

# Uma Estratégia de Tradução Automática de Textos em Língua Portuguesa para Glosa em Língua Brasileira de Sinais

Erickson S. de Oliveira  
Universidade Federal da  
Paraíba  
Cidade Universitária, João  
Pessoa  
Paraíba, Brasil  
erickson.silva@lavid.ufpb.br

Manuella A. C. B. Lima  
Universidade Federal da  
Paraíba  
Cidade Universitária, João  
Pessoa  
Paraíba, Brasil  
manuella.lima@lavid.ufpb.br

Tiago M. U. de Araújo  
Universidade Federal da  
Paraíba  
Cidade Universitária, João  
Pessoa  
Paraíba, Brasil  
tiagomaritan@lavid.ufpb.br

## ABSTRACT

Sign languages (SL) are the natural way of communication among the deaf. Unlike the spoken languages that uses sound in communication, the SL uses a visual channel, i.e., a set of hand, facial and body linguistic elements for presenting the signs. As a result, many deaf people have difficulty to understand and communicate using spoken languages. In addition, the Information and Communication Technologies (ICTs) rarely consider the requirements and needs of deaf. Therefore, to reduce these problems, we developed a computational solution for translating general texts for glosa (a Brazilian Sign Language textual representation) efficiently.

## Categories and Subject Descriptors

D.2.11 [Software Architectures]: Domain-specific architectures; H.5.1 [Multimedia Information Systems]: Animations, Evaluation/methodology; K.4.2 [Social Issues]: Assistive technologies for persons with disabilities

## General Terms

Languages, Performance

## Keywords

Machine translation, Multimedia Platforms, Sign language

## 1. INTRODUÇÃO

As línguas de sinais são a forma natural de comunicação entre os surdos. Diferentemente das línguas orais que utilizam o som na comunicação, as línguas de sinais utilizam um canal visual, isto é, um conjunto de elementos linguísticos manuais, corporais e faciais para articular os sinais [5]. Nesse modelo de comunicação, o emissor constrói uma sentença a partir desses elementos linguísticos e o receptor utiliza o sistema visual para compreender o que está sendo comunicado.

As línguas orais representam para os surdos uma segunda língua, logo, muitos deles têm dificuldade de compreender e se comunicar através de textos em línguas orais, uma vez que essas línguas possuem uma grafia baseada em sons. A maioria dos surdos mesmo passando vários anos na escola não são bem sucedidos em aprender a ler e escrever na língua oral de seu país [17]. No Brasil, por exemplo, segundo o censo demográfico do IBGE de 2000 [7], cerca de 97% dos surdos não concluem o ensino médio. Um outro exemplo

que pode ser citado é um estudo realizado por Wauters [19] com crianças e adolescentes surdos holandeses de 7 a 20 anos de idade, que mostrou que apenas 25% deles possuem uma capacidade de leitura igual ou superior ao de uma criança de 9 anos sem deficiência.

Outro aspecto é que as Tecnologias da Informação e Comunicação (TICs) quando são desenvolvidas, raramente levam em consideração os requisitos e necessidades das pessoas com deficiência [6]. O suporte para línguas de sinais, na TV, por exemplo, é, em geral, limitado a uma janela com um intérprete de língua de sinais, mixado ao vídeo original do programa. Essa solução além de possuir altos custos operacionais para geração e produção dos conteúdos (câmeras, estúdio, equipe, etc.), necessita de intérpretes humanos em tempo integral, o que acaba restringindo seu uso a uma pequena parcela da programação. Essas dificuldades resultam em uma grande barreira para a comunicação com outras pessoas, o acesso a informações, a aquisição de conhecimentos, dentre outros.

Considerando que as línguas de sinais possuem gramáticas próprias, com regras específicas em seus níveis linguísticos, morfológico e sintático, um processo de tradução direta da língua oral para língua de sinais, ou seja, sem utilização de processamento, interpretação ou análise de contexto, não é apropriado. Em virtude disso, um sistema de tradução automática entre os dois tipos de línguas deve levar tanto as mudanças referentes ao canal de comunicação utilizado (isto é, mudança do canal sonoro para o visual), quanto às mudanças gramaticais e estruturais (isto é, mudança de estrutura sequencial das línguas orais para a estrutura paralela das línguas de sinais) [1] [12].

