

# Voice Hero: Um Jogo Digital com Controle por Reconhecimento de Voz para o Auxílio de Terapias Fonoaudiológicas

Nadine Cerqueira Marques<sup>1</sup>, Victor Travassos Sarinho<sup>1</sup>

<sup>1</sup>Laboratório de Entretenimento Digital Aplicado (LEnDA)  
Departamento de Exatas (DEXA)  
Universidade Estadual de Feira de Santana (UEFS)  
Feira de Santana – BA – Brasil

nadymarkes@gmail.com, vsarinho@uefs.br

**Abstract.** *Digital games can be designed to provide specific stimuli for speech evaluation and training. This article presents Voice Hero, a voice-controlled tower defense game developed to assist speech therapy for children. In this game, players must accurately reproduce words that are appropriate for the social and/or linguistic reality of Brazilian Portuguese in order to defeat enemies and save their fortress. As a result, Voice Hero successfully encourages players to pronounce 116 possible words that are randomly presented during a gameplay session, thus creating a tool that can be used for both evaluation and speech training in children with typical or deviant speech acquisition.*

**Keywords:** *Speech Therapy; Serious Games; Voice Recognition.*

**Resumo.** *Jogos digitais podem ser projetados para oferecer estímulos específicos para a avaliação e o treinamento fonoaudiológico. Este artigo apresenta o jogo Voice Hero, um jogo estilo tower defense com controle por reconhecimento de voz desenvolvido com o objetivo de auxiliar terapias fonoaudiológicas em crianças. Trata-se de um jogo onde o jogador deve reproduzir palavras adequadas à realidade social e/ou linguística do português brasileiro a fim de derrotar os inimigos e salvar sua fortaleza. Como resultado, Voice Hero consegue induzir o jogador na pronúncia de 116 palavras possíveis apresentadas de forma aleatória em uma partida, compondo assim um instrumento capaz de ser utilizado tanto na avaliação como no treinamento fonoaudiológico de crianças com aquisição típica ou desviante.*

**Palavras-chave:** *Fonoaudiologia; Serious Games; Reconhecimento de Voz.*

## 1. Introdução

Jogos digitais podem ser definidos como ambientes atrativos e interativos que cativam o jogador ao oferecer desafios que requerem níveis crescentes de habilidades [Balasubramanian and Wilson 2006]. Tratam-se de ferramentas que permitem a criação de simulações capazes de facilitar a aprendizagem de temas diversos por parte de seus jogadores, através da prática de conhecimentos específicos em um ambiente focado e repetível [Aldrich 2003].

Jogos digitais também podem ser projetados para oferecer estímulos específicos para o treinamento fonoaudiológico, como atividades de discriminação auditiva, repetição

de sons, associação entre fonemas e grafemas, entre outras [Lopes et al. 2016], [Saeedi et al. 2022]. Segundo [Wren and Roulstone 2008], o número de softwares que podem ser utilizados para auxiliar terapias da fala e da linguagem em crianças têm crescido de forma significativa ao longo dos anos. [Wren and Roulstone 2008] também afirma que, caso o software correto seja utilizado, novas e estimulantes ferramentas podem ser fornecidas aos clínicos da área da Fonoaudiologia, de modo que o uso de tais recursos poderia acarretar em um maior aproveitamento de tempo do terapeuta, caso seja provado que os avanços da fala e da linguagem da criança ocorram de uma forma mais adequada com a utilização do respectivo recurso.

Diferentes dispositivos de entrada de dados têm sido usados para o controle de interfaces de jogos diversos, tais como: microfones, acelerômetros, câmeras para o reconhecimento de imagens e de movimentos do jogador, entre outros [Chen and Osman 2014]. Tratam-se de dispositivos que detectam aspectos particulares de seus jogadores de tal maneira que se torna possível aplicar os mesmos na produção de abordagens personalizadas e direcionadas de interação com um jogo, mas com o objetivo de atender pacientes em terapias com demandas específicas.

