre números ordinais e cardinais. Dessa forma, a regra poderá sofrer modificação numa futura correção da base de dados e novo treinamento.

**Tabela 15** - Conversão dos números ordinais

| <b>SENTENÇA</b>                  | <b>REGRA APLICADA</b>   |
|----------------------------------|---|
| Fiquei na posição 134º da lista. | EU POSIÇÃO PRIMEIRO&ORDINAL<br>TERCEIRO&ORDINAL QUARTO&ORDINAL LISTA<br>[PONTO] |
| Sou o segundo na fila.           | EU SEGUNDO&ORDINAL FILA [PONTO]   |

Fonte: Dados da Pesquisa, 2020.

Quando um número se refere a uma quantidade, devem ser utilizados sinais específicos na Libras. Referente a hora, na Libras, há duas maneiras diferentes de se referir à hora: “horário cronológico e outro para a duração.” (FELIPE, 2007, p. 78), mais conhecido como incorporação de numerais (DEDINO, 2012), onde há a substituição de uma configuração de mão por outra utilizando quantificador ou números cardinais e ordinais em Libras. (XAVIER; NEVES, 2016) Assim, quando se refere à duração, devem ser usados sinais específicos, com a sinalização em círculo no rosto, denominada hora-duração (1-HORA, 2-HORA). Dessa forma, é possível desambiguar o numeral, diferenciando um numeral quantitativo de outro representativo (ver Tabela 16).

**Tabela 16** - Número e quantidade

| <b>SENTENÇA</b> | <b>REGRA APLICADA</b> |
|-----------------|-----------------------|
| 1 HORA.         | UM_HORA [PONTO]       |
| 1 PESSOA.       | UM_PESSOA [PONTO]     |

Fonte: Dados da Pesquisa, 2020.

É comum na Libras uma representação numérica<sup>45</sup> distinta para cada circunstância. Entender “[...] o significado das fórmulas, compreender as abstrações, constroem o pensamento e dá terreno para o entendimento.” (SILVA, 2018, p.11) Assim, cada divisão dos números em categorias apresenta uma fórmula de tradução. Neste trabalho, será feita a leitura da representação numérica ( $\frac{2}{3}$  = 2 “FRAÇÃO” 3 identifica a fração; % = “PORCENTAGEM” identifica a porcentagem; X = “VERSUS” identificar a comparação quando for referente a jogos ou disputas). Na Tabela 17, são apresentados alguns exemplos de uso envolvendo frações e percentuais.

**Tabela 17** - Representação numérica

| <b>SENTENÇA</b>   | <b>REGRA APLICADA</b>  |
|---|--|
| Temos $\frac{2}{3}$ de aprovação.   | TER DOIS FRAÇÃO TRÊS APROVAR [PONTO]                                   |
| A música teve 25,5% de rejeição.  | MÚSICA TER VINTE E CINCO VÍRGULA CINCO<br>PORCENTAGEM REJEIÇÃO [PONTO] |
| Fluminense ganhou de 4x2 do Vasco. ou<br>Fluminense ganhou de 4 a 2 do Vasco. | FLUMINENSE GANHAR QUATRO VERSUS DOIS VASCO<br>[PONTO]                  |

Fonte: Dados da Pesquisa, 2020.

<sup>45</sup> Até esta etapa de desenvolvimento do corpus não foi definido uma regra para diferenciar número cardinal de quantidade, podendo ser uma regra futuramente elaborada e acrescida.

#### 4.7 Período de tempo (hora)

No conjunto de regras proposto, as marcações de horas devem ser apresentadas respeitando a estrutura gramatical da Libras. Dessa forma, as abreviações comuns presentes no Português devem ser tratadas e descritas usando as glosas de HORA, MINUTO, MANHÃ, TARDE, etc. Na Tabela 18, são apresentados dois exemplos de tratamento de hora em Libras.

**Tabela 18** - Período de tempo (hora)

| SENTENÇA                 | REGRA APLICADA  |
|--------------------------|---|
| São 3h30.                | AGORA TRÊS HORA TRINTA MINUTO MANHÃ [PONTO]                 |
| A reunião será às 14h30. | REUNIÃO DOIS HORA TRINTA MINUTO TARDE [PONTO] <sup>46</sup> |

Fonte: Dados da Pesquisa, 2020.

#### 4.8 Palavras compostas

As palavras compostas ou expressões representadas por um único sinal devem estar separadas por \_ (underline). Para algumas palavras é “[...] importante mencionar que as palavras em português são compostas por mais de um elemento, entretanto, quando entram na Libras, são consideradas como sinais simples.” (NASCIMENTO, 2011, p. 52). Esta regra é aplicada às saudações<sup>47</sup> para que haja uniformidade da sinalização, visto que o tradutor automático utiliza esta transcrição. Como regra, esta notação também será aplicada às saudações, pois, na Libras, elas são caracterizadas como palavras simples (ver Tabela 19).

**Tabela 19** - Palavras compostas

| SENTENÇA                      | REGRA APLICADA                    |
|-------------------------------|-----------------------------------|
| Comprei um novo guarda-roupa. | COMPRAR GUARDA_ROUPA NOVO [PONTO] |
| Bom dia!                      | BOM_DIA [EXCLAMAÇÃO]              |
| Boa tarde!                    | BOA_TARDE [EXCLAMAÇÃO]            |
| Boa noite!                    | BOA_NOITE [EXCLAMAÇÃO]            |

Fonte: Dados da Pesquisa, 2020.

No entanto, ressalta-se que os nomes próprios de pessoas não serão modificados, exceto aqueles referentes a pessoas famosas (verificar regra de Nomes pessoais).

<sup>46</sup> Até esta etapa do desenvolvimento do *corpus* foi utilizado esta regra, no entanto, para as próximas correções esse exemplo de tradução pode sofrer alterações quanto à marcação de hora (Exemplo: AGORA 3 MEI@).

<sup>47</sup> Mesmo que seja dois sinais e não tenha concordância de gênero.

#### 4.9 Verbos com concordância número-pessoal

Nos verbos direcionais, a concordância entre emissor e receptor deve ser respeitada, além da expressão facial, que é considerado o quinto parâmetro de um sinal em Libras. Neste sentido, a notação dos verbos em sua forma neutra, infinitivo, recebe os caracteres \*S (“\*” identifica o número e o “S” identifica o singular) e \*P (“\*” identifica o número e o “P” identifica o plural), como nomenclaturas de direção. Sendo \*S para pronomes pessoais do caso reto ou representações do mesmo no singular (1S = EU, 2S = TU/VOCÊ, 3S= ELE/ELA), e \*P para pronomes pessoais do caso reto, ou representações do mesmo, no plural (1P = NÓS, 2P = VÓS, 3P = ELES/ELAS). Nesta regra, não serão abordadas as opções de incorporação de numeral (NÓS-2, 3, 4, GRUPO. VOCÊS- 2, 3 4, GRUPO), podendo ser um tema a ser discutido em pesquisas futuras na área.

Nas Tabelas 20 e 21, são apresentados exemplos de tratamento de verbos direcionais no singular e plural, respectivamente.

**Tabela 20** - Verbos com concordância

| SENTENÇA                       | REGRA APLICADA                     |
|--------------------------------|------------------------------------|
| Eu perguntei a sua idade.      | 1S_PERGUNTAR_2S IDADE [PONTO]      |
| O que você me perguntou?       | 2S_PERGUNTAR_1S QUE [INTERROGAÇÃO] |
| Maria perguntou meu nome.      | MARIA 3S_PERGUNTAR_1S NOME [PONTO] |
| Ele perguntou a ela sua idade. | 3S_PERGUNTAR_3S IDADE [PONTO]      |

Fonte: Dados da Pesquisa, 2020.

**Tabela 21** - Verbos com concordância relacionados aos pronomes pessoais no plural

| SENTENÇA                      | REGRA APLICADA                |
|-------------------------------|-------------------------------|
| Nós perguntamos a sua idade.  | 1P_PERGUNTAR_2S IDADE [PONTO] |
| Vós perguntam minha idade.    | 2P_PERGUNTAR_1S IDADE [PONTO] |
| Eles perguntaram a sua idade. | 3P_PERGUNTAR_2S IDADE [PONTO] |

Fonte: Dados da Pesquisa, 2020.

#### 4.10 Frases negativas

A incorporação da negação se aplica inicialmente aos verbos<sup>48</sup>, quando há modificação do advérbio de negação no objeto direto ou indireto. Nesse caso, é possível simplificar dizendo que, se a sinalização da NEGAÇÃO pode ser feita ao mesmo tempo que uma palavra (verbo) essa regra específica deve ser aplicada. Visto que em frases negativas o “[...] movimento da cabeça (negando), as expressões faciais, são obrigatórias para marcar sentenças negativas, pois

<sup>48</sup> Esta regra poderá ser aplicada em palavras variadas em versões futuras do *corpus*, desde que os sinais sejam incorporados ao dicionário do tradutor automático utilizado.

está diretamente ligada às questões sintáticas, caso contrário, a sentença se tornará agramatical.” (ALMEIDA; ALMEIDA, 2012, p. 628). Na Tabela 22, são apresentados dois exemplos de aplicação desta regra.

**Tabela 22** - Verbos que incorporam negação

| SENTENÇA                   | REGRA APLICADA                   |
|----------------------------|----------------------------------|
| Eu não posso mais esperar. | NÃO_PODER ESPERAR [PONTO]        |
| Ele não consegue ganhar.   | ELE NÃO_CONSEGUIR GANHAR [PONTO] |

Fonte: Dados da Pesquisa, 2020.

De acordo com a Tabela 22, é possível observar que se optou pela utilização da modificação de negação como um prefixo para facilitar a leitura da descrição.

#### 4.11 Homônimos e palavras homógrafas

No conjunto de regras proposto, as palavras que possuem uma grafia idêntica, denominadas homônimas, receberão uma nomenclatura descritiva no final das palavras. Essa descrição estará separada das palavras pelo caractere especial “&”. Esse caractere especial foi escolhido por questões de melhor adaptação aos aspectos computacionais do modelo de aprendizagem de máquina.

Esta regra possibilita a tradução correta considerando os aspectos semânticos e pragmáticos da Libras pois permite a desambiguação de homônimos e palavras homógrafas, pois no processo de tradução alguns sinais podem ser mais específicos, dessa forma, a tradução deve considerar o significado do item lexical da língua portuguesa para um item lexical correspondente na língua de sinais. Nesta regra, serão tratados os verbos em Libras que possuem concordância com objetos e classificadores (ver Seção 2.3.4 e Tabela 24).

No desenvolvimento desta pesquisa, foram identificadas 651 palavras homônimas (totalizando 2.064 definições, entre elas os verbos acrescidos de descrição, onde o contexto pode alterar a forma como manuseamos um objeto ou realizamos uma ação na língua de sinais)<sup>49</sup>.

Na Tabela 23, é ilustrada a palavra COLAR e algumas variações em Libras usando a regra proposta.

<sup>49</sup> As palavras homônimas podem ser acessadas através do endereço: <https://docs.google.com/spreadsheets/d/14-J-MaxPK-7Cji43AETbp29MP63OVQWJ71SVnEKJKBk/edit?usp=sharing>. Onde a tabela, contendo todas as descrições referentes às variações de significado e contextualização, estará disponível para consulta. Sendo estas variações, além daquelas relacionadas à desambiguação, aos sinais locativos e incorporação de formato. (XAVIER; NEVES, 2016)

**Tabela 23 - Homônimos (palavra Colar)**

| HOMÔNIMO | VARIÇÕES                      |                           |                            |                                 |
|----------|-------------------------------|---------------------------|----------------------------|---------------------------------|
| COLAR    | COLAR&ACESSÓRIO <sup>50</sup> | COLAR&FILAR <sup>51</sup> | COLAR&GRUDAR <sup>52</sup> | COLAR&INFORMÁTICA <sup>53</sup> |

Fonte: Dados da Pesquisa, 2020.

De acordo com a Tabela 23, é possível notar que uma palavra considerada simples pode apresentar três significados diferentes na língua de sinais, sendo a contextualização um dos aspectos de distinção estabelecido. Na Tabela 24, são apresentados alguns exemplos de sentenças usando a palavra COLAR com diferentes significados.

**Tabela 24 - Aplicação de homônimos em frases (palavra colar)**

| SENTENÇA                     | REGRA APLICADA                         |
|------------------------------|--|
| Ganhei um colar da minha vó. | GANHAR COLAR&ACESSÓRIO MEU VÓ [PONTO]  |
| É errado colar na prova.     | ERRADO COLAR&FILAR PROVA [PONTO]       |
| O papel está todo colado.    | PAPEL COLAR&GRUDAR [PONTO]             |
| Copie e cole o texto.        | COPIAR COLAR&INFORMÁTICA TEXTO [PONTO] |

Fonte: Dados da Pesquisa, 2020.

Nas Tabela 25 e 27, são apresentadas as palavras ARTIGO e APAGAR, respectivamente, e suas respectivas variações em Libras usando a regra proposta, enquanto que, nas Tabela 26 e 28, são apresentados alguns exemplos de sentenças usando essas palavras com diferentes significados, apresentando a incorporação do objeto, locativo e contextualização.

**Tabela 25 - Homônimos (palavra Artigo)**

| HOMÔNIMO | VARIÇÕES                   |                             |                                |                          |
|----------|----------------------------|-----------------------------|--------------------------------|--------------------------|
| ARTIGO   | ARTIGO&TEXTO <sup>54</sup> | ARTIGO&OBJETO <sup>55</sup> | ARTIGO&GRAMÁTICA <sup>56</sup> | ARTIGO&LEI <sup>57</sup> |

Fonte: Dados da Pesquisa, 2020.

<sup>50</sup> Ornato usado ao redor do pescoço que consiste em uma fita, faixa, cadeia ou enfiada de contas, pérolas, sementes ou coisas semelhantes.

<sup>51</sup> Copiar clandestinamente de anotações, livros ou receber de outrem as respostas às questões constantes de um exame.

