

Projeto e implementação de uma arquitetura baseada em jogos digitais para coleta de áudios e avaliação de fluência na leitura infantil

Leandro Dornela Ribeiro, Mateus G. do Nascimento, Warley Almeida Silva, Jorão Gomes Jr.,
Igor de Oliveira Knop, Eduardo Barrere, Jairo Francisco de Souza

*Departamento de Ciência da Computação
Universidade Federal de Juiz de Fora*

Juiz de Fora, Brasil

{leandrodornela, mateus.goncalo, warley.almeida, joraojunior, igorknop, eduardo.barrere, jairo.souza}@ice.uff.br

Resumo—A fluência na leitura permite a inserção do indivíduo na sociedade e é também um dos requisitos da Base Nacional Comum Curricular. O *Talk2Me* é uma abordagem para realizar a coleta e classificação de áudios para auxiliar educadores a atuarem de forma mais eficiente nas resolução de problemas de leitura. Este trabalho apresenta como mecanismo de coleta uma dinâmica em sala de aula na forma de um jogo digital que atua como interface para tornar o processo de captura mais lúdico. Uma arquitetura de software é proposta e o conjunto de regras é implementado como um protótipo composto por dois clientes diferentes e serviços via internet.

Keywords—Fluência em leitura; Jogo digital; Reconhecimento automático de fala.

I. INTRODUÇÃO

A fluência em leitura é importante para inserção de um indivíduo na sociedade [1] e um critério de avaliação obrigatório pelo Ministério da Educação através da Base Nacional Comum Curricular [2]. Dessa forma, o desenvolvimento da fluência em leitura deve ser adequadamente conduzido principalmente durante a fase de alfabetização. No entanto, o acompanhamento em sala de aula do desenvolvimento dos alunos pelos professores não é uma tarefa trivial, pois cada aluno possui especificidades no que diz respeito à forma de aprendizagem e às dificuldades encontradas, tornando-se uma tarefa praticamente impossível [3].

Nesse sentido, sistemas para reconhecimento automático de fala (do inglês, *Automatic Speech Recognition* ou ASR) tornam-se uma alternativa para automatização do processo. Em geral, sistemas ASR tem como objetivo primário a transformação de palavras ditas em áudio em texto [4]. Entretanto, trabalhos recentes mostram que é possível construir sistemas para avaliação de fluência em leitura utilizando essa abordagem [5].

Este trabalho apresenta a abordagem computacional *Talk2Me*, a qual realiza a coleta e classificação de áudios de leituras para auxiliar educadores no monitoramento e resolução de problemas durante o processo de obtenção de fluência. O jogo digital *SpaceGEMS* transforma a coleta de áudio em dinâmicas em sala de aula para motivar o processo de captura dos áudios para as crianças.

II. REFERENCIAL TEÓRICO E TRABALHOS CORRELATOS

Ser fluente em leitura consiste em decodificar oralmente um texto com velocidade e precisão adequadas, segmentando corretamente os fonemas pronunciados [6]. São considerados para avaliação de fluência de um indivíduo: (i) a precisão de leitura, que consiste na quantidade de acertos sobre a quantidade de palavras lidas; (ii) a taxa de leitura de palavras, que consiste na quantidade de palavras lidas em um determinado período de tempo; e (iii) a prosódia, que consiste na correta pronúncia da sílaba tônica de uma palavra segundo a norma padrão. Ler corretamente é uma habilidade chave para a integração e participação do ser humano na sociedade [1], mas o acompanhamento do desenvolvimento da leitura durante a alfabetização não é uma tarefa fácil para professores e gestores [3]. Entre as abordagens para aliviar essas dificuldades, aproveita-se do lúdico, que é um elemento importante na educação infantil [7]. Desta forma, a produção de jogos específicos para o auxílio no processo de avaliação de fluência na leitura têm alto potencial para melhorar o processo de alfabetização.

O termo *jogos sérios* [8], [9] começou a ser atribuído a jogos nos quais há um objetivo fim não relacionado ao divertimento autocontido [10], [11]. Seu uso como ferramenta para educação, treinamento e trabalho é aceito e bem documentado na literatura nas últimas décadas [12]. Já a gamificação difere dos jogos sérios por utilizar elementos presentes em jogos em um contexto diferente, como em uma atividade produtiva, educacional ou social com o objetivo de aumentar a motivação e o envolvimento nas tarefas fim [13]. Tanto os jogos sérios quanto a gamificação têm sido aplicados com sucesso em diversas áreas da educação [14] visando estimular um maior engajamento dos alunos e tornar as atividades menos cansativas e mais interessantes.