Neste sentido, este artigo apresenta o jogo *Voice Hero*, um jogo estilo tower defense com controle por reconhecimento de voz desenvolvido com o objetivo de auxiliar terapias fonoaudiológicas em crianças. Trata-se de um jogo onde o jogador deve reproduzir palavras adequadas à realidade social e/ou linguística do português brasileiro a fim de derrotar os inimigos e salvar sua fortaleza. Como resultado, *Voice Hero* consegue induzir o jogador na pronúncia de 116 palavras possíveis apresentadas de forma aleatória em uma partida, compondo assim um instrumento capaz de ser utilizado tanto na avaliação como no treinamento fonoaudiológico de crianças com aquisição típica ou desviante.

## 2. Trabalhos Relacionados

Como exemplos de jogos digitais voltados para terapias fonoaudiológicas, o Fofuuu Fono: Terapia Divertida<sup>1</sup> é uma plataforma para desenvolvimento da fala que propõe exercícios de fonoaudiologia por meio da realização de missões, jogos e brincadeiras apresentadas ao jogador (Figura 1). A inclusão de novas lições personalizadas para o jogador por parte do profissional de fonoaudiologia também pode ser realizada no jogo proposto.

O aplicativo Boca Feliz<sup>2</sup> tem como objetivo ajudar crianças com distúrbios de motricidade oral (e.g., apraxia verbal) através de 15 exercícios demonstrados por um mascote representado por um pato. Através de animações, o patinho pede as crianças que executem movimentos com a língua e lábios, tais como: mover a língua para cima, para direita e esquerda; sorrir; soltar um beijo; tentar tocar o nariz com a língua, dentre outros.

Por fim, o jogo sPeAK-MAN [Tan et al. 2013] traz a proposta de um jogo sério baseado no clássico Pac-Man com a mecânica de incorporar a vocalização de palavras comumente utilizadas em sessões de terapias de fonoaudiologia. Essa mecânica consiste em mostrar palavras sobre os fantasmas, de modo que, quando uma palavra é pronunciada, o fantasma associado a ela se enfraquece, permitindo assim que o mesmo possa ser derrotado pelo jogador (Figura 1).

<sup>1</sup><https://play.google.com/store/apps/details?id=com.fofuuu.apps.therapist>

<sup>2</sup><http://www.ipadfono.com/boca-feliz/>



Figura 1. Imagens dos jogos Fofuu, Boca Feliz e sPeAK-MAN.

### 3. Metodologia

Voice Hero segue o gênero *tower defense* em que é necessário defender uma torre de hordas de inimigos que vão em direção a ela. Para realizar um ataque aos inimigos que tentam invadir a torre, é necessário que o jogador diga as palavras que aparecem na tela do jogo de forma correta. Essas palavras são identificadas por um módulo de reconhecimento de voz e enviadas para a aplicação do jogo usando um protocolo de comunicação. Caso a palavra reconhecida seja igual à palavra apresentada no jogo, o inimigo recebe dano e o jogador ganha um ponto.

Neste sentido, o ponto inicial de desenvolvimento do jogo Voice Hero ocorre com a produção do módulo para reconhecimentos das palavras. Este foi desenvolvido com a linguagem de programação *Python*<sup>3</sup> utilizando o pacote *SpeechRecognition*<sup>4</sup> que faz uso do motor de reconhecimento de voz *Google Speech Recognition*. Como este motor faz o reconhecimento de frases inteiras, foi necessário separar as palavras da frase reconhecida, bem como criar um servidor UDP que faz o envio da palavra reconhecida para um cliente com *host* e porta específicos (Figura 2).



Figura 2. Código parcial e fluxograma da função para reconhecimento e envio das palavras - Fonte: Autoria Própria

<sup>3</sup><https://www.python.org/>

<sup>4</sup><https://pypi.org/project/SpeechRecognition/>

Para a implementação do jogo em si, este foi desenvolvido no *Godot*<sup>5</sup> tendo como base na linguagem de programação *GDScript* que segue uma sintaxe semelhante a da linguagem Python utilizada no desenvolvimento do módulo de reconhecimento de voz. Godot é um motor *open source* para a produção de jogos *cross* plataforma 2D e 3D, o qual permite a criação de clientes UDP capazes de se conectarem ao módulo de reconhecimento de voz implementado.