<sup>52</sup> Fazer aderir com cola; grudar.

<sup>53</sup> Inserir, em um arquivo, texto ou gráficos que foram copiados ou removidos de outro arquivo ou, ainda, de uma parte para outra do mesmo arquivo.

<sup>54</sup> Subcategoria gramatical de determinantes do nome que indicam a maneira como deve ser apreendido o ser ou a coisa designada pelo substantivo. Pode assumir o aspecto semântico de definido (o, a, os, as) ou indefinido (um, uma, uns, umas) e vir precedido de um pré-artigo (toda a sala) ou seguido de um pós-artigo (a própria sala). Em português, ocupa posição que antecede o nome e concorda com ele em gênero e número; artigo.

<sup>55</sup> Objeto, mercadoria, produto.

<sup>56</sup> Texto de uma publicação (jornal, revista, tese etc.).

<sup>57</sup> Parte que corresponde a uma divisão ou subdivisão de um código, lei, tratado, portaria etc. (em geral, é marcada por uma numeração); parágrafo de uma articulação.

**Tabela 26** - Aplicação de homônimos em frases (palavra Artigo)

| SENTENÇA                           | REGRA APLICADA                             |
|------------------------------------|--|
| Aprendi muito com esse artigo.     | ARTIGO&TEXTO APRENDER(+) [PONTO]           |
| Onde estão os artigos do banheiro? | ONDE ARTIGO&OBJETO BANHEIRO [INTERROGAÇÃO] |
| Não usamos artigo na Libras.       | NÃO_USAR ARTIGO&GRAMÁTICA LIBRAS [PONTO]   |
| Um artigo do código penal.         | ARTIGO&LEI CÓDIGO_PENAL [PONTO]            |

Fonte: Dados da Pesquisa, 2020.

**Tabela 27** - Homônimos (verbo Apagar)

| HOMÔNIMO | VARIACÕES                      |                            |                                  |
|----------|--------------------------------|----------------------------|----------------------------------|
| APAGAR   | APAGAR&EXTINGUIR <sup>58</sup> | APAGAR&FOGO <sup>59</sup>  | APAGAR&INFORMÁTICA <sup>60</sup> |
|          | APAGAR&ESCREVER <sup>61</sup>  | APAGAR&LOUSA <sup>62</sup> | APAGAR&LUZ <sup>63</sup>         |
|          | APAGAR&VELA <sup>64</sup>      |                            |                                  |

Fonte: Dados da Pesquisa, 2020.

**Tabela 28** - Homônimos (verbo Apagar)

| SENTENÇA                          | REGRA APLICADA                                       |
|-----------------------------------|--|
| Cuidado, apague o fogo!           | CUIDADO APAGAR&FOGO [EXCLAMAÇÃO]                     |
| Apague essa frase.                | APAGAR&INFORMÁTICA FRASE [PONTO]                     |
| Apague meu nome da lista.         | APAGAR&ESCREVER NOME LISTA [PONTO]                   |
| Posso apagar o quadro?            | PODER APAGAR&LOUSA [INTERROGAÇÃO]                    |
| Entre e apague a luz.             | ENTRAR APAGAR&LUZ [PONTO]                            |
| Apague as velas e faça um pedido. | APAGAR&VELA FAZER PEDIDO&PEDIR <sup>65</sup> [PONTO] |

Fonte: Dados da Pesquisa, 2020.

#### 4.12 Classificação de lugares - Cidade, Estado e País

Para representação de lugares (cidades, estados e países), a regra proposta também prevê a inclusão do caractere '&' para marcação de contexto (ver Tabela 29). Mais especificamente, nos nomes de cidades, será acrescida a marcação &CIDADE, enquanto que, nos nomes de estados e países, serão acrescidas as marcações &ESTADO e &PAÍS, respectivamente.

**Tabela 29** - Regra de representação de lugares

| SENTENÇA                  | REGRA APLICADA                      |
|---------------------------|-------------------------------------|
| Eu moro em São Paulo.     | MORAR SÃO_PAULO&ESTADO [PONTO]      |
| Viajo para Salvador hoje. | VIAJAR SALVADOR&CIDADE HOJE [PONTO] |

Fonte: Dados da Pesquisa, 2020.

<sup>58</sup> Extinguir(-se), acabar(-se).

<sup>59</sup> Fazer desaparecer (fogo).

<sup>60</sup> Sinal próprio da Libras, deletar algo digitado.

<sup>61</sup> Sinal próprio da Libras, apagar com borracha.

<sup>62</sup> Sinal próprio da Libras, apagar o quadro ou lousa.

<sup>63</sup> Sinal próprio da Libras, apagar a luz.

<sup>64</sup> Sinal próprio da Libras, apagar a vela, assoprar.

<sup>65</sup> Ato de pedir; solicitação.

Quando uma palavra que é homônima a um nome de um lugar está representando um nome próprio de pessoa (ou uma característica própria de uma pessoa) na sentença, este deve ser mantido em sua escrita normal sem a necessidade de desambiguação, pois, na Libras, eles são sinalizados em datilologia por não possuírem sinal próprio (verificar regra de Nomes pessoais). Na Tabela 30, pode-se observar um exemplo de desambiguação com nome próprio.

**Tabela 30** - Classificação de lugar: contextualização

| <b>SENTENÇA</b>                      | <b>REGRA APLICADA</b>  |
|--------------------------------------|--|
| A Argentina estuda de noite.         | ARGENTINA ESTUDAR NOITE [PONTO]                              |
| Viajei para a Argentina ano passado. | VIAJAR ARGENTINA&PAÍS ANO&DATA <sup>66</sup> PASSADO [PONTO] |

Fonte: Dados da Pesquisa, 2020.

Nas palavras que possuem o sentido de nacionalidade, deverá ser considerada a fragmentação na tradução. Neste caso, utilizaremos a “justaposição de dois itens lexicais, ou seja, dois sinais que formam uma terceira forma livre”. (FELIPE, 2006, p. 207) Como, por exemplo, nos itens lexicais NASCER+PAÍS/ESTADO/CIDADE (NASCER + BRASIL = brasileiro). Na Tabela 31, pode-se observar a regra de decomposição de palavras aplicada em exemplos.

**Tabela 31** - Decomposição de palavras: classificação de lugar

| <b>SENTENÇA</b>            | <b>REGRA APLICADA</b>                        |
|----------------------------|--|
| Ele é canadense.           | ELE NASCER CANADÁ&PAÍS [PONTO]               |
| Meu amigo é estadunidense. | MEU AMIGO NASCER ESTADOS&UNIDOS&PAÍS [PONTO] |

Fonte: Dados da Pesquisa, 2020.

A composição de sinais, quando referentes a classificação de nacionalidade, deverá seguir essa alteração de palavras na tradução, seguindo a ordem apresentada. Assim, o contexto será respeitado.

#### **4.13 Nomes pessoais**

Na Libras, a datilologia é utilizada na soletração de nomes próprios de pessoas, ou palavras que não possuem sinal definido na língua de sinais. (FERNANDES; ROMEIRO, 2016). Assim, buscando limitar o uso da datilologia pelo tradutor automático, foi desenvolvida

<sup>66</sup> Espaço de tempo da revolução completa de qualquer planeta em torno do Sol.

uma regra que analisasse nomes pessoais que possuem sinal específico, de notório reconhecimento e catalogados no Dicionário Capovilla, bem como, no LibrasGov<sup>67</sup>. Dessa forma, o tradutor faz a leitura da frase buscando em seu acervo o sinal correspondente aos nomes pessoais apresentados, com nome e sobrenome.

Para simplificar esse processo, os nomes pessoais referentes a pessoas famosas recebem o sufixo &FAMOSO. Na Tabela 32, são apresentadas sentenças onde a regra é aplicada.

**Tabela 32** - Nomes pessoais

| <b>SENTENÇA</b>                     | <b>REGRA APLICADA</b>                           |
|-------------------------------------|---|
| Helen Keller é uma escritora.       | HELEN_KELLER&FAMOSO ESCRITOR [PONTO]            |
| Luciano Huck trabalha na televisão. | LUCIANO_HUCK&FAMOSO TRABALHAR TELEVISÃO [PONTO] |

Fonte: Dados da Pesquisa, 2020.

#### 4.14 Amplificação e redução de intensidade

Os intensificadores são usados quando o verbo, adjetivo, substantivo ou expressão está próximo (podendo ser como sucessor ou antecessor) de algum advérbio de intensidade. Sendo essa intensidade caracterizada pela “[...] duração, energia, variância e velocidade média” (PASSOS, 2014, p.19) da sinalização. Neste trabalho, a regra proposta envolve a inserção dos marcadores “(+)” ou “(-)” no final da palavra, para representar o intensificador. Nesse caso, os advérbios de intensidade "muito", "mais", "muitíssimo" e seus sinônimos serão representados pelo marcador "(+)" ao final da palavra, enquanto que os advérbios "pouco", "menos", “pouquíssimo” e seus sinônimos serão representados pelo marcador “(-)” ao final da palavra.

Na Tabela 33, são apresentados alguns exemplos da aplicação da regra em frases com ampliação e redução de intensidade.

**Tabela 33** - Amplificação e redução de intensidade

| <b>SENTENÇA</b>           | <b>REGRA APLICADA</b>      |
|---------------------------|----------------------------|
| Ele está muito deprimido. | ELE DEPRIMIDO(+) [PONTO]   |
| Ela é tão bonita!         | ELE BONITO(+) [EXCLAMAÇÃO] |
| Você está nervoso demais. | VOCÊ NERVOSO(+) [PONTO]    |
| Estou um pouco doente.    | EU DOENTE(-) [PONTO]       |

Fonte: Dados da Pesquisa, 2020.

Para uma total aplicação das regras, foi necessária a inserção de sinais específicos para cada intensidade representada na Libras ao dicionário do tradutor automático utilizado, visto que as expressões faciais e corporais devem estar de acordo com a sinalização.

<sup>67</sup> Disponível em: <https://youtu.be/Pw6yWmTnWIQ> Acesso em: 03 ago. 2020.

#### 4.15 Expressões idiomáticas

Foi observado que em alguns casos, ao traduzir expressões idiomáticas na língua portuguesa para a língua de sinais, podemos manter referente a palavra ao seu significado. Freire (2012), afirma que

[...] uma expressão idiomática ou provérbio traduzidos literalmente seriam desadequados e até possivelmente incompreensíveis. Deve, então, o intérprete procurar um equivalente linguístico e cultural que transmitirá a mensagem de forma eficaz e adequada. No caso de não existir uma expressão equivalente, que veicule a mesma ideia, deve-se recorrer a uma adaptação, usando uma paráfrase que transmita a ideia central do provérbio ou expressão idiomática. (FREIRE, 2012, p. 125)

Nesse caso, entram os xingamentos, que muitas vezes possuem a mesma grafia de substantivos simples e verbos. Assim, esta regra permite que o tradutor automático possa diferenciar o contexto ao qual a palavra está inserida. Pode-se usar como exemplo a expressão “chutar o pau da barraca”, que juntos possuem um significado<sup>68</sup> e separados possuem outro.

Lembrando que a sinalização destas expressões na língua de sinais está disponível no Dicionário Capovilla, escolhido por apresentar um conteúdo atualizado sistematicamente. Na Tabela 34, pode-se observar alguns desses casos.

**Tabela 34** - Expressões idiomáticas

| SENTENÇA | REGRA APLICADA                             |
|----------|--|
| Merda!   | MERDA_EXPRESSÃO [EXCLAMAÇÃO]               |
| Pare!    | PARAR_EXPRESSÃO [EXCLAMAÇÃO] <sup>69</sup> |

Fonte: Dados da Pesquisa, 2020.

Na Libras, como em outras línguas de modalidade oral-auditiva na fala presencial, pode ocorrer também a substituição de frases por apenas uma expressão face-corporal, o que facilitaria o entendimento do usuário de qualquer língua. Assim, toda frase que puder ser substituída por uma expressão (ou apenas uma palavra) deve ser repassada para inclusão na lista de expressões idiomáticas<sup>70</sup>.

<sup>68</sup> Os significados denotativos são os mesmos, mas a frase pode ter um significado conotativo.

<sup>69</sup> Não há a sinalização da [EXCLAMAÇÃO], sendo utilizada apenas como marcação de contexto e finalização da frase.

<sup>70</sup> Como uma atividade futura pode ser desenvolvido um inventário de expressões idiomáticas.

A Tabela 35 traz exemplos de sentenças que apresentam a substituição de uma frase por uma representação na língua de sinais. É importante mencionar, no entanto, que essas substituições não devem modificar o contexto da sentença inicialmente apresentada.

**Tabela 35** - Substituição de frases por uma expressão

| <b>SENTENÇA</b>      | <b>REGRA APLICADA</b> |
|----------------------|-----------------------|
| Não está claro.      | NÃO_CLARO [PONTO]     |
| Não vou tomar banho. | NÃO_BANHAR [PONTO]    |

Fonte: Dados da Pesquisa, 2020.

#### 4.16 Contexto

Outro fator importante na tradução é a semântica, que estuda o significado. Na Libras, com qualquer processo de tradução de uma língua para outra, essa relação torna-se mais complexa, pois está diretamente ligada ao sentido da palavra<sup>71</sup>. Nesse caso, pode-se fazer a substituição, adição ou remoção de palavras, desde que seja repassado o motivo para tal troca.

A utilização da função semiótica é essencial para descrever o contexto e adotar as palavras que fariam mais sentido na Libras, a substituição de termos e frases tem como princípio apresentar os parâmetros da língua de sinais (níveis fonológico e morfológico).

Neste caso, há os ditados populares que são, muitas vezes “[...] expressões tão usadas na cultura ouvinte, não fazem parte do dia a dia do Surdo sendo até algo desconhecido para muitos deles.” (SANTOS, 2018, p.12)

Diante disso, na Libras deve-se interpretar a frase pois seu significado pode ser totalmente diferente do mostrado na glosa em Libras apresentada pela tradução mecânica, visto que para uma tradução da frase é necessário conhecer os seus constituintes utilizados em contextos específicos. Podemos observar, na Tabela 36, o significado de um ditado popular transcrito em glosa.

