

# SISTEMA DE TRADUÇÃO DE VOZ PARA DATILOLOGIA EM LIBRAS EM TEMPO REAL

GAZOLLA, Victor Gabriel Modesto<sup>1</sup>; BAIA, Joás Weslei<sup>2</sup>; PEREIRA, Ana Amélia de Souza<sup>2</sup>; CARMO, Amanda Juliana<sup>2</sup>



<sup>1</sup> Graduando em Ciência da Computação - UNIFAGOC

<sup>2</sup> Docente do curso de Ciência da Computação - UNIFAGOC

victormgazolla@gmail.com  
joas.baia@unifagoc.edu.br  
ana.amelia@unifagoc.edu.br

## RESUMO

*A comunicação é uma parte essencial da vida humana, portanto pessoas surdas ou com deficiência auditiva enfrentam problemas diários na comunicação, ocasionando falta de oportunidades ou dificuldade em interagir. O objetivo deste trabalho foi proporcionar uma comunicação eficaz por meio de um software de tradução de voz para datilogia em Libras. Para o desenvolvimento do trabalho, foram utilizadas a plataforma Figma voltada para o design, a linguagem de programação Python, bibliotecas como o React JS e a biblioteca SpeechRecognition para captura e processamento de áudio.*

**Palavras-chave:** Comunicação. Libras. Tradução de voz. Inclusão. Processamento de linguagem natural. Acessibilidade.

## INTRODUÇÃO

A comunicação é uma parte indispensável da vida humana; é a expressão de pensamentos, emoções e o compartilhamento de informações vitais. No entanto, os surdos e as pessoas com deficiência auditiva enfrentam desafios diários nesse aspecto. Para essas pessoas, a Língua de Sinais, como a Língua Brasileira de Sinais (Libras), é a principal forma de se comunicar.

A Libras é uma língua visual-motora que faz a utilização de gesticulações, expressões faciais e movimentos corporais, para fins de comunicação. A Libras possui sua própria gramática e estrutura linguística, sendo reconhecida oficialmente no Brasil, de acordo com o Art. 1º da Lei nº 10.436, de 24 de abril de 2002 (Brasil, 2002).

No âmbito da Libras, o alfabeto manual, também conhecido como datilogia, desempenha um papel crucial. Esse recurso utiliza gestos específicos com as mãos para representar cada letra do alfabeto escrito em português. Essa técnica é essencial para soletrar palavras, nomes próprios, endereços e outras informações que não possuem um sinal específico na língua de sinais (Foggetti, 2023).

É importante destacar que, embora o alfabeto manual seja valioso, a Libras vai além da datilogia, contando com mais de 14 mil sinais e cinco parâmetros fundamentais: configuração da mão, ponto ou local de articulação, movimento, orientação da palma das mãos e expressão facial ou corporal (Foggetti, 2023). Esses elementos são essenciais para uma comunicação eficaz e rica em nuances na Língua Brasileira de Sinais.

Por meio da Libras, os surdos podem expressar suas ideias, compartilhar informações complexas e estabelecer conexões sociais com outros falantes da língua. Entretanto, enfrentar barreiras de comunicação com pessoas que não conhecem a

Libras é uma realidade comum para os surdos. Essas barreiras podem resultar em exclusão, dificuldades no acesso a serviços essenciais e limitações nas oportunidades educacionais e profissionais. É aqui que surge a necessidade de um sistema que traduza a voz para Libras.

As pessoas surdas ou com deficiência auditiva enfrentam dificuldades significativas na comunicação cotidiana, principalmente em situações em que não há intérpretes de Libras disponíveis. Encontros profissionais, consultas médicas, atendimentos bancários e até mesmo interações sociais simples podem se tornar desafiadoras e frustrantes. Essa falta de acesso a uma comunicação efetiva impede que essas pessoas sejam incluídas plenamente na sociedade e tenham acesso a informações e serviços essenciais.

Além disso, a necessidade de um sistema de tradução de voz para Libras impacta as oportunidades de emprego para pessoas surdas. Durante entrevistas de emprego, a comunicação pode ser um desafio, prejudicando sua capacidade de serem compreendidas e avaliadas corretamente pelos empregadores. No ambiente de trabalho, a dificuldade de comunicação pode limitar o desenvolvimento profissional dessas pessoas.