Por fim, com relação a definição das palavras a serem reconhecidas durante uma partida, estas foram escolhidas a partir do trabalho de [Savoldi et al. 2013], o qual indica 116 palavras possíveis para compor um instrumento que avalie a fonoaudiologia de crianças com aquisição típica ou desviante. Assim, para cada nova partida, estas palavras são selecionadas e apresentadas na tela do jogo de forma aleatória conforme as mecânicas sugeridas para o mesmo.

#### 4. Resultados Obtidos

O jogo Voice Hero tem como base duas cenas principais de jogo, a primeira referente ao menu inicial do jogo (Figura 3) e a segunda referente ao campo de batalha onde a torre será defendida (Figura 4). No menu inicial, o jogador tem basicamente a opção de iniciar uma nova partida ou de sair do jogo encerrando a execução do mesmo.

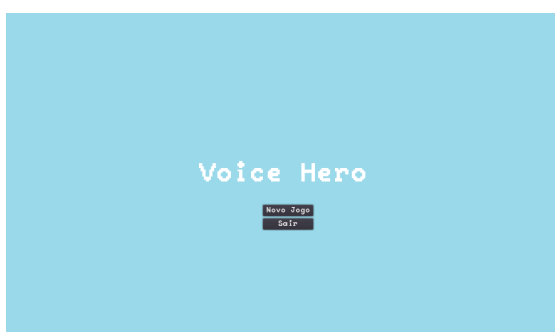


Figura 3. Tela de Menu inicial -  
Fonte: Autoria Própria

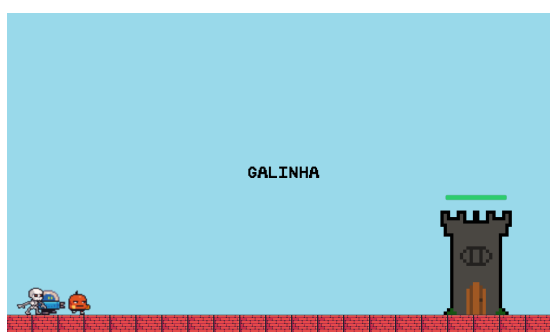


Figura 4. Tela de Novo Jogo -  
Fonte: Autoria Própria

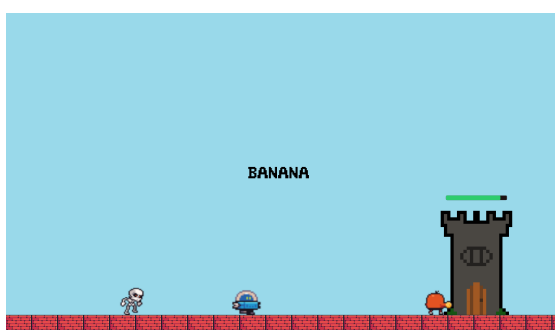


Figura 5. Tela de Dano à Torre -  
Fonte: Autoria Própria

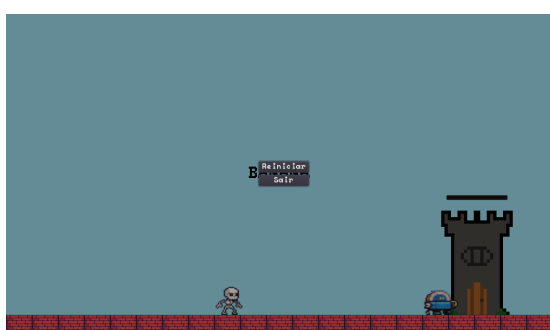
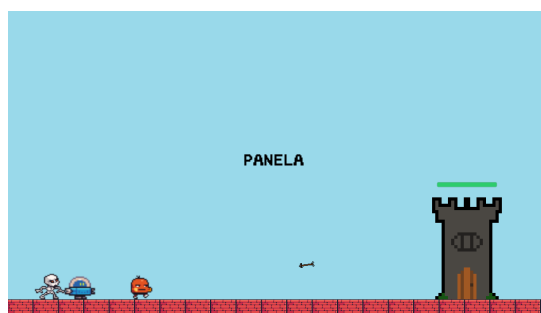


