



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76
Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA

XXIX SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2025

Inclusão através da Voz: Ampliando o Acesso aos Jogos Digitais com o Reconhecimento de Voz na Plataforma AsKME

João Victor Alves Cerqueira¹; Victor Travassos Sarinho²

1. João Victor Alves Cerqueira, PIBIC/FAPESB, ENGENHARIA DE COMPUTAÇÃO, e-mail: jvacerqueira@ecompu.uefs.br

2. Victor Travassos Sarinho, DEPARTAMENTO DE CIÊNCIAS EXATAS, e-mail: vsarinho@uefs.br

PALAVRAS-CHAVE: Inclusão Digital, Reconhecimento de Voz, Tecnologia Assistiva

INTRODUÇÃO

A acessibilidade em jogos digitais é um desafio crescente no contexto da educação e da tecnologia, especialmente para pessoas com deficiência visual ou motora. A ausência de recursos inclusivos limita o acesso dessas pessoas a ferramentas de aprendizado e entretenimento digital. O projeto “Inclusão através da Voz: Ampliando o Acesso aos Jogos Digitais com o Reconhecimento de Voz na Plataforma AsKME” teve como objetivo desenvolver e integrar funcionalidades de leitura de tela e reconhecimento de voz à plataforma AsKME, permitindo que usuários possam interagir com jogos educativos de forma autônoma. A proposta justifica-se pela necessidade de promover equidade no acesso ao conhecimento, utilizando tecnologias assistivas como meio de inclusão digital. A pesquisa baseou-se na combinação de tecnologias nativas de navegadores, como Web Speech API e SpeechSynthesis, com uma arquitetura híbrida baseada em FastAPI e Google Cloud Speech-to-Text, ampliando o suporte a diferentes ambientes e dispositivos.

MATERIAL E MÉTODOS OU METODOLOGIA (ou equivalente)

O desenvolvimento do sistema de acessibilidade foi conduzido por meio de uma abordagem incremental, estruturada em ciclos de implementação, teste e refinamento contínuo. A metodologia combinou análise de código, integração tecnológica e experimentação prática, com foco em garantir compatibilidade, desempenho e usabilidade dentro da plataforma AsKME, destinada a jogos educativos interativos.

A primeira etapa consistiu em compreender a estrutura interna do código da AsKME, mapeando variáveis e funções responsáveis pela entrada e saída de dados. Essa análise foi essencial para permitir a substituição das entradas de teclado por comandos de voz. Para isso, os valores recebidos pela fala foram normalizados e convertidos para o mesmo formato aceito pelo sistema, simulando o comportamento esperado quando o jogador digita uma opção no teclado. Assim, o comando “opção dois”, por exemplo, era interpretado, convertido para o número “2” e processado automaticamente pelo jogo.

Em paralelo, foi implementado o recurso de leitura de tela, utilizando a API `SpeechSynthesis`, responsável por transformar textos exibidos na interface em áudio narrado em tempo real. A cada nova mensagem apresentada ao usuário — como instruções, `feedbacks` ou menus — o sistema gravava automaticamente uma reprodução sonora, garantindo que usuários com deficiência visual tivessem acesso completo às informações da tela sem depender de visão.

A etapa seguinte envolveu a integração do reconhecimento de voz nativo via `Web Speech API (webkitSpeechRecognition)`, configurado para operar em português brasileiro e no modo contínuo. Esse modo permite que o microfone permaneça ativo e capture a fala do usuário em tempo real, interpretando palavras e frases à medida que são pronunciadas. Contudo, observou-se que nem todos os navegadores oferecem suporte a essa API, o que motivou o desenvolvimento de um módulo complementar em Python, baseado no framework `FastAPI`, com uso da biblioteca `SpeechRecognition` integrada ao serviço `Google Cloud Speech-to-Text`. Essa camada de `fallback` garantiu a compatibilidade da solução em qualquer ambiente, mesmo fora do ecossistema Chrome.