Para minimizar esses problemas, neste trabalho é apresentada uma solução computacional para tradução automática de textos em língua portuguesa para uma representação textual em Língua Brasileira de Sinais (LIBRAS), denominada glosa. Essa solução é responsável por efetuar as adaptações gramaticais dos textos em Português para LIBRAS, e poderá compor soluções de caráter mais geral capazes de traduzir fluxos de voz ou de texto em língua portuguesa para vídeos em LIBRAS [2] [3], especialmente quando intérpretes de LIBRAS não estiverem disponíveis.

Dessa forma, espera-se que os usuários surdos brasileiros possam acessar conteúdos e informações usando sua língua natural de comunicação, a LIBRAS, em diferentes plataformas como, por exemplo, TV Digital, Web e Cinema.

```

Python 2.7.3 (default, Sep 26 2013, 20:03:06)
[GCC 4.6.3] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>> from Aelius import AnotacCorpus, Toqueniza, Extras
>>> h = Extras.carrega("AeliusHunPos")
>>> texto = "Laboratio de Aplicacoes de Video Digital".decode("utf-8")
>>> tokens = Toqueniza.TOK_PORT.tokenize(texto)
>>> list = AnotacCorpus.anoa_sentencas([tokens], h, 'hunpos')
>>> print list
[[('Laboratio', 'N'), ('de', 'P'), ('Aplicacoes', 'NPR'), ('de', 'P'), ('Video', 'NPR'), ('Digital', 'NPR')]]
>>>

```

Figura 1: Uso do Aelius para classificação morfossintática

## 2. TRABALHOS RELACIONADOS

Na literatura científica, existem alguns trabalhos direcionados para as necessidades comunicativas dos surdos [8] [9] [16]. Esses trabalhos oferecem soluções tecnológicas para atividades cotidianas que permitem que pessoas surdas assistam e compreendam televisão, interajam com outras pessoas ou escrevam uma carta. O uso de legendas com informações emotivas em filmes e programas de televisão e o desenvolvimento de jogos para crianças surdas [8] [9] são exemplos desse tipo de solução.

No tocante a geração automática de conteúdos nas diversas línguas de sinais, existem na literatura um conjunto de trabalhos relacionados ao uso de estratégias de tradução automática de fluxos de voz ou texto para animações ou vídeos em línguas de sinais [2] [3] [4] [10] [15] [13] [18] [20].

Veale et al. [18], por exemplo, propõe um sistema de tradução automática multilíngue para traduzir textos em inglês para língua americana de sinais (American Sign Language - ASL), língua irlandesa de sinais (Irish Sign Language - IrishSL) e língua japonesa de sinais (Japanese Sign Language - JSL). Esse sistema é baseado numa arquitetura de quadro negro (blackboard control architecture) [11] e possui um conjunto de agentes que cooperam para gerar os conteúdos traduzidos. Esse trabalho explora e estende alguns conceitos de Inteligência Artificial (IA) para línguas de sinais como, por exemplo, representação do conhecimento, raciocínio metafórico, arquiteturas baseadas em quadro negro [10]. Contudo, nenhum teste ou experimento foi realizado para avaliar a viabilidade e qualidade da solução e das hipóteses elaboradas.

Gallo et al. [4], San-segundo et al. [14] [15] [13] propuseram uma arquitetura para traduzir fluxos de voz em língua espanhola de sinais (LSE) com enfoque em ajudar pessoas surdas no atendimento em serviços públicos como, por exemplo, na renovação da carteira nacional de habilitação. A ideia do sistema é facilitar o diálogo entre surdos e prestadores de serviços públicos neste domínio específico. Além disso, o tempo médio reportado para traduzir cada sentença de voz para LSE foi de cerca de 8 segundos, o que torna a solução inviável em cenários que exigem tradução em tempo real, como, por exemplo, na TV.

## 3. SOLUÇÃO PROPOSTA

Conforme mencionado na Seção 1, a ferramenta proposta de um modo geral é responsável por receber um texto como entrada, e após algumas etapas, como classificação morfossintática e aplicação de regras de tradução definidas por especialistas, converter o texto resultante em uma representação textual em LIBRAS, denominada glosa. Nas Seções 3.1,

3.2 e 3.3 serão descritos os principais processos inerentes ao tradutor.