**Tabela 36** - Ditados populares

| <b>SENTENÇA</b>                   | <b>REGRA APLICADA</b>                              |
|-----------------------------------|--|
| Um olho no peixe e outro no gato. | PRECISAR ESPERTO ABRIR&OLHOS <sup>72</sup> [PONTO] |

Fonte: Dados da Pesquisa, 2020.

<sup>71</sup> A questão do sentido está relacionada ao nível pragamático; portanto, é a partir do discurso em um determinado contexto que se pode apreender o sentido de um PALAVRA OU FRASE. Por outro lado, todas o léxico de uma língua tem significado. Signo linguístico = significante + significado (Saussure)

<sup>72</sup> Sinal próprio da Libras disponível na lista de palavras homônimas apresentada no tópico 4.11 Homônimos e palavras homógrafas deste capítulo.

Nessa mesma categoria serão inseridos os “sinônimos” (ver Tabela 37). Para Pessanha e Cozendey (2011), o “[...] uso de sinônimos apresenta-se como uma possibilidade de superação de obstáculos linguísticos no compartilhamento de sentido.” (p. 9) Assim, a substituição de uma palavra para outra mais adequada ao contexto da frase inicial deve ser utilizado, evitando que o tradutor faça a datilologia de palavras não incorporadas ao seu dicionário (ver Tabela 37).

**Tabela 37** - Substituição por sinônimos

| <b>PALAVRA</b> | <b>SINÔNIMO</b> |
|----------------|-----------------|
| IMPRESCINDÍVEL | IMPORTANTE      |

Fonte: Dados da Pesquisa, 2020.

No capítulo 5, serão apresentados os resultados e discussões obtidos no processo de desenvolvimento do *corpus* de tradução, modelado e desenvolvido usando as regras descritas neste capítulo.

## 5 RESULTADOS E DISCUSSÕES

Neste Capítulo, serão apresentados os resultados obtidos na modelagem e desenvolvimento de um *corpus* de tradução que utiliza regras especificadas no Capítulo 4, aplicadas no tradutor automático VLibras. Mais especificamente, será apresentado o processo de criação da base de dados, a quantidade de sentenças adaptadas, o tempo envolvido na modelagem e desenvolvimentos do *corpus* Português Brasileiro - Libras, e outras informações adicionais. Em seguida, serão apresentados os resultados relacionados à aplicação dos questionários e sua análise.

### 5.1 Processo de criação da base de dados

Para criação do corpus de tradução, as sentenças foram inicialmente retiradas da base de dados do Datafolha<sup>73</sup>. No entanto, foram realizados experimentos na base de artigos do Wikipedia acessível por meio do Wikidata<sup>74</sup>, para a construção de um corpus mais específico. Assim,

A partir dela foram extraídas aproximadamente 5,8 milhões de sentenças em Português, contendo cerca de 900 milhões de tokens, e geradas glosas equivalentes usando a Suíte VLibras (o tradutor da Suíte VLibras utiliza a estratégia de Tradução Baseada em Regras e possui código aberto). O corpus bilíngue Português Brasileiro/LIBRAS produzido com esse processo foi usado como um dos níveis do fator (5,8M). Para o outro nível, foram selecionadas aleatoriamente 200 mil sentenças do corpus maior para a definição de um corpus de treinamento menor (200K). (VERÍSSIMO et al. 2019, p. 4)

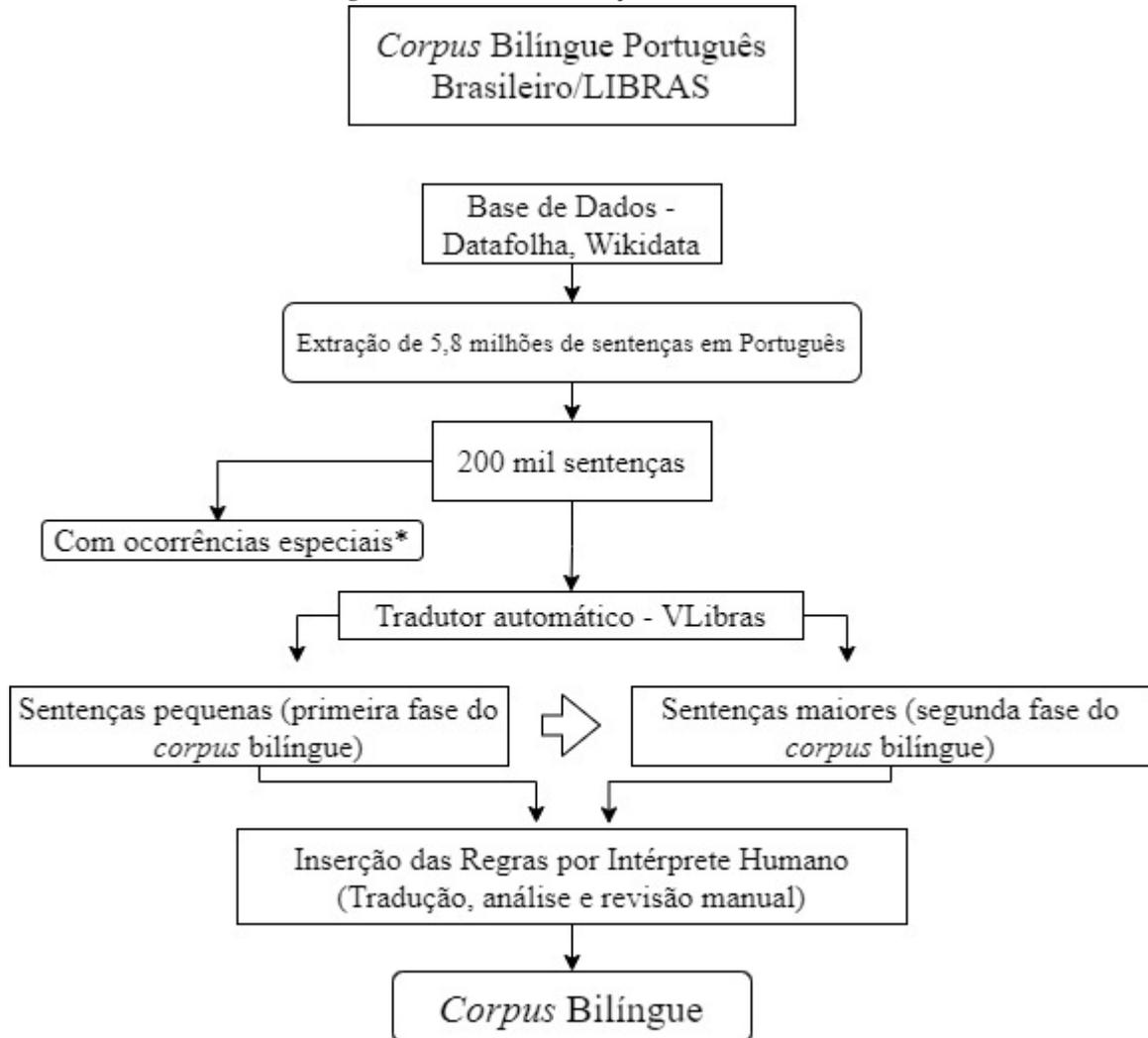
Partindo dessas 200 mil sentenças, foram selecionados os casos em que havia ocorrências especiais (tais como verbos com concordância, intensificador, homônimo e negação), dado que o tradutor automático baseado em regras tem dificuldades para tratar adequadamente este tipo de sentença. Dessa forma, o *corpus* foi gerado com a intenção de validar a aprendizagem de máquina para estes contextos específicos, e depois ampliar para o contexto geral, sendo necessário um *corpus* mais amplo. Inicialmente, foram trabalhadas sentenças pequenas (primeira fase do *corpus* bilíngue) e, em seguida, foram trabalhadas sentenças maiores (segunda fase do *corpus* bilíngue), permitindo alcançar uma maior dimensão de casos especiais devido a variedade de sentenças e assuntos abordados.

---

<sup>73</sup> <http://datafolha.folha.uol.com.br/>

<sup>74</sup> [https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)

Figura 8 - Processo de criação da base de dados



\*Verbos direcionais, intensificador, homônimo e negação

Fonte: Dados da Pesquisa, 2021.

O *corpus* de tradução proposto foi desenvolvido dentro do escopo do projeto VLibras<sup>75</sup> e contou com uma equipe de 8 intérpretes de Libras contendo a seguinte formação:

- **Formação técnica em Libras:** 6 integrantes possuem curso de formação em Tradutor/Intérprete de Libras pela Funad, 1 integrante é formado no Curso de Extensão em Libras pela UFPB e 1 integrante formado em Libras pelo Senac.

<sup>75</sup> O *corpus* bilíngue pode ser utilizado por qualquer tradutor automático Português-Libras, o VLibras foi escolhido para a validação dos resultados, servindo como estudo de caso.

- **Formação de nível superior:** 3 integrantes possuem graduação em Letras-Libras, 2 integrantes são graduandos em Letras-Libras e 3 integrantes são graduados em Pedagogia;
- **Especialização:** 2 integrantes possuem especialização em Libras e 1 integrante possui especialização em Gestão e Avaliação da Educação Superior.

A faixa etária da equipe de intérpretes variava entre 21 e 51 anos, sendo todas do gênero feminino. Com relação à experiência profissional, 5 integrantes possuem entre 7 a 13 anos de experiência como Intérprete de Libras, 3 integrantes possuem de 2 a 5 anos de experiência como professor de Libras, e 1 integrante possui 9 anos de experiência na educação de Surdos como pedagoga.

Essa equipe de intérpretes atuou traduzindo, analisando e revisando manualmente sentenças provenientes dessas fontes de dados durante 10 meses de trabalho utilizando as regras propostas no Capítulo 4. As intérpretes trabalhavam, no ambiente do laboratório LAVID, em média 20hs semanais, distribuídas entre os horários matutino e vespertino, havendo sempre a presença mínima de duas intérpretes, garantindo apoio profissional caso surgissem dúvidas sobre a tradução e revisão das sentenças. A equipe analisava 2.800 sentenças mensais na primeira fase do *corpus* bilíngue (com sentenças pequenas), e 600 sentenças mensais na segunda fase do *corpus* bilíngue (com sentenças grandes).

Após esse período foi adicionado um mês para a realização de uma revisão geral das sentenças, pois muitos homônimos foram anexados durante o desenvolvimento do *corpus*. Dessa forma, a correção permitiu que todas as palavras homônimas fossem corrigidas e recebessem a descrição correta.

Do mesmo modo, o processo de criação do *corpus* de tradução contou com a presença de dois surdos, com 21 e 37 anos. Os surdos possuem Licenciatura em Letras-Libras com Especialização em Libras - com 10 anos de experiência como professora de Libras, e graduação em andamento em Tecnologia em Sistemas para Internet, respectivamente. Eles atuaram revisando o processo de criação do *corpus*, o que possibilitou que as regras e a tradução final estivessem em consonância com os requisitos da comunidade surda. Como base da pesquisa em Libras, foi utilizado como material de apoio o Dicionário Enciclopédico da Língua de Sinais, o Capovilla, sendo este

[...] uma obra ímpar, de surpreendentemente vasto escopo e notável profundidade, para fazer jus à grande complexidade linguística da Libras e a um dos maiores desafios postos à ciência neuropsicológica que é o de desenvolver novos modelos científicos e

ferramentas tecnológicas capazes de subsidiar uma guinada científica na educação brasileira rumo à eficiência na produção da competência. (CAPOVILLA; RAPHAEL; MAURICIO, p.11)

Para a validação dos significados das palavras homônimas na língua portuguesa, foi utilizado como material de apoio o Dicionário Online Michaelis<sup>76</sup>.

Como resultado desse processo, foi produzido com um *corpus* de tradução contendo 57.206 sentenças em Português e em Libras alinhadas e seguindo o conjunto de regras proposto. Na Tabela 38, pode-se observar a ocorrência das regras no *corpus* de tradução, a quantidade de frases que possuem modificação.

**Tabela 38 - Ocorrência das regras no *Corpus***

| REGRA                           | QUANTIDADE DE FRASES |
|---------------------------------|----------------------|
| Verbos com concordância         | 1.930                |
| Frases negativas                | 7.569                |
| Homônimos e palavras homógrafas | 36.828               |
| Números cardinais               | 9.155                |
| Números ordinais                | 5                    |
| Intensidade                     | 1.170                |
| Palavras compostas              | 2.485                |
| Classificação de lugar          | 3.832                |
| Famosos                         | 72                   |
| Expressões                      | 8                    |
| Sem ocorrências especiais       | 28.594               |

Fonte: Dados da Pesquisa, 2020.

O *corpus* de tradução desenvolvido apresenta um léxico variado, tendo como mais frequentes as palavras: NÃO (9.485 ocorrências), EU (6.607 ocorrências), VOCÊ (5.622 ocorrências), TER (4.872 ocorrências), PAÍS (2.683 ocorrências), ELE (2.599 ocorrências).

As sentenças variam de tamanho, como pode ser observado na Tabela 39, permitindo que o treinamento seja realizado com uma maior diversidade de exemplos<sup>77</sup>.