Diante dessas situações, um sistema de tradução de voz para Libras em tempo real surge como uma solução eficiente. Esse sistema permite que as pessoas com deficiência auditiva tenham acesso imediato a informações cruciais, além de abrir um mundo de conhecimento e oportunidades. Ao eliminar as barreiras de comunicação, o sistema tornaria possível a inclusão social e no mercado de trabalho, tornando as pessoas surdas ou com deficiência auditiva capazes de participarem plenamente da sociedade e aproveitarem todas as oportunidades oferecidas por ela.

Assim, o objeto deste trabalho foi desenvolver um aplicativo de software para traduzir a comunicação oral pela comunicação por sinais do alfabeto em Libras.

## REFERENCIAL TEÓRICO

### Processamento de Linguagem Natural

De acordo com Indurkhyá e Damerau (2010 *apud* Barbosa *et al.*, 2017), o Processamento de Linguagem Natural (PLN) é um campo da ciência da computação que visa compreender e interpretar textos escritos em linguagem natural, tendo o objetivo de extrair informações e significados abrangentes desses textos.

O PLN emergiu como uma resposta crescente à necessidade de se estabelecer uma interação automatizada e uma compreensão mútua entre seres humanos e computadores (Santos, 2015).

Segundo Liu e Zhang (2013 *apud* Barbosa *et al.*, 2017), o PLN vai além do tratamento de texto apenas como uma sequência de caracteres, ou seja, deve também ser levada em consideração a estrutura hierárquica da linguagem. Por meio da análise do significado da linguagem, os sistemas de PLN possuem uma ampla aplicação em diversas áreas, como correção da gramática, conversão de fala para texto, tradução automática entre idiomas e análise de sentimentos.

Compreender a linguagem natural envolve mais do que simplesmente palavras; é necessário compreender os conceitos subjacentes e a maneira como eles se conectam para formar significados. Embora humanos apresentem facilidade no aprendizado da linguagem natural, a ambiguidade inerente a ela torna o

Processamento de Linguagem Natural (PLN) um desafio complexo para computadores (Barbosa *et al.*, 2017).

A fim de que um sistema possa compreender e processar a linguagem natural, é essencial identificar e extrair características (features). As features são elementos presentes nos dados que desempenham um papel crucial no reconhecimento de padrões pelo algoritmo. É a partir dessas características que se determina quais aspectos do texto serão utilizados como entrada para o classificador (Filho, 2022).

## React

React é uma biblioteca JavaScript utilizada para criar interfaces de usuário. Com essa biblioteca é possível desenvolver componentes e combiná-los para formar telas, páginas e aplicativos (React, 2023). É utilizado para o desenvolvimento de sites. Além disso, a biblioteca React permite adicionar interatividade aos componentes, possibilitando a atualização da interface em resposta a ações do usuário (React, 2023).

## MÉTODO DE DESENVOLVIMENTO

Para desenvolver o aplicativo de software que traduz a comunicação oral para a comunicação por sinais do alfabeto em Libras, foram utilizadas as seguintes tecnologias:

- Prototipagem com Figma;
- Arquitetura cliente-servidor;
- SpeechRecognition para capturar e transformar a voz em texto;
- React Three Fiber.

## DESENVOLVIMENTO

A partir do método de desenvolvimento proposto para atingir nosso objetivo de pesquisa, foram executadas as atividades descritas a seguir.