### 3.1 Classificação

O módulo de classificação baseia-se em duas abordagens: morfológica e sintática. Na classificação morfológica, são identificadas as classes gramaticais de cada palavra que compõe a sentença, enquanto que na classificação sintática é analisada a função que as palavras desempenham dentro da oração, ou seja, a relação sintática entre as palavras.

Para a classificação morfológica, foi utilizado a ferramenta Aelius desenvolvida por Alencar (2010) que é um pacote em Python que, utilizando a biblioteca Natural Language Toolkit (NLTK) e se destina ao pré-processamento de textos, construção de etiquetador morfossintático, anotação de corpora e auxílio de revisão humana de anotação automática.

### 3.2 Aplicação das regras

O resultado é uma lista de tuplas gerada para cada sentença de entrada. Cada tupla é composta pela palavra e sua respectiva classificação morfológica. Na Figura 1 é apresentada uma captura de tela do uso do Aelius para classificação morfológica.

A partir da lista de tuplas gerada é aplicado um processo de classificação sintática que pode produzir dois tipos de saída: uma árvore sintática ou uma árvore sintática nula. A árvore sintática é produzida para sentença reconhecida pela gramática gerativa enquanto que a árvore sintática nula é produzida pela sentença não reconhecida pela gramática. No entanto, a árvore sintática nula nem sempre retrata a inadequação da estrutura formal do Português, podendo esta ser uma limitação da gramática gerativa utilizada pelo parser.

```

<rule name = "P+D-P"><!-- Eliminar preposição especifica -->
  <active>true</active>
  <count>1</count>
  <class>
    <title>P+D-P</title>
    <specific>x</specific>
    <action>remove</action>
  </class>
</rule>

```

Figura 2: Exemplo de uma regra de tradução.

A LIBRAS possui diferentemente do Português Brasileiro não expressa alguns morfemas, tais como: artigos, preposições, dentre outros. Sendo assim, é necessário realizar uma adequação textual para a gramática de Libras. Para isso,

faz-se uso de um conjunto de regras de tradução definida por especialistas de Libras. Essas regras são aplicadas após o processo de classificação e realizam, por exemplo, o tratamento do tempo verbal, de substantivos comuns de dois gêneros, de eliminação dos artigos, de adequação semântica do uso dos advérbios de intensidade e negação e o tratamento dos verbos de ligação, dentre outros. Na Figura 2, é ilustrada uma regra de tradução modelada em linguagem XML para aplicação no tradutor.

### 3.3 Pós-processamento

Um pós-processamento é realizado para refinar a glosa gerada no processo de aplicação das regras. Nessa etapa, números e datas são convertidos para representação por extenso, palavras no plural são convertidas para singular, dentre outros. Ele é importante pois melhora a qualidade da saída, reduzindo a possibilidade de datilologia, que é um processo de de soletração manual, durante a apresentação dos sinais.

## 4. RESULTADOS E DISCUSSÕES

Para validar a estratégia de tradução, a solução proposta foi incorporada a um serviço de tradução automática de Libras, denominado VLibras. Devido a utilização de diferentes linguagens de programação, foi necessário a criação de uma classe em C++ com Python embarcado para que a comunicação pudesse ser realizada.

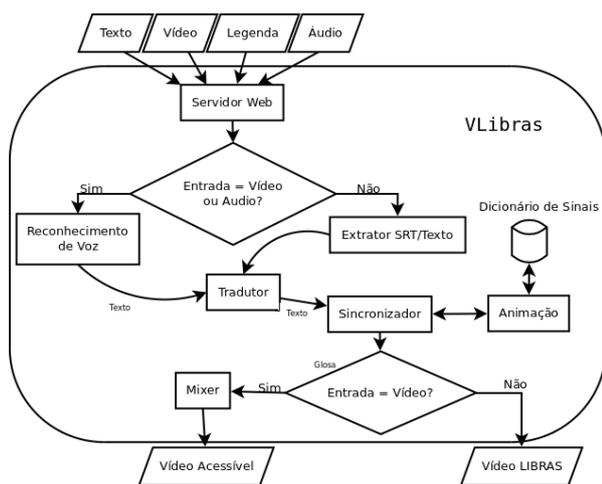


Figura 3: Visão esquemática do VLibras.

O VLibras é um projeto do Laboratório de Aplicações de Vídeo Digital (LAViD) da Universidade Federal da Paraíba (UFPB) que objetiva a geração automática de janelas em LIBRAS a partir do áudio ou legenda do vídeo, tornando-os acessíveis. Os vídeos de LIBRAS são representados usando um avatar-3D (Figura 6).