RESULTADOS E/OU DISCUSSÃO (ou Análise e discussão dos resultados)

O desenvolvimento do sistema de acessibilidade foi conduzido de forma incremental e experimental, priorizando a integração progressiva de recursos e a avaliação prática de cada etapa. A metodologia combinou análise de código, integração tecnológica e testes de usabilidade, garantindo compatibilidade com a plataforma `AsKME` e a manutenção de seu funcionamento original.

Inicialmente, foi realizada uma análise estrutural do código da `AsKME` para mapear as variáveis e funções responsáveis pelas interações com o usuário. Esse mapeamento permitiu que as entradas de teclado fossem estendidas para também receber comandos de voz, convertendo a fala em texto e processando-a de acordo com as regras já existentes no sistema. Assim, o jogador pôde realizar ações como escolher opções ou iniciar partidas apenas com a voz.

Em seguida, implementou-se o recurso de leitura de tela utilizando a API `SpeechSynthesis`, que converte textos exibidos na interface em áudio narrado. Essa funcionalidade garante que mensagens, instruções e menus sejam lidos automaticamente sempre que atualizados, permitindo que usuários com deficiência visual acompanhem o jogo em tempo real.

A principal camada de interação foi desenvolvida com o reconhecimento de voz, integrando a `Web Speech API (webkitSpeechRecognition)`, configurada para o idioma português brasileiro e operação contínua. Para ampliar a compatibilidade, foi criado um módulo backend em Python com `FastAPI`, utilizando a biblioteca `SpeechRecognition` e o serviço `Google Cloud Speech-to-Text` como alternativa para navegadores sem suporte nativo. Essa abordagem híbrida assegurou o funcionamento universal da ferramenta em diferentes dispositivos e ambientes.

Durante a implementação, também foram desenvolvidos elementos de interface acessível, como botões com feedback visuais e efeitos sonoros que indicam o estado das funções de leitura e transcrição de voz. O ícone de microfone verde, acompanhado de um feedback auditivo, foi adicionado para sinalizar quando o sistema está pronto para receber comandos, proporcionando ao usuário maior clareza e segurança durante a interação.

Além disso, foi adotado um mecanismo de escuta contínua, que mantém o sistema atento à voz do usuário sem exigir reativações manuais. Esse mecanismo permite captar comandos em tempo real, garantindo uma experiência fluida e natural, sem comprometer o desempenho da aplicação.

Por fim, a metodologia envolveu testes de usabilidade e acessibilidade que orientaram ajustes no tempo de resposta, na clareza dos feedbacks e na estabilidade geral da ferramenta. Cada etapa foi validada individualmente antes da integração ao sistema completo, resultando em uma solução funcional, acessível e compatível com os objetivos de inclusão digital do projeto.