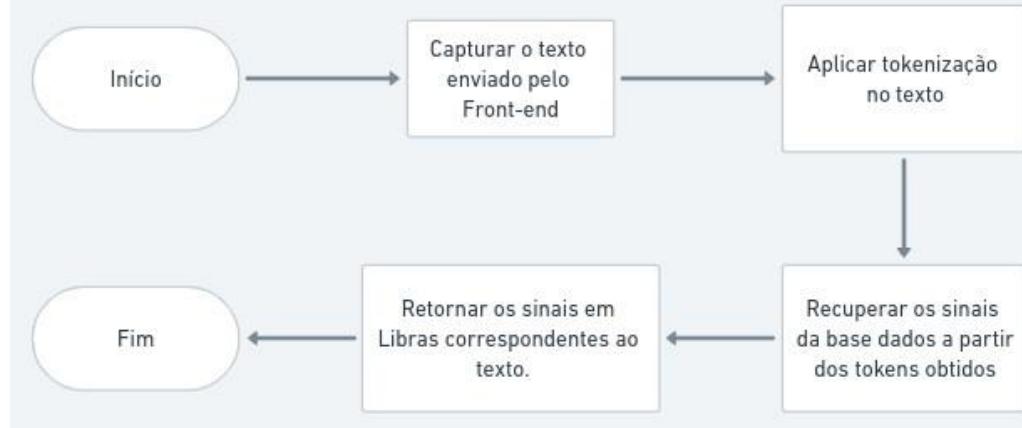
### Design

Primeiro utilizando o Figma, foram desenvolvidos protótipos interativos que simulam a navegação e a funcionalidade do site. Esses protótipos foram iterados e refinados durante o processo de desenvolvimento, permitindo visualizar o resultado esperado do projeto de forma visual.

### Tradução de texto para datilologia em Libras

A tradução de texto para datilologia Libras foi implementada com a linguagem Python, utilizando o WebSocket como meio de comunicação com o Front-end. A sintaxe de fácil entendimento da linguagem facilitou o desenvolvimento de um Back-end robusto.

**Figura 1:** Representação do fluxo de tratamento das informações no Back-end



Fonte: dados da pesquisa.

No processo de tradução do alfabeto para o alfabeto em Libras, cada letra foi incorporada aderindo à estrutura JSON. Cada caractere foi configurado como um objeto contendo o conjunto de valores essenciais para a manipulação do modelo 3D.

**Figura 2:** Letra C em Libras representada em um objeto JSON

```

1  {
2    "letter_name": "c",
3    "duration": 500,
4    "action": [
5      {
6        "name": "upper_armR",
7        "rotationX": 0.5,
8        "rotationY": 0
9      },
10     {
11       "name": "handR",
12       "rotationX": 0.7,
13       "rotationY": 0
14     },
15     {
16       "name": "forearmR",
17       "rotationX": 1.8,
18       "rotationY": 0
19     }
20   ]
21 }
  
```

Fonte: dados da pesquisa.

Após a definição da estrutura de cada letra seguindo o formato JSON, foi implementado o processo de tokenização que é aplicado no texto proveniente do Front-end, visando determinar a letra a ser utilizada. No contexto do Processamento de Linguagem Natural (PLN), a tokenização consiste em dividir o texto em partes menores chamadas de tokens. Cada token pode representar uma palavra, uma parte de uma palavra ou até mesmo um caractere, dependendo da abordagem adotada. Esse procedimento facilita a análise e compreensão do texto, permitindo a identificação eficiente das letras relevantes para a tradução.