O VLibras é composto por um conjunto de sub-serviços cujo fluxo dos processos são representados na Figura 3.

Na Figura 3 observa-se que inicialmente, o usuário submete um arquivo (texto, vídeo, legenda, áudio) para o serviço. Esse arquivo é então recebido por um servidor web que o encaminha, de acordo com o seu tipo, para o serviço de reconhecimento de voz ou extrator de legenda visando obter

a informação necessária à tradução, ou seja, o texto. Em seguida, o texto é repassado para o tradutor, no qual realiza os processos citados na Seção 3. O tradutor gera como saída uma glosa, semelhante a visualizada na Figura 4, e repassa para o sincronizador que, junto com o componente de animação, é responsável pela geração do vídeo de LIBRAS. Finalmente, o vídeo de LIBRAS pode ser devolvido ao usuário ou mixado no vídeo original, produzindo um único vídeo acessível e encaminhado ao usuário.

```

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

Figura 4: Exemplo da glosa gerada pela saída do tradutor.

Com o objetivo de investigar o nível de compreensão dos conteúdos gerados pela solução proposta incorporada ao VLibras, foi realizado um teste na Universidade Federal da Paraíba, com 20 usuários surdos brasileiros. Nesse teste, inicialmente os usuários assistiram um vídeo informativo de dois minutos com a tradução automática gerada pelo VLibras.

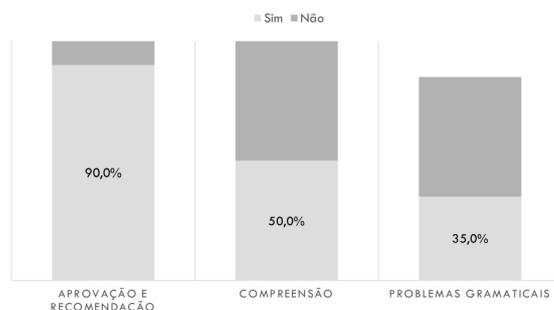


Figura 5: Resultados do teste com surdos na UFPB.

Após os surdos assistirem o vídeo acessível, eles foram convidados a responder um questionário sobre a experiência com a ferramenta acessível. Como resultado, 90% dos usuários indicaram que usariam a ferramenta para traduzir textos ou vídeos de português para LIBRAS, e recomendam a ferramenta para outros surdos. Além disso, 50% dos usuários apontaram ter compreendido bem os conteúdos gerados pela solução. Um dos problemas apontados, no entanto, foi que 35% dos usuários identificaram algum problema gramatical nos conteúdos gerados.

Para endereçar esses problemas, um trabalho resultante foi a definição de uma linguagem formal para gerar regras de descrição que reflitam os aspectos sintático-semântico e pragmático no tradutor automático do Português Brasileiro para LIBRAS.

## 5. CONCLUSÕES

O trabalho apresenta o desenvolvimento de uma solução computacional para traduzir textos de domínio geral para glosa de forma eficiente. Esta solução combina estratégias de simplificação textual, métodos de classificação morfossintática e um conjunto de regras de tradução, definidas por



Figura 6: Avatar 3D do VLibras.

especialistas, para permitir a tradução eficiente de conteúdos do Português Brasileiro para LIBRAS, viabilizando a sua aplicação em diferentes plataformas como TV Digital, Web e Cinema Digital.

De acordo com as análises dos resultados obtidos através do teste com o VLibras e da tradução de diversos conteúdos, a solução proposta possui um bom nível de aceitação pelos usuários surdos. No entanto, é importante ressaltar a necessidade de se investigar novas estratégias para incorporar aspectos sintático-semântico e pragmático e melhorar a qualidade de tradução.

Por fim, apesar de se encontrar ainda em fase de desenvolvimento, é possível perceber a existência de uma contribuição científica, tecnológica e social da proposta apresentada neste trabalho, uma vez que essa solução pode trazer grandes benefícios para aproximadamente 9,7 milhões de surdos brasileiros.

Como proposta de trabalhos futuros pretende-se incorporar aspectos semânticos e pragmáticos na solução proposta, além de modelar um corpus bilingue português-Libras; tentar agregar estratégias de tradução estatística, além de realizar um conjunto de testes mais abrangentes com usuários surdos brasileiros e especialistas em Libras.