**Tabela 39 - Variação de sentenças**

| CORPUS                |                           |
|-----------------------|---------------------------|
| MENOR SENTENÇA        |                           |
| SENTENÇA EM PORTUGUÊS | Ri.                       |
| APLICAÇÃO DAS REGRAS  | RIR [PONTO]               |
| CARACTERÍSTICAS       | 1 palavras; 3 caracteres. |

<sup>76</sup> <https://michaelis.uol.com.br/>

<sup>77</sup> Todas as palavras que apresentam caractere '&' para marcação de contexto, estão disponíveis da Tabela de Palavras Homônimas.

| <b>MAIOR SENTENÇA</b>       |  |
|-----------------------------|--|
| <b>TEXTO EM PORTUGUÊS</b>   | <p>Para os caninos em si, que se resfriam como muitos outros animais, existem sete doenças comuns e fatais, que também podem ser evitadas através de vacinação anual ou combatidas por meio de fortes tratamentos: a tosse canina, doença infecto-respiratória causada por uma bactéria, é tratável com antibióticos e comum onde muitos cães vivem juntos; a coronavírus, contraída quando um cão entra em contato com as fezes ou outras excreções de espécimes infectados, causa apatia e vômitos, e é tratada com abundância de líquidos e medicação; a cinomose, vista como a doença infecciosa mais fatal, é causada por vírus transmitido diretamente pelo ar, com tratamento eficaz apenas no primeiro de seus dois estágios e se em cães de sistema imunológico saudável; a hepatite infecciosa canina, doença viral espalhada por contato direto, causa febre, inchaço das amígdalas e dores estomacais, e é tratada com medicação adequada e tratamento intravenoso, cujos resultados aparecem em quatro dias; a leptospirose, causada por bactéria e transmitida através da urina, causa febre, depressão, letargia e perda de apetite, além de úlcera na boca e na língua, e tem como tratamento a internação e o uso de antibióticos; a parvovirose, doença altamente contagiosa que se espalha através das patas, pelo, saliva e fezes de um cão infectado ou por sapatos de pessoas, é considerada altamente fatal nos filhotes, tem como sintomas a diarreia aguda e o dano ao músculo cardíaco, e seu tratamento é feito através de medicação adequada e a obrigatória esterilização do ambiente, para que o cão não adoeça novamente com maior gravidade.</p> |
| <b>APLICAÇÃO DAS REGRAS</b> | <p>CANINO&amp;CACHORRO RESFRIAR COMO OUTRO ANIMAL EXISTIR 7 DOENÇA COMUM FATAL TAMBÉM PODER&amp;POSSIBILIDADE EVITAR VACINAÇÃO ANUAL COMBATER&amp;SUPRIMIR TRATAMENTO TOSSE CANINO&amp;CACHORRO DOENÇA INFECTO RESPIRATÓRIA BACTÉRIA TRATAR&amp;CUIDAR ANTIBIÓTICO COMUM VÁRIOS CÃO VIVER&amp;MORAR JUNTO CORONAVÍRUS CONTRAIR&amp;AFETAR QUANDO CÃO CONTATO&amp;TOQUE FEZES OUTRO EXCREÇÃO ESPÉCIME INFECTAR CAUSAR APATIA VÔMITO TRATAR&amp;CUIDAR ABUNDÂNCIA LÍQUIDO MEDICAÇÃO CINOMOSE DOENÇA INFECCIOSA MAIS&amp;QUANTIDADE FATAL CAUSAR VÍRUS TRANSMITIR&amp;PROPAGAR AR TRATAMENTO EFICAZ PRIMEIRO&amp;ORDINAL ESTÁGIO SE CÃO SISTEMA_IMUNOLÓGICO SAUDÁVEL HEPATITE INFECCIOSA CANINO&amp;CACHORRO DOENÇA VIRAL ESPALHAR&amp;DISSEMINAR CONTATO&amp;TOQUE CAUSAR FEBRE INCHAÇO AMÍGDALA DOR&amp;BARRIGA TRATAR&amp;CUIDAR MEDICAÇÃO ADEQUAR TRATAMENTO INTRAVENOSO RESULTADO 4 DIA LEPTOSPIROSE CAUSAR BACTÉRIA TRANSMITIR&amp;PROPAGAR URINA CAUSAR FEBRE DEPRESSÃO LETARGIA PERDER APETITE ÚLCERA BOCA LÍNGUA&amp;ÓRGÃO TRATAMENTO INTERNAÇÃO USAR ANTIBIÓTICO PARVOVIROSE</p>  |

---

DOENÇA PERIGO CONTÁGIO ESPALHAR&DISSEMINAR  
 PATA&ANIMAL SALIVA FEZES CÃO INFECTAR SAPATO  
 PESSOA FATAL FILHOTE TER SINTOMA DIARREIA  
 FORTE&INTENSIDADE DANO MÚSCULO CARDÍACO  
 TRATAMENTO MEDICAÇÃO ADEQUAR OBRIGATÓRIO  
 ESTERILIZAÇÃO AMBIENTE CÃO NÃO\_ADOECER  
 NOVAMENTE GRAVIDADE&SERIEDADE [PONTO]

---



---

**CARACTERÍSTICAS**

---

257 palavras; 1.609 caracteres.

Fonte: Dados da Pesquisa, 2020.

---

Baseado nos resultados apresentados (ver Tabela 38), pode-se observar que a regra aplicada de maior ocorrência foram as palavras homônimas e homógrafas (com 36.828 ocorrências), os numerais cardinais (com 9.155 ocorrências), e frases negativas (com 7.569 ocorrências).

Dando continuidade ao projeto, a equipe do VLibras realizou o treinamento de uma nova implementação do componente de tradução automática do VLibras que utilizou o *corpus* de tradução desenvolvido para tentar aprimorar a qualidade da sua tradução. Em seguida, foram realizados alguns testes com usuários surdos para avaliar se realmente o *corpus* desenvolvido era capaz de melhorar a qualidade da tradução do VLibras. Esses testes serão apresentados na Seção 5.2 e a análise dos resultados será apresentada na Seção 5.3.

## 5.2 Instrumentação

Para esta pesquisa foi realizada uma comparação objetiva entre duas versões do VLibras: uma versão que utiliza o *corpus* desenvolvido, que se denominou solução proposta, e outra versão descrita em Reis et al. 2020. Esta pesquisa empregou a mesma metodologia e instrumentos utilizados por Reis (2020), no entanto, foram realizadas algumas alterações na organização do questionário e outro tipo de *corpus*, com o intuito de melhorar a acessibilidade, pois de acordo com Guimarães (2014) “a utilização de escalas objetivando precisão dos resultados pode ser uma necessidade de avaliação, contudo avalia-se que respostas curtas sejam mais adequadas ao perfil de surdos.” (p. 76). O questionário foi dividido em 3 seções com 10 perguntas relacionadas ao perfil do usuário e 24 sentenças traduzidas. O processo para a seleção do corpus de sentenças foi diferente do utilizado anteriormente do estudo de caso que

[...] veio de uma análise do Dicionário Enciclopédico Ilustrado Trilíngue de Libras [Capovilla, Raphael e Mauricio 2009] proposto por Capovilla e Raphael, e de exemplos retirados de livros de Curso Básico de Libras de Felipe [Felipe 2007] e Quadros [Quadros 2009]. Mais especificamente, foi realizada uma pesquisa manual

para identificar sentenças que abordassem os aspectos gramaticais avaliados. Como resultado, foram obtidos exemplos de frases de Homonímia, Advérbios de Negação, Verbos Direcionais, Advérbios de Modo e Advérbios de Intensidade, totalizando mais de 527 exemplos de frases que abordam esses aspectos gramaticais. A seleção das sentenças a partir deste corpus, deu-se devido ao Corpus de Sentenças de Referência utilizado no teste computacional, não estar pronto para ser utilizado, neste momento da pesquisa. (REIS, 2020, p. 68)

Os participantes foram divididos em dois grupos: um com 6 intérpretes de Libras e outro com 10 surdos. Em relação aos vídeos de tradução, “[...] as sentenças de entrada em Português foram inseridas no VLibras para gerar as glosas em Libras de saída do tradutor.” (REIS, 2020, p. 63) Utilizando um script desenvolvido em linguagem de informação, gerando as traduções do VLibras a serem avaliadas. Apresentando sentenças de referência (chamadas GLOSA OURO) traduzidas para Libras e revisadas por um intérprete da língua de sinais.

Para esta pesquisa, os usuários surdos foram acompanhados por um mediador (intérprete de Libras) através de videoconferência, assim todas as dúvidas foram tiradas no momento da aplicação do questionário, e com a gravação foi feita uma avaliação posterior.

Os questionários foram aplicados entre os dias 09 a 23 de dezembro de 2020, individualmente, através do *Google Meet*<sup>78</sup>. Essa opção de aplicação a distância foi escolhida devido às dificuldades em encontrar os respondentes em um período de distanciamento social. Com a permissão dos participantes, os encontros foram gravados e estão disponíveis para a verificação de autenticidade da aplicação dos questionários<sup>79</sup>. No entanto, não será utilizada nenhuma imagem de caráter pessoal dos respondentes na divulgação dos resultados.

Os questionários foram caracterizados por sua tradução e aplicados a 30 respondentes surdos<sup>80</sup> que foram distribuídos aleatoriamente em três grupos de 10 integrantes. O Grupo A avaliou a tradução de sentenças usando a versão do VLibras que era baseada em Regras, ou seja, antes da criação do *corpus* de tradução desenvolvido neste trabalho. O Grupo B avaliou as sentenças traduzidas com uma nova versão do VLibras baseada apenas em Aprendizagem de Máquina (*Deep Learning*), com as modificações propostas nesta pesquisa. Por outro lado, o Grupo C avaliou as traduções usando uma versão Híbrida do VLibras, com traduções baseadas em regras e na aprendizagem de máquina.

Para realizar o teste, foram criadas três versões diferentes do questionário de avaliação (Questionários A, B e C), cada uma contendo 24 sentenças traduzidas para os Grupos A, B e C

<sup>78</sup> Apenas dois questionários foram aplicados presencialmente.

<sup>79</sup> <https://drive.google.com/drive/folders/1MrjWn87UPuXSBOwQVC9VjgNz06i3wT31?usp=sharing>

<sup>80</sup> Não foi solicitado no momento da aplicação do questionário explicações sobre o tipo de surdez/uso de prótese.

usando a versão do VLibras correspondente. Para cada sentença, três perguntas foram realizadas relacionadas ao tratamento dos aspectos gramaticais apresentados pelas sentenças.

|   |                       |                       |                       |                       |                       |                       |          |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|----------|
| a) Avalie a tradução quanto a corretude gramatical? |                       |                       |                       |                       |                       |                       |          |
|   | 1                     | 2                     | 3                     | 4                     | 5                     | 6                     |          |
| Péssimo   | <input type="radio"/> | Perfeito |
| b) Avalie a sua compreensão da tradução gerada?     |                       |                       |                       |                       |                       |                       |          |
|   | 1                     | 2                     | 3                     | 4                     | 5                     | 6                     |          |
| Péssimo   | <input type="radio"/> | Perfeito |
| c) Avalie a qualidade da tradução?                  |                       |                       |                       |                       |                       |                       |          |
|   | 1                     | 2                     | 3                     | 4                     | 5                     | 6                     |          |
| Péssimo   | <input type="radio"/> | Perfeito |

Adicionalmente, as sentenças foram agrupadas de acordo com os aspectos gramaticais apresentados na tradução: Traduções com palavras homônimas, com advérbios de negação, com verbos direcionais, com frases interrogativas, com frases exclamativas, com advérbios de modo e com advérbios de intensidade. As traduções do 1 ao 6, apresentam sentenças com palavras homônimas. Sendo elas S1: “A manga desta camisa está rasgada.”, S2: “O menino escalou o muro para fugir do cachorro.”, S3: “A ação do governo não conseguiu combater às drogas.”, S4: “A manga é uma fruta típica do verão.”, S5: “O diretor do hospital escalou o médico para o plantão.” e S6: “Os países decidiram dar uma trégua no combate.”

As sentenças S7, S8 e S9 envolvem o uso do advérbio de negação. Sendo elas S7: “Não adianta gritar, ele não ouviu.”, S8: “Não gosto desse seu jeito de agir.” e S9: “Não posso atendê-lo.” As sentenças S10, S11 e S12 compreendem o uso de verbos com concordância. Sendo elas S10: “Eu pergunto para você.”, S11: “Eu disse à vocês.” e S12: “Ele deu para mim.”. As sentenças S13, S14 e S15 apresentam frases interrogativas. Sendo elas S13: “Qual é o seu nome?”, S14: “Você é casado?” e S15: “Você quer água?”.

As traduções do 16 ao 18 compreendem as frases exclamativas. Sendo elas S17: “Carro bonito!”, S18: “Conheci muitos surdos!” e S19: “Saudades, você sumiu!”. As sentenças S19, S20 e S21 apresentam traduções com advérbios de modo. Sendo elas S19: “O homem lê o livro rapidamente.”, S20: “A mulher escova os dentes lentamente.” e S21: “Este método vai ser melhor que o anterior.”. As sentenças S22, S23 e S24 envolvem o uso de advérbios de intensidade. A S22: “Ela é pouco inteligente.”, S23: “Ela está mais gorda do que ele.” e S24: “O carnaval é muito alegre.”. Na seção (5.3.1), pode-se observar as informações relacionadas

aos 30 respondentes Surdos, lembrando que cada questionário foi avaliado por 10 usuários da língua de sinais.

### 5.3 Análise dos resultados

Nessa seção será apresentada a análise dos principais resultados obtidos nos testes com os usuários surdos. Na Seção 5.3.1, será analisado o perfil dos usuários envolvidos. Na Seção 5.3.2, serão apresentados os resultados dos testes das sentenças traduzidas pelo VLibras nas três versões avaliadas.

#### 5.3.1 Análise do Perfil dos Usuários

Na Tabela 40, são apresentadas as informações sobre a média de idade e a distribuição por gênero dos três grupos de usuários surdos que participaram do teste. Considerando todos os usuários do teste, juntando os três grupos, a média geral de idade geral dos participantes foi de 29,4.

**Tabela 40 - Idade e identidade de gênero**

|                    | <b>GRUPO A - REGRAS</b> | <b>GRUPO B - DL</b> | <b>GRUPO C - HÍBRIDO</b> |
|--------------------|-------------------------|---------------------|--------------------------|
| <b>IDADE MÉDIA</b> | 30,1                    | 29,1                | 28,9                     |
| <b>MASCULINO</b>   | 90%                     | 70%                 | 50%                      |
| <b>FEMININO</b>    | 10%                     | 30%                 | 50%                      |

Fonte: Dados da Pesquisa, 2021.