**Figura 3:** Exemplo de código de como o processo de tokenização é aplicado

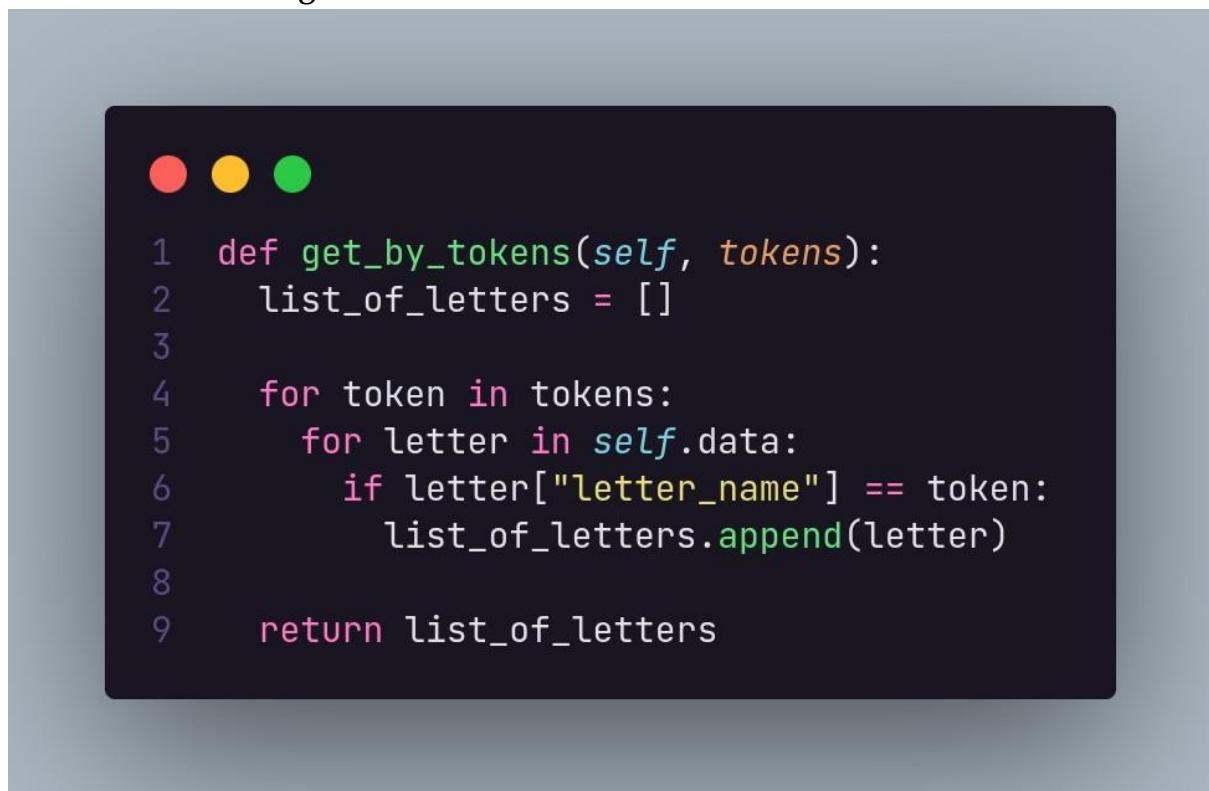


```
 1 import re
 2
 3 class TokenizeTextService:
 4     def tokenize_text(self, text):
 5         text = text.lower()
 6         text = re.sub(r'[^a-zA-Z\s]', '', text)
 7
 8         tokens = list(text)
 9
10     return tokens
11
```

Fonte: dados da pesquisa.

Após a obtenção da lista de tokens, o sistema realiza uma busca na base de dados local para encontrar as letras correspondentes. Durante esse processo, cada token é comparado com as entradas disponíveis na base de dados, resultando em uma lista de letras associadas aos tokens.

**Figura 4:** Busca das letras na base de dados local



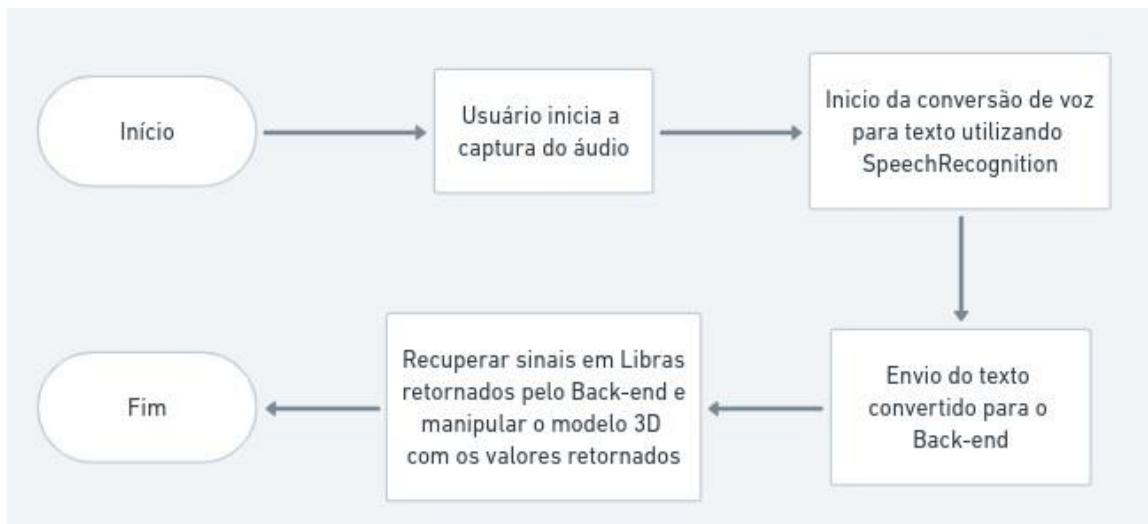
```
1 def get_by_tokens(self, tokens):
2     list_of_letters = []
3
4     for token in tokens:
5         for letter in self.data:
6             if letter["letter_name"] == token:
7                 list_of_letters.append(letter)
8
9     return list_of_letters
```