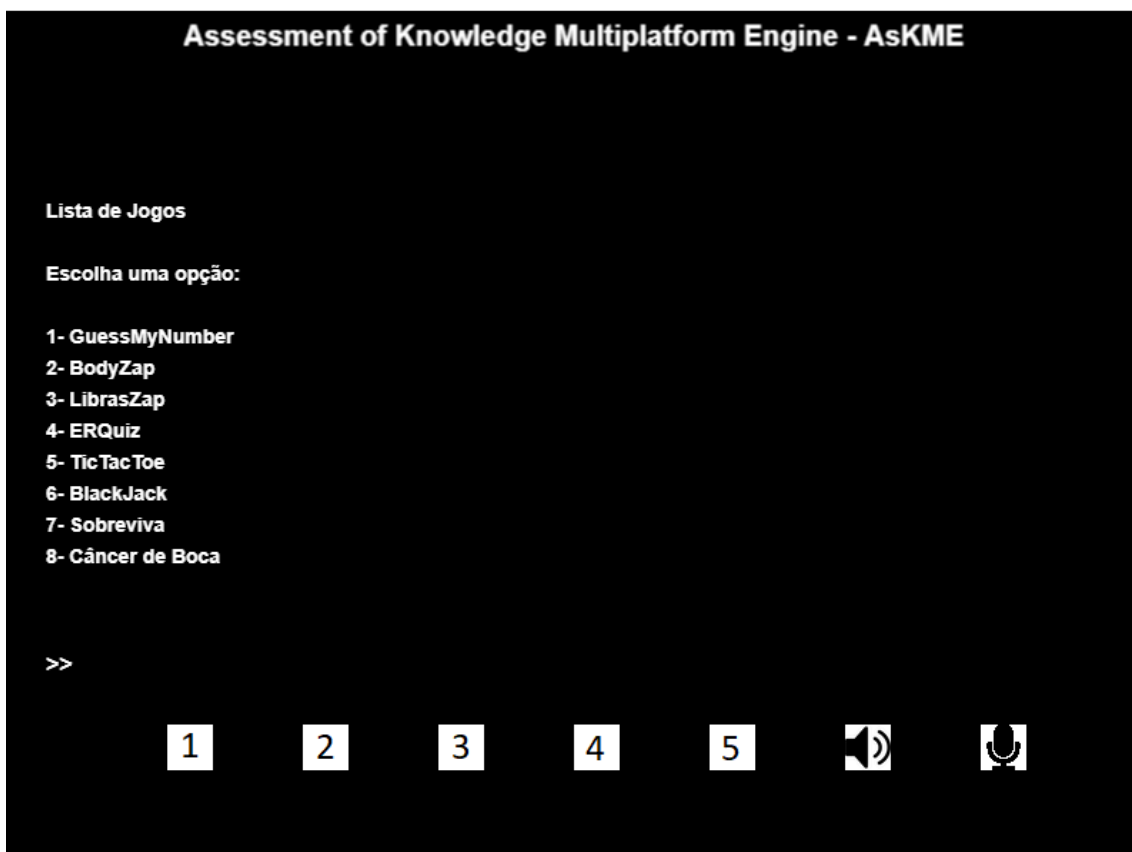


Figura 1: Interface principal da AsKME exibindo a lista de jogos disponíveis e os ícones de acessibilidade na parte inferior: som para leitura de tela e microfone para reconhecimento de voz.

CONSIDERAÇÕES FINAIS (ou Conclusão)

O projeto comprovou a viabilidade da integração entre leitura de tela e reconhecimento de voz em ambientes de jogos educacionais. A abordagem híbrida adotada mostrou-se eficiente para compatibilizar diferentes navegadores e oferecer acessibilidade ampla.

Além de cumprir os objetivos técnicos, o trabalho contribuiu significativamente para a inclusão digital e educacional, permitindo que pessoas com deficiência visual e motora participem de forma mais ativa no universo dos jogos digitais. A iniciativa reforça o papel das tecnologias assistivas como ferramentas essenciais de equidade e inclusão, consolidando a plataforma AsKME como um ambiente mais acessível, interativo e inclusivo.

REFERÊNCIAS

SARINHO, V.; DE AZEVEDO, G.; BOAVENTURA, F. AsKME: A Feature-Based Approach to Develop Multiplatform Quiz Games, 2018, pp. 389-398.

CMU Sphinx, "PocketSphinx: Lightweight speech recognition engine," Carnegie Mellon University, 2023. <https://cmusphinx.github.io/wiki/tutorialpocketsphinx/>.

MDN Web Docs, "SpeechSynthesis – Web Speech API," Mozilla, 2023. <https://developer.mozilla.org/en-US/docs/Web/API/SpeechSynthesis>.

MDN Web Docs, "webkitSpeechRecognition – Web Speech API," Mozilla, 2023. https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API.

A. Zhang, "SpeechRecognition 3.10.0: Speech recognition module for Python," PyPI, 2023. <https://pypi.org/project/SpeechRecognition/>.

Google Cloud, "Speech-to-Text Documentation," Google, 2023. <https://cloud.google.com/speech-to-text/docs>.

Python Software Foundation, "Python 3.11. Documentation," Python.org, 2023. <https://docs.python.org/3/>.

Amazon Web Services, "Alexa Skills Kit (ASK) Documentation," Amazon, 2023. <https://developer.amazon.com/en-US/alexa/alexa-skills-kit>.