

## ACESSIBILIDADE À INTELIGÊNCIA ARTIFICIAL PARA O APRENDIZADO DE LÍNGUAS ESTRANGEIRAS UTILIZANDO SISTEMAS ELETRÔNICOS EMBARCADOS.

ANTUNES, F. V.<sup>1</sup>; HARMUCH, G, D.<sup>1</sup>

<sup>1</sup>Grupo PET Elétrica, UFPR, Campus Centro Politécnico, [pet.eletrica@ufpr.br](mailto:pet.eletrica@ufpr.br); [flavia.victoria@ufpr.br](mailto:flavia.victoria@ufpr.br); [gustavoharmuch@ufpr.br](mailto:gustavoharmuch@ufpr.br). <sup>3</sup>Tutor do Grupo PET Elétrica, Horácio Tertuliano dos Santos Filho.

**RESUMO:** O presente trabalho propõe o desenvolvimento de um sistema de avaliação automática de pronúncia e coesão em conversações em outros idiomas, utilizando dispositivos microcontroladores e técnicas de processamento digital de sinais. O fluxo de funcionamento do sistema se inicia a partir de um microfone genérico, responsável por captar a fala do usuário e enviar para um microcontrolador, que por sua vez converte o áudio em texto, utilizando inteligência artificial. Em seguida, um comparador fonético avalia a similaridade entre a fala reconhecida e o texto esperado, utilizando métricas comparativas, como a similaridade do cosseno entre vetores gerados por *embeddings* de linguagem, assim como métricas fonéticas são analisadas por seus respectivos algoritmos. O resultado disso, é um índice de pronúncia entre 0 e 100%, exibido por LEDs e mensagens em display LCD. O sistema inclui ainda uma etapa de síntese de voz que transforma texto em fala, permitindo interação contínua do usuário por meio de um ciclo de escuta, avaliação e resposta.

**Palavras-chave:** sistemas embarcados; inteligência artificial; língua-estrangeira; eletrônica.

## ACCESSIBILITY TO ARTIFICIAL INTELLIGENCE FOR FOREIGN LANGUAGE LEARNING USING EMBEDDED ELECTRONIC SYSTEMS

**ABSTRACT:** This work proposes the development of a system for automatic evaluation of pronunciation and cohesion in conversations in foreign languages, using microcontroller devices and digital signal processing techniques. The system's workflow begins with a generic microphone, responsible for capturing the user's speech and sending it to a microcontroller, which in turn converts the audio into text using artificial intelligence. Next, a phonetic comparator evaluates the similarity between the recognized speech and the expected text, using comparative metrics such as cosine similarity between vectors generated by language embeddings, as well as phonetic metrics analyzed by their respective algorithms. The result is a pronunciation index ranging from 0 to 100%, displayed via LEDs and messages on an LCD screen. The system also includes a speech synthesis stage that transforms text into speech, enabling continuous user interaction through a cycle of listening, evaluation, and response.

**Keywords:** embedded systems; artificial intelligence, language learning; electronics.

## Introdução

O avanço das tecnologias embarcadas e da inteligência artificial vem cada vez mais permitindo novas estratégias para o aprendizado de idiomas e conseqüentemente, a melhoria da pronúncia. Segundo Gileno e Cruz (2020, p. 234-237), “torna-se imprescindível ampliar as possibilidades de letramentos para multiletramentos em contextos onde as tecnologias [...] podem facilitar o ensino e a aprendizagem de uma língua estrangeira”. Instrumentos que combinam microcontroladores, processamento de áudio e aprendizado de máquina possuem a capacidade de tornar a experiência de aprendizado mais prática e benéfica.

Este trabalho tem como objetivo criar um sistema acessível e prático de estudo de idiomas, que possa capturar e avaliar a fala do usuário, oferecendo feedback em tempo real sobre a pronúncia, usando hardware de custo reduzido e técnicas de processamento digital de sinais e reconhecimento de fala.

## Método

O desenvolvimento do sistema foi dividido em quatro etapas principais: captura e aquisição de sinal, pré-processamento, análise fonética/comparação de similaridade, e interação com o usuário

1. Captura e Aquisição de Sinal:  
O sistema inicia com um microfone qualquer, como por exemplo um KY038 (módulo com opções digitais e analógicas de saída já integrados), responsável por converter a fala do usuário em um sinal. Este sinal é enviado ao microcontrolador (ESP32, STM32, ou similar) que realiza a digitalização por meio de seu conversor analógico-digital (ADC).
2. Pré-processamento de Áudio (DSP):  
Após a captura, o sinal passa por um processo de filtragem passa-baixo para remoção de ruídos agudos e normalização RMS, que ajusta o nível de intensidade da fala garantindo que o sinal de entrada seja limpo e padronizado para as próximas fases. Bibliotecas como *Librosa* e *SciPy Signal* em *Python* são utilizadas para realizar essas operações. Nessa etapa também pode ser aplicada uma equalização de volume, compensando variações naturais de amplitude entre diferentes gravações.
3. Análise Fonética e Comparação de Similaridade:  
A fala pré-processada é submetida a um módulo de reconhecimento de voz, que converte o áudio em texto, utilizando inteligência artificial. Em seguida, o sistema realiza uma comparação fonética e semântica entre o texto reconhecido e o texto esperado. Para isso, são utilizadas três métricas principais:

## INTELIGÊNCIA ARTIFICIAL E DIREITOS HUMANOS: DESAFIOS ÉTICOS PARA O SÉCULO XX

- a. Distância de Levenshtein: mede o número de operações (inserções, deleções e substituições) necessárias para transformar uma palavra em outra, fornecendo uma medida direta de similaridade textual.
- b. Similaridade do cosseno com *embeddings*: converte as sentenças em vetores numéricos utilizando modelos pré-treinados de IA, como os *embeddings* da *OpenAI*, e calcula a similaridade entre eles, identificando a proximidade semântica entre as expressões.
- c. Algoritmo *Metaphone*: realiza uma análise fonética, codificando palavras conforme sua pronúncia, o que permite avaliar semelhanças sonoras mesmo quando há pequenas variações de fonética.