A utilização de jogos para estimular maior engajamento dos alunos em suas atividades não é algo novo: [15] defende o uso de jogos digitais para o desenvolvimento da cognição, enquanto [16] apresenta, compara e analisa exemplos de jogos digitais que auxiliam no ensino de programação. Da mesma forma, [17] traça uma análise da forte presença da

aplicação de gamificação no contexto educacional mas pouca aplicação no que se diz respeito à avaliação da fluência em língua materna.

Para capturar especificidades da língua materna, alguns trabalhos focam nas pessoas com necessidades especiais. Em [18], os autores propuseram um jogo para auxiliar o desenvolvimento da pronúncia de indivíduos autistas. Já [19] trabalhou no desenvolvimento da leitura com indivíduos disléxicos e [20] desenvolveu um jogo para melhorar a pronúncia de crianças que apresentam dificuldades fonológicas.

Em [5], os autores propõem uma ferramenta para avaliação de fluência da língua materna em larga escala pelo uso de um aplicativo convencional, sem contar com alguma interface que estimule o engajamento do aluno. Os autores destacam que a coleta dos áudios é um ponto a ser trabalhado para melhorar da abordagem, uma vez que questões técnicas (como ambiente ruidoso e qualidade dos aparelhos) e questões pessoais (desinteresse da criança e baixa afinidade com tecnologia) influenciam diretamente na qualidade dos resultados finais.

III. MÉTODO E DESENVOLVIMENTO

O método de criação da abordagem *Talk2Me* teve início com uma reunião com um grupo interdisciplinar para expor o cenário de avaliação de fluência. O grupo formado pelos pesquisadores da área de fluência, especialistas em desenvolvimento de software e desenvolvedores de jogos realizaram um *brainstorming* a fim de criar uma dinâmica que agilizasse o processo.

Em experiência de trabalhos anteriores [5], o grupo de fluência construiu com o auxílio do *toolkit open-source* de ASR chamado Kaldi¹ um modelo acústico treinado com videoaulas. Recentemente, ao se deparar com um novo contexto de avaliação de áudios com vozes infantis, tem realizado experimentos no treinamento de modelos acústicos infantis e investigando o processo de avaliação das crianças.

A proposta evoluiu para a criação de uma dinâmica de grupo *in loco*, controlado por um jogo digital disponível via internet através do uso de um único *smartphone* ou computador mas que permitisse compartilhar o estado com outros dispositivos. Esse jogo teria um conjunto de regras mínimo, com uma seleção de mecanismos conhecidos que geram o engajamento em jogos, como objetivos bem definidos, competição (conflito indireto), resposta frequente, recompensas para esforço e aprovação entre pares [21].

Para a dinâmica, foi criado um jogo simples (*Space-GEMS*) de pegar e entregar [22] no qual os jogadores seriam divididos em duas equipes. A disputa pela maior pontuação dentro de um número limitado de rodadas seria o drama principal e as ações dos jogadores discutidas em grupo para coordenar o movimento ao longo da partida.

¹<https://kaldi-asr.org/>

A temática escolhida define que cada equipe representa uma megacorporação espacial buscando aumentar seu faturamento pela entrega de joias entre planetas. As joias em três tipos diferentes (rubis, esmeraldas e safiras) são recolhidas e entregues pelos jogadores ao mover-se entre os planetas. O tema jogo é deixado propositalmente mais abstrato para que o professor assuma o papel narrativo, não mediado, característico entre os jogos de mesa e também para facilitar a readequação uma série de aplicações futuras.

A. Projeto e implementação da arquitetura do *Talk2Me*

A abordagem *Talk2Me* seguiu uma linha modular, para permitir a expansão posterior ou mesmo a substituição de funcionalidades. A Figura 1 apresenta uma visão geral da arquitetura utilizada.

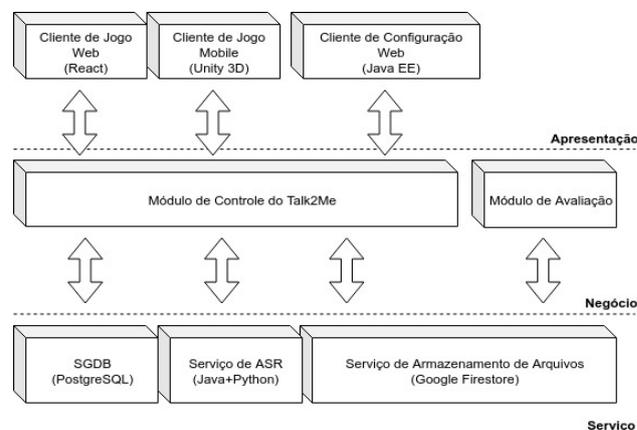


Figura 1. Visão geral da arquitetura do *Talk2Me*.