No Gráfico 1 são apresentadas as informações com relação ao nível de escolaridade dos participantes. De acordo com o Gráfico 1, é possível observar que os respondentes apresentam níveis de escolaridade variados, no entanto, a maioria possui Ensino Médio, seguido do Nível Superior Incompleto.



Fonte: Dados da Pesquisa, 2021.

Com relação ao acesso à tecnologia, apenas dois respondentes informaram que não utilizam computador, sendo um usuário do Grupo A e outro do Grupo B. Todos os demais (28) usuários afirmaram que utilizam computador com frequência.

Na Tabela 41, pode-se perceber que a maioria dos respondentes utilizam tradutor automático para Libras em seu dia-a-dia. Dessa forma, percebe-se que a avaliação das traduções em sua maioria foi realizada por usuários frequentes de aplicativos de tradução, não sendo especificados qual tradutor utilizam.

**Tabela 41** - Uso de aplicativos de Tradução para Libras

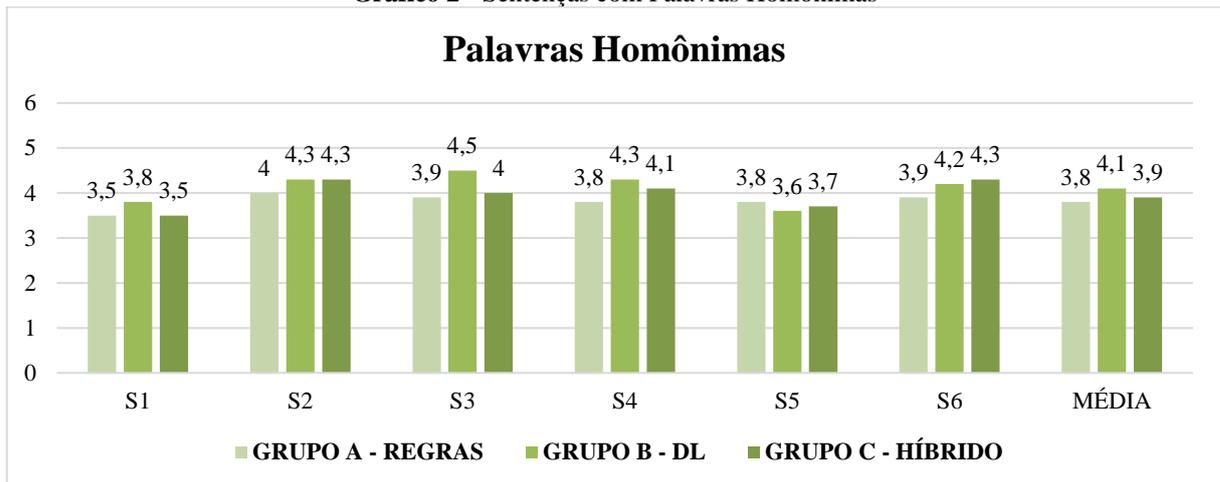
|   | <b>GRUPO A -<br/>REGRAS</b> | <b>GRUPO B - DL</b> | <b>GRUPO C -<br/>HÍBRIDO</b> |
|---|-----------------------------|---------------------|------------------------------|
| <b>Uso de aplicativos de<br/>Tradução para Libras</b> | Sim 8<br>Não 2              | Sim 7<br>Não 3      | Sim 8<br>Não 2               |

Fonte: Dados da Pesquisa, 2021.

Para a avaliação do contexto das sentenças traduzidas, os respondentes afirmaram possuir conhecimento nas duas línguas envolvidas (Português Brasileiro - PB e Libras). Dos 30 participantes, 73,33% informaram possuir nível “Perfeito” de entendimento em Libras, e 26,67% possuem nível “Bom”. Com relação ao nível de entendimento do PB, 10% afirmaram possuir nível “Perfeito”, 43,33% com nível “Bom”, 36,67% entendem “Mais ou menos” e 10% entendem “Pouco”.

### **5.3.2 Análise das sentenças traduzidas automaticamente pelo VLibras**

Com relação às sentenças traduzidas pelo VLibras, inicialmente, a análise dos resultados foi feita de acordo com o aspecto gramatical em Libras das sentenças utilizadas no teste. No Gráfico 2, são apresentados os resultados obtidos nos testes para as sentenças S1 a S6, que envolvem o uso de palavras homônimas.

**Gráfico 2 - Sentenças com Palavras Homônimas**

Fonte: Dados da Pesquisa, 2021.

Como observado, os Grupos B e C, que utilizam a base de dados proposta neste trabalho, apresentaram valores de média geral (4,1 e 3,9, respectivamente) superiores ao Grupo A (3,8), que utiliza apenas um conjunto de regras formais sem uma base de dados envolvida. A média geral inclui os resultados da avaliação relacionados aos aspectos gramaticais, como a correteza gramatical, compreensão e qualidade da tradução de cada tradução.

Durante a aplicação dos questionários, com a utilização da videoconferência, foram observados alguns comportamentos que possam justificar esta avaliação, como, por exemplo, a sinalização das palavras homônimas de acordo com o contexto da sentença.

Com relação às sentenças S1 e S4, por exemplo, a palavra MANGA foi sinalizada com sentido de fruta para os usuários do Grupo A - Regras. Nos Grupos B e C, no entanto, que envolviam versões dos tradutores que utilizam a base proposta neste trabalho, a palavra MANGA foi sinalizada de acordo com o contexto da sentença (MANGA&ROUPA<sup>81</sup> e MANGA&FRUTA<sup>82</sup>). Nas sentenças S2 e S5, o tradutor VLibras gerou uma única sinalização para a palavra ESCALAR, relacionadas a subir (ESCALAR&SUBIR<sup>83</sup>) nas duas sentenças para os três Grupos. Isso poderia justificar os valores próximos apresentados na avaliação.

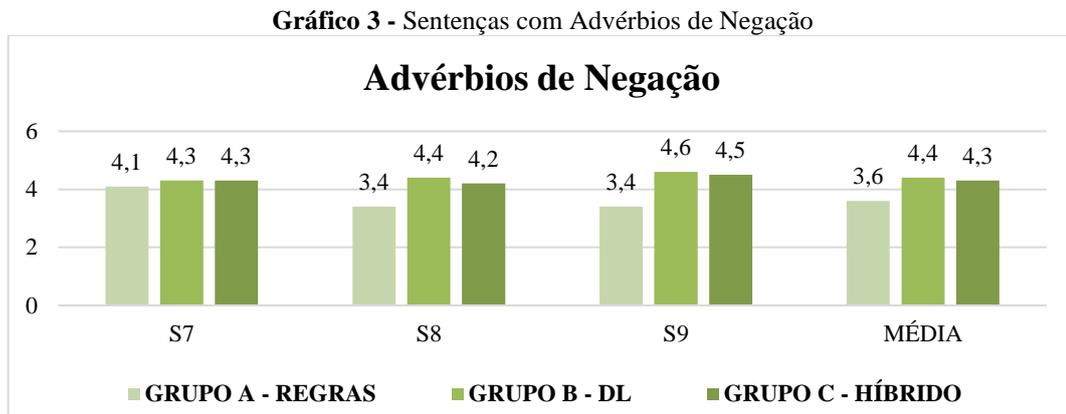
Nas sentenças S3 e S6, os Grupos A, B e C visualizaram a mesma sinalização para a palavra homônima COMBATER. No entanto, as traduções dos Grupos B e C utilizaram a negação adequada do verbo conseguir (NÃO\_CONSEGUIR), com a sinalização da negação simultânea ao verbo, e, portanto, de acordo com as regras gramaticais da Libras.

<sup>81</sup> Parte do vestuário que cobre o braço, cingindo-o, parcial ou totalmente.

<sup>82</sup> O fruto comestível da mangueira, de diferentes formas, com caroço volumoso, cuja polpa, de cor amarelada, é apreciada pelo sabor e aroma; mango.

<sup>83</sup> Subir uma superfície íngreme (muralha, fortaleza, obstáculos etc.), geralmente com a ajuda de escadas.

No Gráfico 3, são apresentados os principais resultados do teste com relação a esse aspecto gramatical.



Fonte: Dados da Pesquisa, 2021.

De acordo com o Gráfico 3, o Grupo B - DL obteve a maior média geral (4,4), seguido do Grupo C - HÍBRIDO (4,3) e do Grupo A - REGRAS (3,6). Como pode ser observado durante a aplicação dos testes, na sentença S7 foi apresentada uma sinalização diferente apenas na tradução do Grupo B, mesmo que essa sinalização seja considerada errada, pois está relacionada ao contexto de ADIANTAR&TEMPO<sup>84</sup>, e não de ADIANTAR&RESOLVER<sup>85</sup> apresentada nas traduções dos Grupos A e C. No entanto, provavelmente os surdos conhecem apenas o primeiro contexto, relacionando a palavra “adiantar” à antecipação, e não ao sentido de resolver algo. Na Figura 9, são apresentadas a sinalização<sup>86</sup> de ADIANTAR e sua negação correspondente (NÃO\_ADIANTAR).

**Figura 9 - Traduções da palavra Adiantar**



Fonte: Captura de tela do App VLibras, 2021.

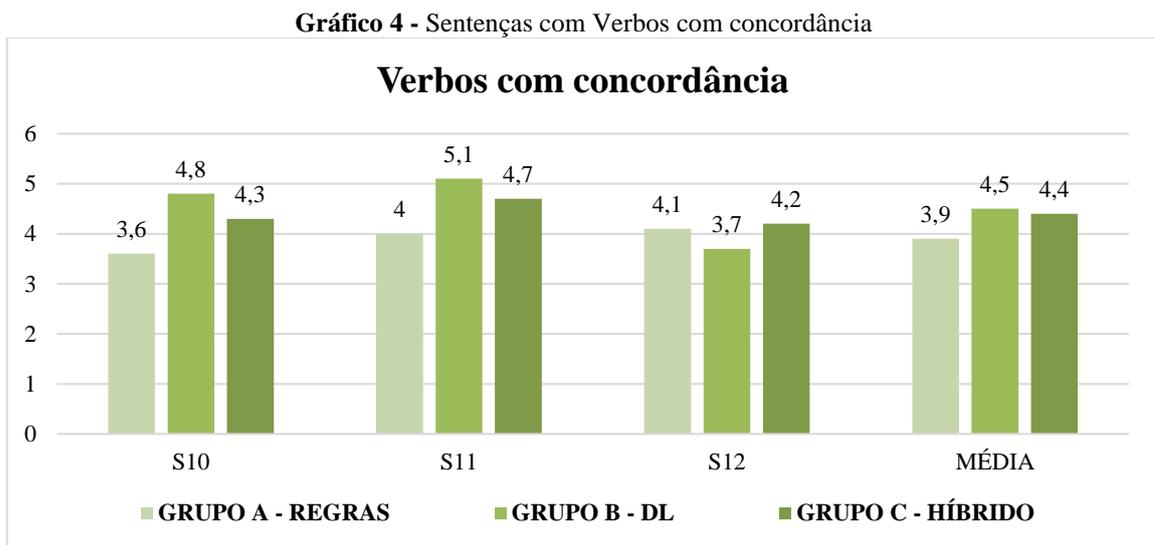
<sup>84</sup> Acontecer mais cedo do que o previsto.

<sup>85</sup> Valer a pena, ser útil ou resolver.

<sup>86</sup> As sinalizações podem ser encontradas no Link: <https://wiki.vlibras.gov.br/>

Nas sentenças S8 e S9, as traduções dos Grupos B e C ocorreram com a sinalização correta dos advérbios de negação (NÃO\_GOSTAR, NÃO\_PODER), enquanto as traduções do Grupo A relacionou a palavra “GOSTO” à sabor e não ao verbo gostar.

As sentenças S10, S11 e S12 envolvem o uso de verbos com concordância. No Gráfico 4, pode-se observar os resultados de acordo com a avaliação dos usuários surdos para as sentenças envolvendo esse tipo de concordância verbal.

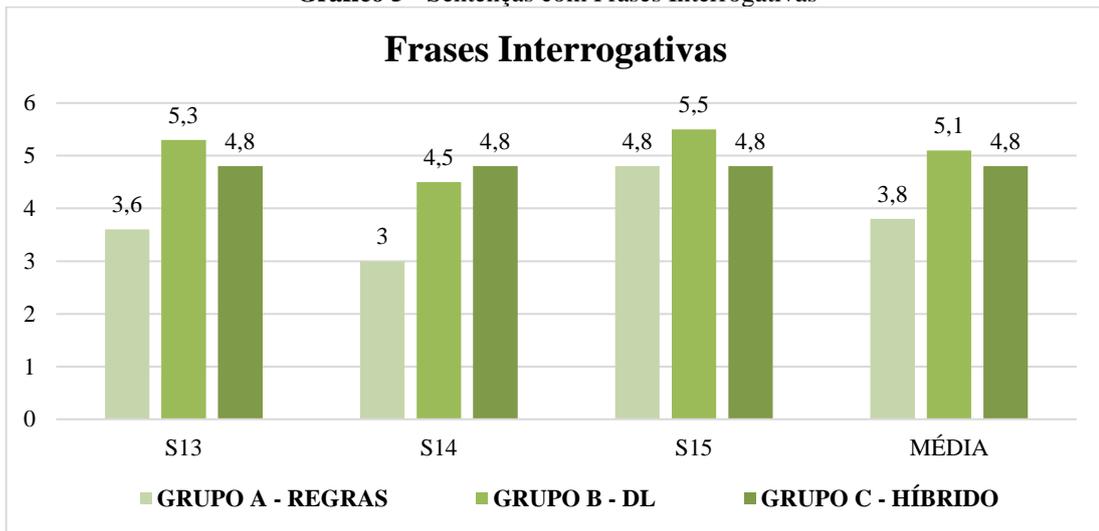


Fonte: Dados da Pesquisa, 2021.