Fonte: dados da pesquisa.

### Tradução de voz para texto

No desenvolvimento do Front-end com a tecnologia React JS para o desenvolvimento do site, foram utilizados recursos nativos, o que trará benefícios significativos para a manutenção do sistema. Essa abordagem evita a necessidade de utilizar dependências externas, simplificando o processo de atualização e correção. Ao aproveitar recursos nativos, teremos maior controle sobre o código, evitando possíveis conflitos entre bibliotecas diferentes e aumentando a eficiência na gestão do sistema. Essa escolha permitirá uma manutenção ágil e reduzirá a complexidade geral do projeto.

**Figura 5:** Representação do tratamento das informações no Front-end



Fonte: dados da pesquisa.

No contexto da captura e processamento de áudio, uma decisão estratégica foi tomada em relação à utilização da biblioteca SpeechRecognition na linguagem Python. Em vez desta dependência, optou-se por explorar o recurso SpeechRecognition nativo da web, cuja função principal é traduzir a voz para texto. Essa escolha foi motivada por diversos fatores que impactam diretamente na eficiência, na flexibilidade e na manutenção do sistema.

O SpeechRecognition nativo da web oferece uma alternativa robusta, eficaz e sem a necessidade de incorporar bibliotecas externas. Ao adotar essa abordagem, o sistema se beneficia das vantagens inerentes às tecnologias web modernas, garantindo um desempenho otimizado e uma maior independência de bibliotecas adicionais.

A decisão de não utilizar a dependência SpeechRecognition em Python também está alinhada com a escolha de utilizar React JS no Front-end. Ao optar por recursos nativos do React, buscamos simplificar o processo de desenvolvimento, garantindo uma maior agilidade na manutenção e reduzindo a complexidade geral do projeto. Essa abordagem visa evitar conflitos entre diferentes bibliotecas e proporcionar um controle mais preciso sobre o código.

Além disso, a decisão de utilizar o SpeechRecognition nativo da web está em consonância com a busca por eficiência e minimização de dependências externas. A integração de recursos nativos simplifica a gestão do sistema, garantindo um desenvolvimento mais coeso e evitando possíveis obstáculos que poderiam surgir com a adoção de bibliotecas externas.

Em resumo, a opção pelo SpeechRecognition nativo da web, com a função de traduzir a voz para texto, foi baseada em considerações de eficiência, flexibilidade e manutenção do sistema, alinhando-se com a abordagem adotada no projeto.

**Figura 6:** Processo de tradução de voz para texto  
Fonte: dados da pesquisa.



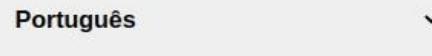
```
1 const startCapture = () => {
2   const recognition = new window.webkitSpeechRecognition();
3   recognition.continuous = true;
4   recognition.interimResults = false;
5   recognition.lang = language;
6
7   recognition.onresult = (event: any) => {
8     for (let i = event.resultIndex; i < event.results.length; i++) {
9       const transcript = event.results[i][0].transcript;
10      if (event.results[i].isFinal) {
11        socket.send(transcript);
12        setTranscription(transcript);
13      }
14    }
15  };
16
17  recognition.start();
18};
```

### Manipulação do modelo 3D

A renderização do modelo 3D foi possibilitada através da utilização da biblioteca React Three Fiber. Os parâmetros essenciais para a manipulação do modelo são provenientes do Back-end. A cada interação do usuário, o áudio correspondente é convertido em texto, o qual é posteriormente encaminhado para o Back-end. Nessa etapa, ocorre a tradução para o alfabeto em Libras. As traduções resultantes, ao serem retornadas pelo Back-end, são então aplicadas ao modelo 3D, ajustando-o de acordo com o conjunto de valores obtidos.