## 6. REFERENCES

- [1] L. F. Brito. *Por uma gramática de língua de sinais*. Editora Tempo Brasileiro, Rio de Janeiro, Brasil, 1995.
- [2] T. M. U. de Araújo. *Uma solução para geração automática de trilhas em Língua Brasileira de Sinais em conteúdos multimídia*. Universidade Federal do Rio Grande do Norte, Natal, RN, 2012.
- [3] T. M. U. de Araújo, F. L. S. Ferreira, D. A. N. dos S. Silva, F. H. Lemos, G. P. Neto, D. Omaia, G. L. de Souza Filho, and T. A. Tavares. Automatic generation of brazilian sign language windows for digital tv systems. *Journal of the Brazilian Computer Society*, 19(2):107–125, Junho 2013.
- [4] B. Gallo, R. San-Segundo, J. M. Lucas, R. Barra, L. F. D’Haro, and F. Fernández. Telecomi+d04: Speech into sign language statistical translation system for deaf people. In *IEEE Latin America Transactions*, number 7, pages 400–404, 2009.
- [5] M. C. R. Goes. *Linguagem, Surdez e Educação*. Editora Autores Associados, Campinas, SP, 1996.
- [6] L. Haddon and G. Paul. Technology and the market: Demand, users and innovation. In *ASEAT Conference Proceedings Series*, pages 201–215, 2001.
- [7] IBGE. *Censo demográfico 2000: Características gerais da população*. Editora Tempo Brasileiro, Brasília, Brasil, 2000.
- [8] D. G. Lee, D. I. Fels, and J. P. Udo. Emotive captioning. In *Computers in Entertainment*, volume 5.
- [9] S. Lee, V. Henderson, H. Hamilton, T. Starner, H. Brasher, and S. Hamil. A gesture based american sign language game for deaf children. In *Conference on Human Factors in Computing Systems*, pages 1589–1592, 2005.
- [10] S. Morrissey. *Data-driven machine translation for sign languages*. Dublin City University, Dublin, Irlanda, 2008.
- [11] A. Othman and M. Jemni. Statistical sign language machine translation: from english written text to american sign language gloss. *International Journal of Computer Science Issues*, 8(5):65–73, YEAR = 2011.
- [12] R. M. Quadros. *Educação de Surdos - Aquisição da Linguagem*. Editora Artes Médicas, Porto Alegre, Brasil, 1997.
- [13] R. San-Segundo, J. M. Montero, R. Córdoba, V. Sama, F. Fernández, L. F. D’Haro, V. López-Ludeña, D. Sánchez, and A. García. Design, development and field evaluation of a spanish into sign language translation system. In *Pattern Analysis and Applications*, number 15, pages 203–224, 2011.
- [14] R. Segundo, J. M. Montero, J. Macías-Guarasa, R. Córdoba, J. Ferreiros, and J. M. Pardo. Proposing a speech to gesture translation architecture for spanish deaf people. *Journal of Visual Languages and Computing*, 19(5):523–538, Outubro 2008.
- [15] R. Segundo, J. M. Montero, J. Macías-Guarasa, R. Córdoba, J. Ferreiros, and J. M. Pardo. Speech to sign language translation system for spanish. *Speech Communication*, 50(11-12):1009–1020, Novembro–Dezembro 2008.
- [16] T. Starner, A. Pnetland, and J. Weaver. Real-time american sign language recognition using desk and wearable computer based video. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, number 20, pages 1371–1375, 1998.
- [17] M. R. Stumpf. Língua de sinais: escrita dos surdos na internet. In *Proceedings of V Congresso Ibero-Americano de Informática na Educação*, number 5, pages 1–8, 2000.
- [18] T. Veale, B. Collins, and A. Conway. The challenges of cross-modal translation: English to sign language translation in the zardoz system. *Machine Translation*, 1(13):81–106, 1998.
- [19] L. N. Wauters. *Reading comprehension in deaf children: The impact of the mode of acquisition of word meanings*. Radboud University, Nijmegen, Holanda, 2005.
- [20] L. Zhao, K. Kipper, W. Schuler, C. Vogler, N. Badler, and M. Palmer. Machine translation system from english to american sign language. In *Proceedings of the 4th Conference of the Association for Machine Translation in the Americas*, number 4, pages 54–67, 2000.