Pode-se observar no Gráfico 4, que o Grupo B mais uma vez apresentou o maior valor médio geral (4,5), seguido do Grupo C (4,4) e do Grupo A (3,9). Dentre as traduções, o Grupo B obteve a maior média nas sentenças S10 e S11, mas obteve uma média menor na sentença S12. Essa avaliação pode ter ocorrido devido a erro na tradução. O avatar sinalizou “1S\_DAR\_2S”, sendo que a tradução correta é “2S\_DAR\_1S” sinalizado na tradução apresentada pelo Grupo C, que obteve o melhor resultado nesta sentença.

Os Gráficos 5 e 6 apresentam os resultados obtidos com relação às frases interrogativas (Sentenças S13 a S15) e exclamativas (Sentenças S16 a S18), respectivamente.

Gráfico 5 - Sentenças com Frases Interrogativas

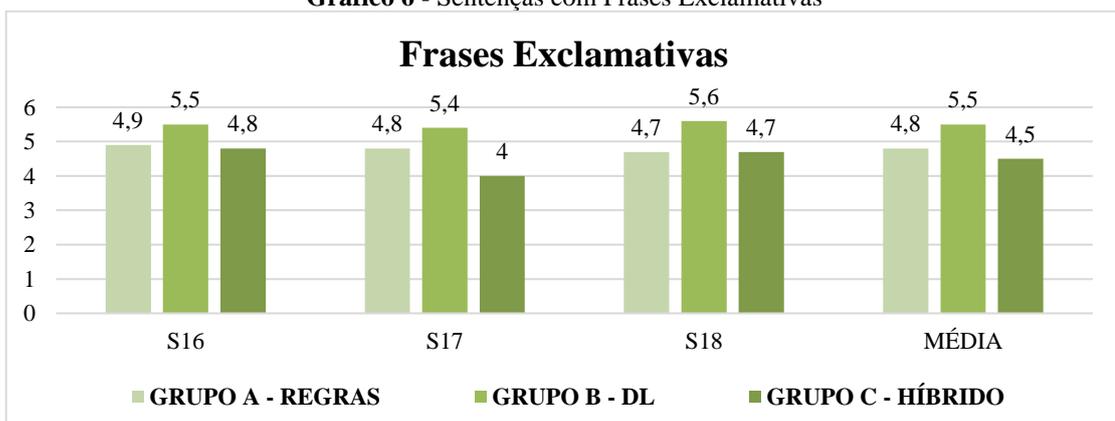


Fonte: Dados da Pesquisa, 2021.

Como observado no Gráfico 5, o Grupo B obteve o maior valor médio geral (5,1), seguido do Grupo C (4,8) e do Grupo A (3,8). Na tradução da S13, o Grupo B obteve um valor médio melhor. Nesse caso, a média baixa na avaliação do Grupo A pode ser justificada pela falta de expressão facial do avatar e sinalização do verbo SER, não sendo usualmente utilizada na Libras.

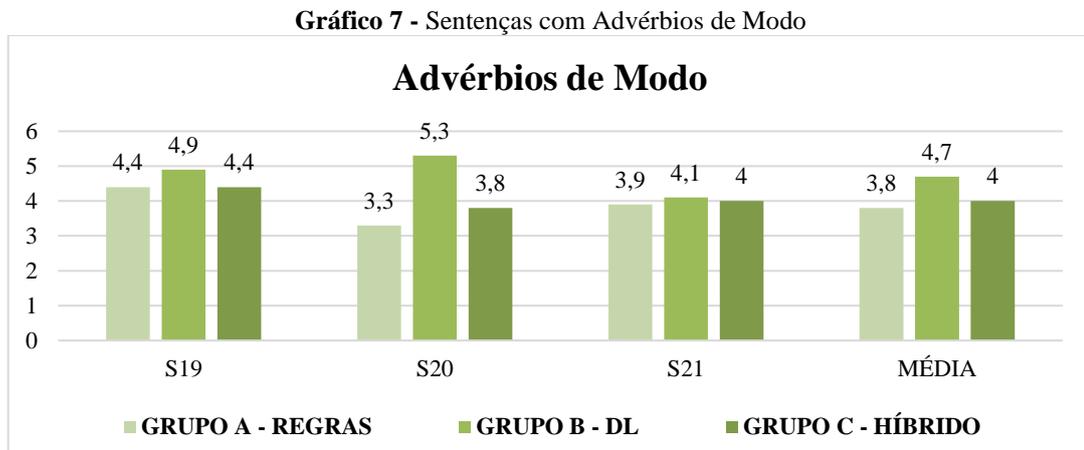
De acordo com o Gráfico, é possível notar que o Grupo B obteve o maior valor médio (5,5), seguido do Grupo A (4,8) e do Grupo B (4,5). Em frases exclamativas, a expressão facial se torna essencial na compreensão do contexto, dessa forma, o desenvolvimento de uma implementação no tradutor que inclua a expressão facial à frase completa e não como uma marcação de finalização da sentença pode ser considerada uma melhoria importante em pesquisas futuras.

Gráfico 6 - Sentenças com Frases Exclamativas



Fonte: Dados da Pesquisa, 2021.

No Gráfico 7, são apresentados os resultados do teste com relação aos advérbios de modo (Sentenças S19, S20 e S21).



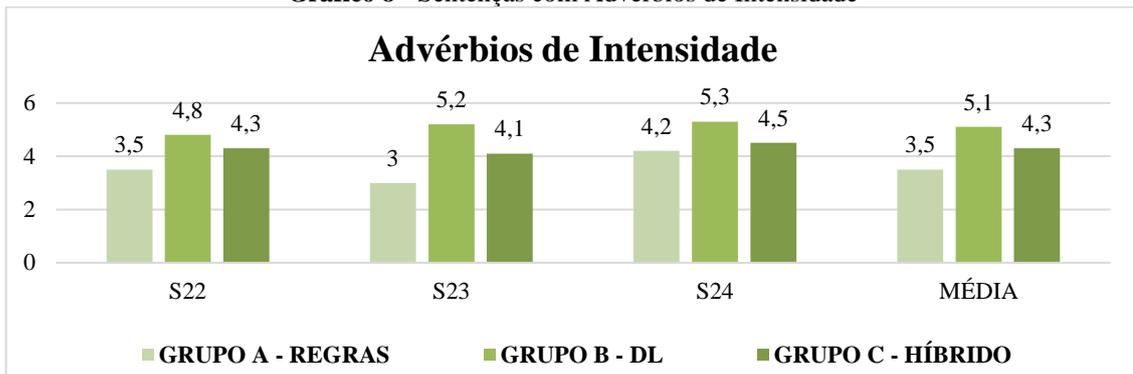
Fonte: Dados da Pesquisa, 2021.

De acordo com o Gráfico 7, pode-se observar que o Grupo B obteve novamente o maior valor médio geral (4,7), seguido do Grupo C (4,0) e do Grupo A (3,8). A S19 apresenta a mesma tradução nos três questionários, sendo esta uma das justificativas para as médias mais próximas alcançadas pela avaliação dos usuários. A S20, por outro lado, apresenta variações nas avaliações. A tradução do Grupo A apresentou uma contextualização equivocada para a palavra ESCOVA, utilizando o contexto de ESCOVA&CABELO<sup>87</sup> e não de ESCOVA&DENTE<sup>88</sup> apresentada nas traduções dos Grupos B e C. A S21 apresentou valores médios próximas. No entanto, a tradução do Grupo A possui um erro na contextualização, pois atribui o verbo “IR” para a tradução da palavra “VAI”, enquanto as traduções dos Grupos B e C são iguais, e removem o "IR" da tradução.

Por fim, no Gráfico 8, são apresentados os resultados relacionados às sentenças com advérbios de intensidade (S22, S23 e S24).

<sup>87</sup> Sinal próprio da Libras para pente de cabelo, concordância de gênero.

<sup>88</sup> Sinal próprio da Libras para escova de dente, concordância de gênero.

**Gráfico 8 - Sentenças com Advérbios de Intensidade**

Fonte: Dados da Pesquisa, 2021.

Como observado no Gráfico 8, o Grupo B obteve a maior média (5,1), seguido do Grupo C (4,3) e do Grupo A (3,5). A S22 apresentou variações nas traduções entre os grupos. Na tradução do Grupo A, o verbo SER foi utilizado. A tradução do Grupo B utilizou a forma correta de sinalização do advérbio de intensidade, enquanto a tradução do Grupo C excluiu a referência ao advérbio de intensidade POUCO. Na S23, a tradução do Grupo A utilizou o contexto de soma para a palavra homônima e advérbio de intensidade “MAIS” (MAIS&SOMAR<sup>89</sup>); a tradução do Grupo B traduziu o “MAIS” relacionado ao contexto de quantidade (MAIS&QUANTIDADE<sup>90</sup>); enquanto que a tradução do Grupo C sinalizou a intensidade junto ao sinal GORDO(+). Na S24, a tradução do Grupo A voltou a empregar o verbo SER, e os Grupos B e C apresentaram a mesma tradução. Apesar disso, a avaliação média desta sentença nos Grupos B e C foi diferente, sendo talvez influenciada pela experiência dos usuários com a língua.

Diante disso, pode-se avaliar (ver Tabelas 42, 43 e 44) o desempenho das três versões do tradutor do VLibras com base na média geral alcançada para todos os aspectos gramaticais avaliados, com relação a corretude gramatical, compreensão e qualidade da tradução. De acordo com a Tabela 43, no Grupo A, que envolveu traduções baseadas em regras, o valor médio geral obtido foi de 3,9. Alguns surdos comentaram sobre a utilização de palavras em contextos errados, e sinalização da palavra SER como os principais motivos para essa avaliação.

**Tabela 42 - Média Final do Grupo A - Regras**

| GRUPO A - REGRAS     | AValiação  |
|----------------------|------------|
| Corretude Gramatical | 3,9        |
| Compreensão          | 3,9        |
| Qualidade            | 3,9        |
| <b>Média Geral</b>   | <b>3,9</b> |

Fonte: Dados da Pesquisa, 2021.

<sup>89</sup> Sinal próprio da Libras, sinalizando somar.

<sup>90</sup> Em maior quantidade ou número.

De acordo com a Tabela 43, no Grupo B, que envolveu traduções baseadas na Aprendizagem de Máquina com base de dados proposta neste trabalho, o valor médio geral para os três aspectos avaliados foi de 4,7. As traduções do Grupo B obtiveram a melhor avaliação tanto na média geral, quanto na análise da corretude gramatical, compreensão e qualidade da tradução. Mesmo alcançando uma média razoável, a proposta apresentada por esta pesquisa ainda apresenta falhas na contextualização e necessidade de melhorias no resultado final. No entanto, para esta pesquisa, as traduções alcançaram seu objetivo, que seria a utilização das regras desenvolvidas aplicadas a um *corpus* de tradução para o aperfeiçoamento da sentença final gerada pelo tradutor automático VLibras.

**Tabela 43 - Média Final do Grupo B- DL**

| <b>GRUPO B - DL</b>  | <b>AValiação</b> |
|----------------------|------------------|
| Corretude Gramatical | 4,7              |
| Compreensão          | 4,8              |
| Qualidade            | 4,6              |
| <b>Média Geral</b>   | <b>4,7</b>       |

Fonte: Dados da Pesquisa, 2021.

De acordo com a Tabela 44, o Grupo C, que avaliou traduções utilizando abordagem Híbrida, combinando traduções baseadas em Regras e Aprendizagem de Máquina, obteve um valor médio geral de 4,3. A versão híbrida do tradutor automático foi mais bem avaliada que a versão de tradução baseada em regras do Grupo A, o que é um indício de que combinar as melhores características de cada modelo de tradução pode resultar em uma opção viável e aceitável pela comunidade dos usuários da língua de sinais.

**Tabela 44 - Média Final do Grupo C - Híbrido**

| <b>GRUPO C - HÍBRIDO</b> | <b>AValiação</b> |
|--------------------------|------------------|
| Corretude Gramatical     | 4,4              |
| Compreensão              | 4,3              |
| Qualidade                | 4,2              |
| <b>Média Geral</b>       | <b>4,3</b>       |

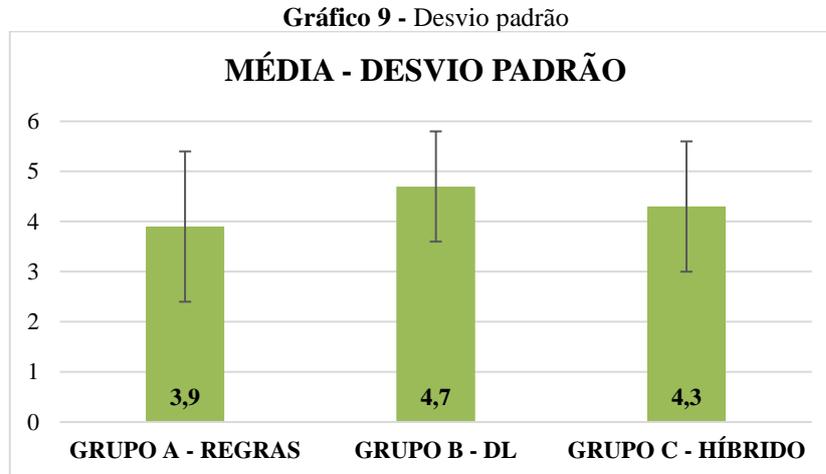
Fonte: Dados da Pesquisa, 2021.

Na Tabela 45 e no Gráfico 9, são apresentados também os resultados com relação a dispersão dos dados nos testes realizados. De acordo com a Tabela 46, apesar de apresentar um desvio padrão alto em todos os grupos, o Grupo B apresentou a maior média e o menor valor de desvio padrão.

**Tabela 45** - Média Final e desvio padrão dos Resultados

|                      | <b>GRUPO A - REGRAS</b> | <b>GRUPO B - DL</b> | <b>GRUPO C - HÍBRIDO</b> |
|----------------------|-------------------------|---------------------|--------------------------|
| <b>DESVIO PADRÃO</b> | 1,5                     | 1,1                 | 1,3                      |
| <b>MÉDIA GERAL</b>   | <b>3,9</b>              | <b>4,7</b>          | <b>4,3</b>               |

Fonte: Dados da Pesquisa, 2021.