Figura 6. Tela de Fim de Jogo -  
Fonte: Autoria Própria

Na cena de batalha do jogo, o cenário com a torre, os inimigos e a palavra a ser reproduzida são apresentados na tela (Figura 4). Caso os inimigos se aproximem da torre

<sup>5</sup><https://godotengine.org/>

antes de serem derrotados, eles provocam dano à torre (Figura 5). Caso a torre perca todos os pontos de vida, simbolizada pela tarja verde, o jogo é finalizado sendo possível iniciar uma nova partida ou encerrar a execução do jogo (Figura 6). Em caso de acerto da palavra apresentada na tela, a torre ataca o inimigo (Figura 7), o jogador ganha pontos e uma nova palavra é apresentada.



**Figura 7. Tela de Ataque do Jogador - Fonte: Autoria Própria**

Com relação à validação do jogo, esta ainda se encontra em fase preliminar e com foco na usabilidade e no engajamento proporcionado pelo mesmo, a qual está sendo realizada tanto pela equipe de desenvolvimento como também por fonoaudiólogos que trabalham com crianças em geral. Neste sentido, melhorias referentes a objetivos a serem alcançadas no jogo, apresentação de estímulos para um jogar contínuo, agrupamento de palavras relacionadas a fonemas específicos, detalhes gráficos referentes ao passar do tempo no jogo, dentre outros, foram devidamente identificados e considerados para implementação futura do mesmo.

## **5. Conclusões e Trabalhos Futuros**

Jogos digitais têm mostrado grande potencial no campo da terapia fonoaudiológica, fornecendo ambientes interativos e envolventes para atividades de treinamento específicas. Neste sentido, o Voice Hero se soma ao crescente conjunto de jogos digitais projetados especificamente para terapias fonoaudiológicas, oferecendo como resultado uma experiência de jogo única baseada no reconhecimento de voz.

De fato, o Voice Hero traz sua contribuição para o campo da fonoaudiologia, reforçando o potencial dos jogos digitais como uma ferramenta para avaliação e treinamento em terapias fonoaudiológicas com base no português brasileiro. Assim, através de uma proposta de solução gamificada para um público infantil, o Voice Hero se apresenta como uma opção na busca pela motivação e pelo engajamento das crianças durante as sessões de pacientes em tratamentos fonoaudiológicos.

Como trabalhos futuros, se faz necessário melhorar a usabilidade e o engajamento de longo termo do jogo, bem como os objetivos de jogabilidade e de qualidade no polimento gráfico do mesmo, aprimorando assim a experiência geral do usuário com o jogo proposto. Além disso, se faz necessário coletar o feedback dos usuários e efetuar a aplicação de estudos clínicos, de modo a avaliar a eficácia do jogo em ambientes de terapia fonoaudiológica voltados para crianças em geral.

## Referências

- Aldrich, C. (2003). *Simulations and the Future of Learning: An Innovative (and Perhaps Revolutionary) Approach to e Learning*. John Wiley & Sons.
- Balasubramanian, N. and Wilson, B. G. (2006). Games and simulations. *Society for information technology and teacher education international conference*.
- Chen, R. and Osman, S. (2014). Blow tracking user interface system and method. US Patent 8,786,698.
- Lopes, M., Magalhães, J., and Cavaco, S. (2016). A voice-controlled serious game for the sustained vowel exercise. In *Proceedings of the 13th International Conference on Advances in Computer Entertainment Technology*, pages 1–6.
- Saeedi, S., Bouraghi, H., Seifpanahi, M.-S., and Ghazisaeedi, M. (2022). Application of digital games for speech therapy in children: a systematic review of features and challenges. *Journal of healthcare engineering*, 2022.
- Savoldi, A., Ceron, M. I., and Keske-Soares, M. (2013). Quais são as melhores palavras para compor um instrumento de avaliação fonológica? *Audiology-Communication Research*, 18:194–202.
- Tan, C. T., Johnston, A., Ballard, K., Ferguson, S., and Perera-Schulz, D. (2013). speak-man: Towards popular gameplay for speech therapy.
- Wren, Y. and Roulstone, S. (2008). A comparison between computer and tabletop delivery in phonology therapy. *International Journal of Speech-Language Pathology*, 10(5):346–363.