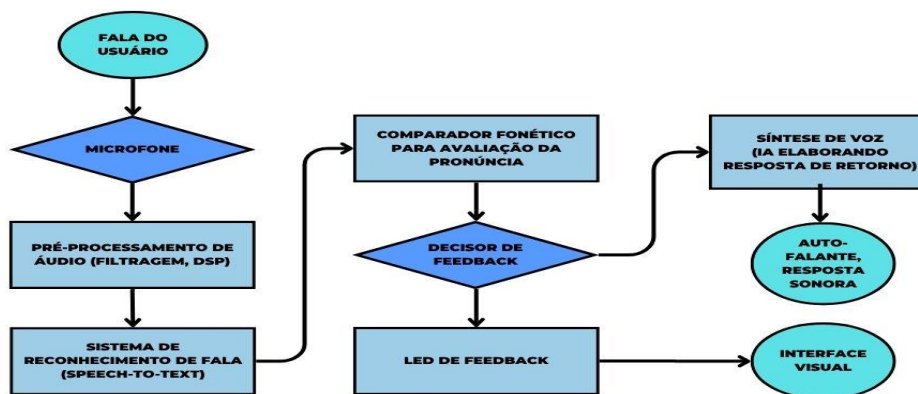
A combinação dessas métricas fornecerá um índice de pronúncia, que varia de 0 a 100%, indicando o grau de semelhança entre a fala do usuário e a pronúncia esperada.

#### 4. Feedback Visual e Auditivo:

Após a análise, o sistema retorna o resultado ao usuário de forma visual e sonora. LEDs de diferentes cores indicam o nível de acerto (por exemplo, verde para alta similaridade e vermelho para baixa). Um display LCD pode exibir mensagens de incentivo. Por fim, a síntese de voz (TTS), implementada com bibliotecas como *gTTS*, gera uma resposta falada, concluindo o ciclo de interação. Esse processo se repete continuamente em um loop de conversação composto pelas etapas: captura de fala; processamento; avaliação e resposta.

A metodologia empregada permite um aprendizado eficiente, corrigindo os possíveis erros do usuário de forma rápida e eficiente.

Figura 1 – Fluxograma do funcionamento



Fonte: Autoria própria, 2025.

## Resultados e Discussão

Espera-se do projeto a capacidade de avaliar a pronúncia em tempo real de forma funcional, fornecendo feedback visual e sonoro ao usuário para autoavaliação e acompanhamento do progresso no aprendizado. Além disso, prevê-se que, os microcontroladores apresentem desempenho adequado para o processamento de sinais de áudio e integração com algoritmos de IA, demonstrando a viabilidade de soluções acessíveis para o aprendizado autônomo de idiomas.

Os resultados esperados indicam que o sistema poderá contribuir para a prática oral e o aperfeiçoamento da pronúncia, aproximando o aprendizado de uma experiência interativa.

Para futuros desenvolvimentos, analisou-se a necessidade de discutir sobre as limitações relacionadas a ruídos no ambiente e capacidade de processamento, visando aprimoramentos na precisão.

## Conclusões

Por fim, conclui-se que, o desenvolvimento do sistema proposto demonstra o potencial de integração entre inteligência artificial e eletrônica embarcada aplicada ao aprendizado de línguas estrangeiras. Dessa forma, vemos que soluções baseadas em microcontroladores e processamento digital de sinais podem ampliar o acesso a métodos tecnológicos de ensino, contribuindo para uma aprendizagem mais prática, contínua e acessível.



## INTELIGÊNCIA ARTIFICIAL E DIREITOS HUMANOS: DESAFIOS ÉTICOS PARA O SÉCULO XX

### Agradecimentos

Como membros de um grupo, agradecemos nossos colegas e tutores pelo apoio. Estendemos também nosso reconhecimento ao MEC e do Departamento de Engenharia Elétrica da UFPR, pelo apoio financeiro e estrutural no decorrer do projeto.

### Referências

DOS SANTOS, Milena Meira Ramos. Ensino de língua estrangeira: os métodos. Revista EntreLínguas, Araraquara, v. 6, n. 2, p. 249-265, jul./dez. 2020. Disponível em: <https://periodicos.fclar.unesp.br/entrelinguas/article/view/14272/9757>. Acesso em: 7 out. 2025.

CABRAL, Raquel da Silva; CAMARA NETO, Vilar Fiuza da. Arquiteturas DSP e aplicações. Universidade Federal de Minas Gerais, Departamento de Ciência da Computação, 2006. Disponível em: <https://www.verlab.dcc.ufmg.br/old/media/cursos/arquitetura/2006-2/grupo1/artigoarquiteturaraquelvilar.pdf>. Acesso em: 2 out. 2025.

SILVA, André Luiz Viana da. Aplicação de filtros digitais e efeitos sonoros em sinais de áudio, utilizando NI SPEEDY 33 e módulo LabView DSP. 2021. 53 f. Monografia (Graduação em Engenharia de Controle e Automação) - Escola de Minas, Universidade Federal de Ouro Preto, Ouro Preto, 2021.