Fonte: Dados da Pesquisa, 2021.

A partir dos dados apresentados, pode-se observar que a solução proposta de um *corpus* bilíngue desenvolvido com regras gramaticalmente pré-definidas, alcançou uma melhor média na avaliação dos aspectos gramaticais realizados pelos usuários da Libras. Por fim, verificou-se a necessidade de continuar o desenvolvimento de um *corpus* bilíngue maior e mais abrangente, visto que as regras desenvolvidas não englobam todas as características gramaticais da língua de sinais.

## 6 CONSIDERAÇÕES FINAIS

Esta pesquisa possibilitou a modelagem e desenvolvimento de uma base de dados de tradução baseada em exemplos, com o propósito de auxiliar na melhoria de tradutores automáticos para Libras que atuam no processo de tradução de conteúdos presentes em plataformas digitais do Português Brasileiro para a Libras. Essa proposta incluiu o desenvolvimento de uma linguagem formal de descrição com pares de exemplos, compreendendo a estrutura gramatical das línguas envolvidas, e adaptadas às necessidades dos tradutores automáticos. Neste caso, sendo o Português Brasileiro e a Libras, de modo que todas as disposições foram desenvolvidas seguindo alguns dos principais aspectos da estrutura gramatical da língua de sinais.

As regras desenvolvidas e aplicadas no *corpus* de tradução procuraram abranger algumas das principais características da estrutura gramatical da Libras. No entanto, houve a necessidade de alterar descrições e uso de caracteres, devido a questões computacionais, para facilitar o treinamento. Essas modificações evidenciaram a importância das adequações morfológicas, sintáticas e semânticas da tradução de uma língua oral para outra gesto-visual, garantindo, assim, a contextualização da informação.

Durante o desenvolvimento deste trabalho, foi possível perceber que a construção de uma base de dados de tradução não deve ser considerada como uma atividade limitada, mas sim em constante evolução. Pois, as regras aplicadas no *corpus* desenvolvido não conseguem explorar todos os casos existentes de determinado assunto. No entanto, ela apresenta a linguagem formal de descrição com regras morfológicas viabilizando uma melhoria no estado da arte e da técnica dos tradutores automáticos para Libras, que atualmente são, em geral, baseados em regras.

Para validar as regras criadas e o *corpus* desenvolvido, foi aplicado o *corpus* de tradução desenvolvido na próxima versão do componente de tradução automática do VLibras que contém um módulo de tradução automática neural, utilizando o *corpus* proposto para melhorar a qualidade das traduções geradas.

Como proposta de trabalhos futuros pode-se citar a expansão da base de dados do *corpus* bilíngue, visto que para o treinamento de máquina a quantidade atual de sentenças ainda é considerada relativamente pequena. Assim como, melhorar as regras desenvolvidas nesta pesquisa, com o intuito de abranger mais aspectos gramaticais da língua de sinais. Nesta mesma

linha de pesquisa, há a necessidade de ampliar o Wikilibras<sup>91</sup>, com estratégias de engajamento visando a participação da comunidade surda na avaliação do dicionário do tradutor automático VLibras.

Como outro meio de aperfeiçoar as regras desenvolvidas, há a possibilidade de uma participação mais ativa da comunidade surda, como a elaboração e aplicação de entrevistas para avaliar a qualidade das sentenças traduzidas. Dessa forma, pode-se ampliar as características da língua de sinais exploradas na base de dados, e inseri-las no *corpus* bilíngue.

---

<sup>91</sup> <https://wiki.vlibras.gov.br/>

## REFERÊNCIAS

- AGUIAR, Thiago Cardoso; CHAIBUE, Karime. Histórico das Escritas de Línguas de Sinais. **Revista Virtual de Cultura Surda**, v. 15, 2015.
- ALBUQUERQUE, Dilainne D. **Inclusão digital e informacional de estudantes surdos: ouvindo o silêncio**. 2015. 84 f. Monografia (Graduação em Biblioteconomia) – Universidade Federal da Paraíba, João Pessoa, 2015.
- ALMEIDA, Magno Pinheiro; ALMEIDA, Miguel Eugênio. História de LIBRAS: característica e sua estrutura. **Revista Philologus**, Ano 18, Nº 54 – Suplemento: Anais da VII JNLFLP. Rio de Janeiro: CiFEFiL, 2012.
- ARAÚJO, Tiago M. U. et al. Automatic generation of Brazilian sign language windows for digital TV systems. **Journal of the Brazilian Computer Society**, v. 19, n. 2, p. 107-125, 2013.
- BATISTA, Gustavo Enrique de Almeida Prado et al. **Pré-processamento de dados em aprendizado de máquina supervisionado**. 2003. Tese de Doutorado. Universidade de São Paulo.
- BECKER, Karin; TUMITAN, Diego. **Introdução à mineração de opiniões: Conceitos, aplicações e desafios**. Simpósio brasileiro de banco de dados, v. 75, 2013.
- BISOL, Leda (Ed.). **Introdução a estudos de fonologia do português brasileiro**. EdIPUCRS, 2005.
- BRITO, Lucinda Ferreira. **A gramática de Libras**. 2013.
- \_\_\_\_\_. Uma abordagem fonológica dos sinais da LSCB. **Revista Espaço**, n. 1, 1990.
- \_\_\_\_\_. Estrutura lingüística da Libras. **Introdução à Língua Brasileira de Sinais**. Brasília: APADA/DF, 2002.
- BRITO, FB de; NEVES, Sylvia Lia Grespan; XAVIER, André Nogueira. O movimento surdo e sua luta pelo reconhecimento da Libras e pela construção de uma política linguística no Brasil. **Libras em estudo: política linguística**, v. 1, p. 67-104, 2013.
- CAPOVILLA, Fernando César; RAPHAEL, Walkiria Duarte. **Enciclopédia da língua de sinais brasileiras: o mundo do surdo em libras**. Edusp, 2004.
- CAPOVILLA, Fernando César; RAPHAEL, Walkiria Duarte; MARTINS, Antonielle Cantarelli; TEMOTEO, Janice Gonçalves. **Dicionário da Língua de Sinais do Brasil: a libras em suas mãos**. [S.l: s.n.], 2017.
- CAPOVILLA, Fernando C. Filosofias educacionais em relação ao surdo: do oralismo à comunicação total ao bilinguismo. **Revista Brasileira de Educação Especial**, v. 6, n. 1, p. 99-116, 2000.

CASELI, Helena M.; NUNES, Israel A. **Tradução Automática Estatística baseada em Frases e Fatorada**: Experimentos com os idiomas Português do Brasil e Inglês usando o toolkit Moses. 2009.

CASSIANO, Paulo Victor. O surdo e seus direitos: os dispositivos da lei 10.436 e do decreto 5.626. **Revista virtual de cultura surda**, v. 21, n. 1, p. 1-28, 2017.

CHINEA-RIOS, Mara. **Advanced techniques for domain adaptation in Statistical Machine Translation**. 2019. Tese de Doutorado.

CHINEA-RIOS, Mara; SANCHIS-TRILLES, Germán; CASACUBERTA, Francisco. Vector sentences representation for data selection in statistical machine translation. **Computer Speech & Language**, v. 56, p. 1-16, 2019.

CHIU, Yu-Hsien et al. Joint optimization of word alignment and epenthesis generation for Chinese to Taiwanese sign synthesis. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 29, n. 1, p. 28-39, 2006.

COSTA, A. B.; ZOLTOWSKI, A. P. C. Como escrever um artigo de revisão sistemática. **Manual de produção científica**, p. 55-70, 2014.

COTOVICZ, Marcio et al. Libras: algumas reflexões sobre a sintaxe. **Revista Odisseia**, v. 3, n. 1, p. 16-35, 2018.

DICIONÁRIO Michaelis. Dicionário online Uol. **Editora Melhoramentos**, 2020.

D'HARO, Luis Fernando et al. Language model adaptation for a speech to sign language translation system using web frequencies and a map framework. In: **Ninth Annual Conference of the International Speech Communication Association**. 2008.

EBLING, Sarah. Building a parallel corpus of German/Swiss German Sign Language train announcements. **International Journal of Corpus Linguistics**, v. 21, n. 1, p. 116-129, 2016.

EL MAAZOUZI, Zakaria et al. Development of a bilingual corpus of Arabic and Arabic Sign Language based on a signed content. In: **2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)**. IEEE, 2016. p. 349-354.

FARJADO, I. et al. Mapeamento Estruturado de LIBRAS para utilização em sistemas de comunicação. In: **1st International Workshop on Assistive Technology**. 2015.

FELIPE, Tanya Amara; MONTEIRO, Myrna. **Libras em Contexto: Curso Básico - Livro do Professor**. ed. 6. Brasília/DF: Programa Nacional de Apoio à Educação dos Surdos, MEC: SEEP, 2007.

FELIPE, Tanya Amara. Os processos de formação de palavras na Libras. **ETD-Educação Temática Digital**, v. 7, n. 2, p. 200-217, 2006.

\_\_\_\_\_. **A relação sintático-semântica dos verbos na Língua Brasileira de Sinais (LIBRAS)**. Tese de Doutorado em Linguística e Filologia. 1998. V.1: 01-143; V.2: 144- 298;

V.3: Transcrição dos dados das Fitas: 1,2,3,4,5,6 (Banco de dados). Universidade Federal do Rio de Janeiro. Rio de Janeiro.

\_\_\_\_\_. **Libras em contexto:** curso básico - Livro do estudante. 8ª. edição, 2007.

\_\_\_\_\_. Sistema de Flexão Verbal na LIBRAS: Os classificadores enquanto marcadores de flexão de gênero. In: **Anais do Congresso Surdez e Pós-Modernidade: Novos rumos para a educação brasileira.** 2002. p. 37-58.

\_\_\_\_\_. Diferentes políticas e diferentes contextos educacionais: educação bilíngue para educandos surdos x educação bilíngue inclusiva. **Revista Espaço**, p. 189-220, 2018.

FERNANDES, Carla Caballero; ROMEIRO, Camila de A. A contribuição da datilologia como estratégia metodológica no processo de alfabetização. **Revista diálogos**, v. 4, n. 1, p. 170-185, 2017.

FERNANDES, Sueli; STROBEL, Karin Lilian. **Aspectos lingüísticos da LIBRAS.** Curitiba: SEED/SUED/DEE, 1998.

FINAU, Rossana Aparecida. **Os sinais de tempo e aspecto na LIBRAS.** 2004.

FONSECA, Ricardo Tadeu Marques da. O novo conceito constitucional de pessoa com deficiência: um ato de coragem. **Revista do TRT da 2ª Região**, São Paulo, n. 10, 2012.

FREIRE, Maria José Duarte. A tradução e interpretação de provérbios e expressões idiomáticas em Língua gestual: equivalentes lingüísticos e culturais. **Exedra: Revista Científica**, p. 121-128, 2012.

FREITAS, Victor A.; MEDEIROS, Sheyla Natália; SOUSA, Hercilio M. Importância do software Vlibras no processo de aprendizagem de pessoas com deficiência auditiva. **Nuevas Ideas en Informática Educativa**, Volumen 13, p. 578 – 581, 2017. Disponível em: <http://www.tise.cl/volumen13/TISE2017/81.pdf> Acesso em: 02 ago. 2021.

GOLDFELD, Márcia. **A criança surda:** linguagem e cognição numa perspectiva sociointeracionista. Plexus Editora, 1997.

GUIMARÃES, Cayley et al. A expressão facial é parte integrante da língua de sinais-libras como l2. **Educere et Educare**, v. 13, n. 28, p. 10-17648/educare. v13i28. 18372.

HORST, Paulo Sergio. **Avaliação do conhecimento adquirido por algoritmos de aprendizado de máquina utilizando exemplos.** 1999. Tese de Doutorado. Universidade de São Paulo.

HUMBLÉ, Philippe. Os estudos da tradução e os dicionários. **Trabalhos em Linguística Aplicada**, v. 44, n. 2, p. 233-246, 2005.

KARNOPP, Lodenir B. **Fonética e fonologia.** Florianópolis: UFSC, 2006. Disponível em: [https://www.libras.ufsc.br/colecaoLetrasLibras/eixoFormacaoBasica/foneticaEFonologia/assets/359/FoneticaFonologia\\_TextoBase.pdf](https://www.libras.ufsc.br/colecaoLetrasLibras/eixoFormacaoBasica/foneticaEFonologia/assets/359/FoneticaFonologia_TextoBase.pdf) Acesso em: 11 mar. 2021.

KOUREMENOS, Dimitrios et al. Statistical Machine Translation for Greek to Greek Sign Language Using Parallel Corpora Produced via Rule-Based Machine Translation. **CEUR Workshop Proceedings**, 2252, pp. 28-42, 2018.

LACERDA, Cristina B.F. de. Um pouco da história das diferentes abordagens na educação dos surdos. **Cad. CEDES**, Campinas, v. 19, n. 46, p. 68-80, Sept. 1998. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0101-32621998000300007&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-32621998000300007&lng=en&nrm=iso) Acesso em: 08 mar. 2021.

LIMA, Manuella Aschoff Cavalcanti Brandão. **Tradução automática com adequação sintático-semântica para LIBRAS**. 2015.

KLEIN, Madalena. Movimentos surdos e os discursos sobre surdez, educação e trabalho: a constituição do surdo trabalhador. **Sinais**, v. 10, p. 2, 1995.

KUBASKI, Cristiane; MORAES, Violeta Porto. O bilinguismo como proposta educacional para crianças surdas. In: **IX Congresso Nacional de Educação–EDUCERE–III Encontro Sul Brasileiro de Psicopedagogia, PUCPR, PR**. 2009. p. 3415. Disponível em: [http://www.educadores.diaadia.pr.gov.br/arquivos/File/artigos\\_edespecial/bilinguismo.pdf](http://www.educadores.diaadia.pr.gov.br/arquivos/File/artigos_edespecial/bilinguismo.pdf) Acesso em: 09 mar. 2021.