**Figura 7:** Modelo 3D realizando o movimento da letra C

Português



C

Iniciar captura

Fonte: dados da pesquisa.

## CONSIDERAÇÕES FINAIS

Espera-se que o resultado deste projeto seja capaz de promover uma inclusão mais abrangente e igualdade de oportunidades para pessoas surdas ou com deficiência auditiva. A comunicação será mais acessível, permitindo que essas pessoas se envolvam plenamente na sociedade, tenham acesso a informações cruciais, participem ativamente de atividades profissionais e sociais, e desfrutem de uma melhor qualidade de vida. O investimento nesse sistema é de extrema importância para construir uma sociedade mais inclusiva, em que todos tenham a oportunidade de se expressar e ser compreendidos, independentemente de suas habilidades auditivas.

## REFERÊNCIAS

- BARBOSA, Jardeson Leandro Nascimento; VIEIRA, João Paulo Albuquerque; SANTOS, Roney Lira de Sales; JUNIOR, Gilvan Veras Magalhães; MUNIZ, Mariana dos Santos; MOURA, Raimundo Santos. Introdução ao processamento de linguagem natural usando Python. III Escola Regional de Informática do Piauí. **Livro Anais - Artigos e Minicursos**, v. 1, n. 1, p. 336-360, jun. 2017. Disponível em: [https://www.facom.ufu.br/~wendelmelo/terceiros/tutorial\\_nltk.pdf](https://www.facom.ufu.br/~wendelmelo/terceiros/tutorial_nltk.pdf). Acesso em: 6 jul. 2023.
- BRASIL. **Lei nº 10.436, de 24 de abril de 2002**. Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/2002/l10436.htm](https://www.planalto.gov.br/ccivil_03/leis/2002/l10436.htm). Acesso em: 26 jun. 2023.
- FIGMA. **What is Figma?** Disponível em: <https://help.figma.com/hc/en-us/articles/14563969806359-What-is-Figma->. Acesso em: 1 jul. 2023.
- FILHO, Nelson Silva de Freitas. **Aprendizado de máquina e processamento de linguagem natural aplicado na predição de marcas a partir de dados de marketplaces**. Universidade Federal do Rio Grande do Norte. Disponível em: [https://repositorio.ufrn.br/bitstream/123456789/49080/4/AprendizadoDeM%c3%a1quina\\_Freitas%20Filho\\_2022.pdf](https://repositorio.ufrn.br/bitstream/123456789/49080/4/AprendizadoDeM%c3%a1quina_Freitas%20Filho_2022.pdf). Acesso em: 7 jul. 2023.
- FOGGETTI, Fernanda. **Alfabeto em Libras**: o que é, para que serve e importância. Disponível em: <https://www.handtalk.me/br/blog/alfabeto-em-libras-o-que-e-para-que-servir-e-importancia/#:text=O%20alfabeto%20manual%20em%20Libras,portugu%C3%AAs%20e%20outras%20l%C3%ADnguas%20faladas>. Acesso em: 6 dez. 2023.
- PYTHON. **Doc**. Disponível em: <https://www.python.org/doc>. Acesso em: 1 jul. 2023.
- REACT. Disponível em: <https://docs.pmnd.rs/react-three-fiber/getting-started/introduction>. Acesso em: 19 nov. 2023.
- SANTOS, Cedric Michael. **Classificação de documentos com processamento de linguagem natural**. Disponível em: <https://comum.rcaap.pt/bitstream/10400.26/15293/1/Cedric-Michael-Santos.pdf>. Acesso em: 6 jul. 2023.
- SPEECH RECOGNITION. Disponível em: <https://developer.mozilla.org/en-US/docs/Web/API/SpeechRecognition>. Acesso em: 18 nov. 2023.