MARCON, Andréia Mendiola. O Papel do Tradutor/intérprete de Libras na compreensão de conceitos pelo Surdo. **ReVEL**, v. 10, n. 19, 2012.

MARCONI, Marina de Andrade. LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

MÁXIMO, Nídia Nunes; DE SOUZA, Wilma Pastor. **Argumentação na Libras: reflexões sobre estratégias visuo-espaciais**. 2014.

MCCLEARY, Leland; VIOTTI, Evani. **Semântica e pragmática**. Florianópolis: UFSC-CCE, 2009.

MISHRA, Gouri Sankar; SAHOO, Ashok Kumar; RAVULAKOLLU, Kiran Kumar. Word based statistical machine translation from English text to Indian sign language. **ARN Journal of Engineering and Applied Sciences**, v. 12, n. 2, 2017.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, v. 1, n. 1, p. 32, 2003.

MONTEIRO, Myrna Salerno. História dos movimentos dos surdos e o reconhecimento da Libras no Brasil. **ETD-Educação Temática Digital**, v. 7, n. 2, p. 292-305, 2006.

NASCIMENTO, Vinícius. Interpretação da Libras para o português na modalidade oral: considerações dialógicas. **Tradução & Comunicação**, v. 24, 2015.

NIH, **National Institutes of Health**.

NUNES, Sylvia da Silveira et al. Surdez e educação: escolas inclusivas e/ou bilíngues? **Psicol. Esc. Educ.**, Maringá, v. 19, n. 3, p. 537-545, Dec. 2015. Disponível em:

[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-85572015000300537&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-85572015000300537&lng=en&nrm=iso) Acesso em: 09 mar. 2021.

NUNES, Maria das Graças Volpe. Processamento de línguas naturais: para quê e para quem? **ICMC-USP**, 2008.

OTHMAN, Achraf; JEMNI, Mohamed. Statistical sign language machine translation: from english written text to american sign language gloss. **arXiv preprint arXiv:1112.0168**, 2011.

PAIVA, Francisco Aulísio dos Santos et al. Um sistema de transcrição para língua de sinais brasileira: o caso de um avatar. **Revista do GEL**, 2016.

PASSOS, Rosana. **Parâmetros Físicos do Movimento em Libras**: um estudo sobre intensificadores. Tese (doutorado) – Universidade Federal de Minas Gerais, 2014.

PERLIN, Gladis. STROBEL, Karin. (Professoras Autoras). **Fundamentos Da Educação De Surdos**. Universidade Federal de Santa Catarina. Florianópolis, 2008. Texto adaptado.

PESSANHA, Márlon Caetano Ramos; COZENDEY, Sabrina Gomes. Significação e Sentido no ensino inclusivo de Física mediado por intérpretes de Libras: uma perspectiva Bakhtiniana. **VIII ENPEC** (Encontro Nacional de Pesquisa em Educação em Ciências), 2011.

PIZZIO, Aline Lemos. et al. Tópicos de linguística aplicados à Língua de Sinais: Sintaxe. **Língua Brasileira de Sinais II**. Florianópolis, Brasil: Universidade Federal de Santa Catarina, 2008.

PORTA, Jordi et al. A rule-based translation from written Spanish to Spanish Sign Language glosses. **Computer Speech & Language**, v. 28, n. 3, p. 788-811, 2014.

QUADROS, Ronice M.; KARNOPP, Lodenir Becker. **Língua de sinais brasileira**: estudos linguísticos. Porto Alegre: Artmed, v. 1, p. 222, 2004.

QUADROS, R M. et al. **Língua Brasileira de Sinais I**, Texto-base para o curso de licenciatura em Letras-Libras a distância. Florianópolis: UFSC, 2009.

QUADROS, Ronice M.; CRUZ, Carina Rebello. **Língua de sinais**. Artmed Editora, 2009.

QUADROS, Ronice M. **Um capítulo da história do SignWriting**. v. 10, p. 03-05, 2004.

\_\_\_\_\_. A transcrição de textos do Corpus de Libras. **Leitura**, v. 1, n. 57, p. 8-34, 2016.

\_\_\_\_\_. **Educação de surdos**: a aquisição da linguagem. Artmed Editora, 2009.

RAMOS, Clélia Regina. **LIBRAS**: a língua de sinais dos surdos brasileiros. Disponível para download na página da **Editora Arara Azul**: <http://www.editora-arara-azul.com.br/pdf/artigo2.pdf>, 2004.

REIS, Luana Silva et al. Assessment of the Treatment of Grammatical Aspects of Machine Translators to Libras. In: **Anais Estendidos do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web**. SBC, 2018. p. 73-76.

REIS, Luana Silva et al. **Uma metodologia multinível de avaliação da qualidade da tradução gerada por tradutores automáticos para libras**. 2020.

ROCHA, Carlos Alberto de M.; ROCHA, Carlos Eduardo Penna de M. **Dicionário de locuções e expressões da língua portuguesa**. LEXIKON Editora, 2011.

RODRIGUES, Carlos Henrique; SANTOS, S. A. A interpretação e a tradução de/para línguas de sinais: contextos de serviços públicos e suas demandas. **Tradução em Revista**, v. 24, n. 1, p. 2-29, 2018.

SALLES, Heloisa Maria Moreira Lima et al. **Ensino de língua portuguesa para surdos: caminhos para prática pedagógica**. Brasília: MEC, SEESP, 2004. v. 2.

SANTOS, Carlene da Penha. **Políticas inclusivas e a formação do Trabalhador Intérprete da Libras (TILS) atuante no Ensino Superior**. Dissertação (Mestrado) - UFPB/Educação, 2018.

SANTOS, Jaquelma Teles Pereira. **Avaliação da interpretação literal e cultural de ditados populares brasileiros para a Língua de Sinais Brasileira**. 2018.

SANTOS, Guilherme Spolavori dos et al. **Produção de textos paralelos em língua portuguesa e uma interlíngua de Libras**. 2009.

SANTOS, E F. et al. Morfologia da Libras e a (re)afirmação linguística: o óbvio que ainda precisa ser dito. **Cadernos do CNLF**, Vol. XVII, Nº 12. Rio de Janeiro: CiFEFiL, 2013

SILVA, Alan David Sousa et al. **Variação fonológica e lexical em Libras**. 2020. 130f. Dissertação (mestrado em Linguística e Literatura) - Universidade Federal de Alagoas, Maceió, 2020.

SILVA, Giselli Mara da. **Parâmetros da LIBRAS**. Universidade Federal de Minas Gerais. Ed. Belo Horizonte, 2007. Disponível em: [http://www.lettras.ufmg.br/padrao\\_cms/documentos/eventos/dialogosdeinclusao/Parametros\\_da\\_Libras.pdf](http://www.lettras.ufmg.br/padrao_cms/documentos/eventos/dialogosdeinclusao/Parametros_da_Libras.pdf) Acesso em: 03 ago. 2021.

SILVA, Marília da Piedade Marinho. A semântica como negociação dos significados em Libras. **Trabalhos em Linguística Aplicada**, v. 45, n. 2, p. 255-269, 2006.

SILVA, T. P. da. **Tecnologias Assistivas e Inclusão Digital: A importância do computador no processo de aprendizagem de crianças com deficiência intelectual da APAE- Campina Grande/ PB**. 2014. 33f. Trabalho de Conclusão de Curso (Graduação em Pedagogia) - Universidade Estadual da Paraíba, Campina Grande, 2014.

SOUSA, Danielle Vanessa Costa. Um olhar sobre os aspectos linguísticos da língua brasileira de sinais. **Littera**, v. 1, n. 2, 2011.

- STEIN, Daniel; SCHMIDT, Christoph; NEY, Hermann. Analysis, preparation, and optimization of statistical sign language machine translation. **Machine Translation**, v. 26, n. 4, p. 325-357, 2012.
- STROBEL, Karin Lilian. **Surdos: vestígios culturais não registrados na história. Florianópolis**, 2008. Tese de Doutorado. Tese de Doutorado em Educação–UFSC- Universidade Federal de Santa Catarina.
- STUMPF, Marianne Rossi. **Aprendizagem de escrita de língua de sinais pelo sistema signwriting: língua de sinais no papel e no computador**. Tese (Doutorado) – Porto Alegre: UFRGS, CINTED, PGIE, 2005.
- STUPIELLO, É.N.A. A aplicação de sistemas de memória de tradução como ferramentas de produtividade para o tradutor. In: **Ética profissional na tradução assistida por sistemas de memórias** [online]. São Paulo: Editora UNESP, 2014, pp. 75-139.
- SU, Hung-Yu; WU, Chung-Hsien. Improving structural statistical machine translation for sign language with small corpus using thematic role templates as translation memory. **IEEE Transactions on Audio, Speech, and Language Processing**, v. 17, n. 7, p. 1305-1315, 2009.
- SUTTON, Valerie. **Um sistema de escrita para língua de sinais**. Trad. Marianne Rossi Stumpf, 2000.
- TEIXEIRA, Vanessa Gomes; LEITÃO, Catarina Modesto de Carvalho. Flexão verbal em libras e em língua portuguesa: análise contrastiva. **Revista Philologus**, p. 31-43.
- TMAR, Zouhour; OTHMAN, Achraf; JEMNI, Mohamed. A rule-based approach for building an artificial English-ASL corpus. In: **2013 International Conference on Electrical Engineering and Software Applications**. IEEE, 2013. p. 1-4.
- VERÍSSIMO, Vinícius et al. A study on the use of sequence-to-sequence neural networks for automatic translation of brazilian portuguese to LIBRAS. In: **Proceedings of the 25th Brazillian Symposium on Multimedia and the Web**. 2019. p. 101-108.
- VOSGERAU, Dilmeire Sant'Anna Ramos; ROMANOWSKI, Joana Paulin. Estudos de revisão: implicações conceituais e metodológicas. **Revista Diálogo Educacional**, v. 14, n. 41, p. 165-189, 2014.
- XAVIER, André Nogueira; NEVES, Sylvia Lia Grespan. Descrição de aspectos morfológicos da Libras. **Revista Sinalizar**, v. 1, n. 2, p. 130-151, 2016.
- XAVIER, André Nogueira; BARBOSA, Plínio Almeida. Diferentes pronúncias em uma língua não sonora? Um estudo da variação na produção de sinais da Libras. **DELTA: Documentação de Estudos em Linguística Teórica e Aplicada**, v. 30, n. 2, p. 371-413, 2014.
- ZUFFO, Darci. O aluno surdo e a aprendizagem de Inglês como língua estrangeira. **O professor PDE e os desafios da escola pública paranaense**, [S.l.], v. 1, p. 1-21, jan. 2010.

**ANEXO A**  
**QUESTIONÁRIO PROPOSTO:**

**PRIMEIRA PARTE:**

1- Qual a sua idade? \_\_\_\_\_

2- Por favor, selecione seu gênero?

- Masculino
- Feminino
- Prefiro não dizer
- Outro (Por favor especifique)

3- Qual seu nível de escolaridade?

- Primária
- Primária não concluída
- Ensino fundamental
- Ensino médio
- Formação profissional / vocacional / técnica
- Superior completo
- Superior incompleto
- Mestrado
- Doutorado

4- Você é:

- surdo
- ouvinte

5- Você usa computador?

- sim
- não

6- Você usa smartphone?

- sim
- não

7 - Você usa aplicativos de tradução para Libras?

- sim
- não

8- Você entende português:

- Perfeito
- Bem
- Mais ou menos
- Pouco
- Não sei português

9- Você entende Libras:

- Perfeito
- Bem
- Mais ou menos
- Pouco
- Não sei Libras

### **SEGUNDA PARTE:**

1- Avalie a tradução da sentença "A manga desta camisa está rasgada".

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

2) Avalie a tradução da sentença "O menino escalou o muro para fugir do cachorro."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

3) Avalie a tradução da sentença "A ação do governo não conseguiu combater às drogas".

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

4) Avalie a tradução da sentença "A manga é uma fruta típica do verão".

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

5) Avalie a tradução da sentença "O diretor do hospital escalou o médico para o plantão."

- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?
- 6) Avalie a tradução da sentença "Os países decidiram dar uma trégua no combate."
- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?
- 7) Avalie a tradução da sentença "Não adianta gritar, ele não ouve."
- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?
- 8) Avalie a tradução da sentença "Não gosto desse seu jeito de agir."
- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?
- 9) Avalie a tradução da sentença "Não posso atendê-lo."
- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?
- 10) Avalie a tradução da sentença "Eu pergunto para você."
- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?
- 11) Avalie a tradução da sentença "Eu disse à vocês."
- a) Avalie a tradução quanto a corretude gramatical?
  - b) Avalie a sua compreensão da tradução gerada?
  - c) Avalie a qualidade da tradução?

12) Avalie a tradução da sentença "Ele deu para mim."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

13) Avalie a tradução da sentença "Qual é o seu nome?"

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

14) Avalie a tradução da sentença "Você é casado?"

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

15) Avalie a tradução da sentença "Você quer água?"

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

16) Avalie a tradução da sentença "Carro bonito!"

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

17) Avalie a tradução da sentença "Conheci muitos surdos!"

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

18) Avalie a tradução da sentença "Saudades, você sumiu!"

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

19) Avalie a tradução da sentença "O homem lê o livro rapidamente."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

20) Avalie a tradução da sentença "A mulher escova os dentes lentamente."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

21) Avalie a tradução da sentença "Este método vai ser melhor que o anterior."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

22) Avalie a tradução da sentença " Ela é pouco inteligente."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

23) Avalie a tradução da sentença "Ela está mais gorda do que eu."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?

24) Avalie a tradução da sentença "O carnaval é muito alegre."

- a) Avalie a tradução quanto a corretude gramatical?
- b) Avalie a sua compreensão da tradução gerada?
- c) Avalie a qualidade da tradução?