Estão disponíveis três clientes de acesso diferentes: *i*) um cliente web de página única para o jogo que pode ser acessado por qualquer navegador em um dispositivo móvel em uma estação de trabalho do educador; *ii*) um cliente para o jogo para dispositivos Android; e *iii*) um cliente web para configuração e gestão das salas de alunos.

O Módulo de Controle *Talk2Me* orquestra todos os demais serviços, persistência e comunicação direta com os clientes de jogo. A cada ação de um jogador recebida, recebe o áudio capturado no cliente, envia o arquivo para o Serviço de armazenamento de arquivos, gera um pedido para análise do áudio no Serviço de ASR e avança o estado do jogo.

O serviço ASR é especializado na avaliação da leitura e permite levantar informações sobre os fonemas pronunciados pela criança e os fonemas que deveriam ser pronunciados. A lista de seqüências de fonemas corretos para cada palavra podem ser alteradas para abranger diferentes pronúncias por conta de diferenças regionais de acordo com o local de aplicação do jogo. Essa lista de seqüências de fonemas por palavras devem ser editadas previamente pelo professor. O Módulo de Avaliação é responsável por avaliar as jogadas,

os resultados do ASR, os arquivos de áudio e as informações colhidas durante e após uma partida.

IV. RESULTADOS E DISCUSSÃO

No Módulo de Controle foi implementado como uma aplicação Java EE e o funcionamento de sua interface web pode ser observado pela telas da Figura 2, na qual uma partida é cadastrada e seu estado é exibido.



Figura 2. Cliente do Módulo de Controle permite criar previamente a partida e acompanhar os alunos da atividade: a) cadastro de palavras da partida; e b) acompanhamento dos alunos.

O primeiro cliente da Figura 3, chamado de *Galaxy Reader* foi desenvolvido exclusivamente com tecnologias para internet para ter um público mais abrangente, em várias arquiteturas de hardware e sistemas operacionais, desde que um navegador e atualizado esteja disponível.

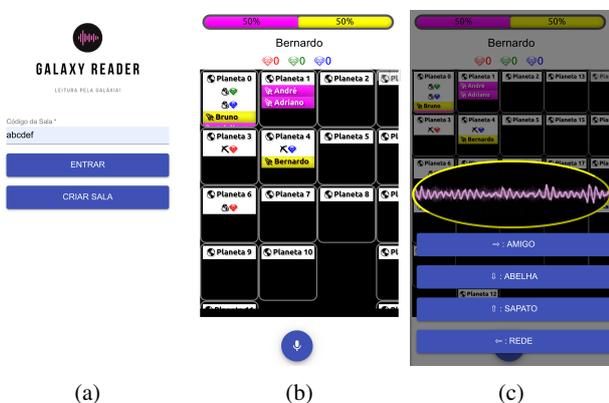


Figura 3. Cliente *Galaxy Reader* em ReactJS com baixo consumo de recursos: a) tela de entrada; b) visão geral da partida; e c) diálogo de captura de áudio dos jogadores.

Um segundo cliente, chamado de *SpaceGEMS* e apresentado na A Figura 4, foi desenvolvido buscando oferecer um melhor apelo entre as crianças, utilizando animações e modelos tridimensionais.

O cliente *SpaceGEMS* aproveita do poder de processamento do cliente para reimplementar as regras de gestão do jogo localmente. Ele realiza a captura local e armazenar os áudios no dispositivo para posterior envio para o servidor e subsequente avaliação.

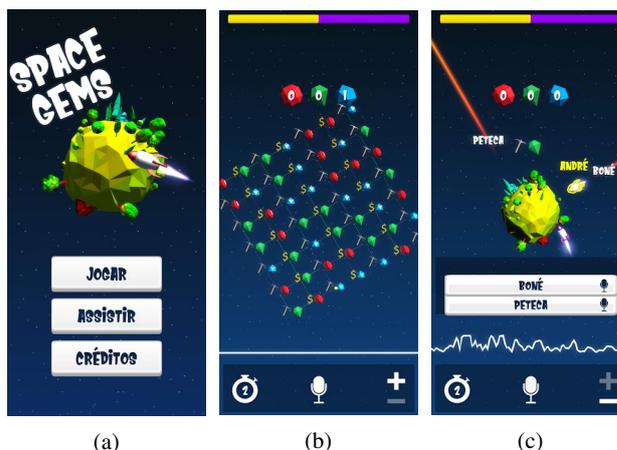


Figura 4. Cliente *SpaceGEMS* em Unity3d com recursos de animação e música: a) tela de entrada no jogo; b) visão geral da partida; e c) diálogo de captura de áudio de jogadores.

O Serviço de ASR, foi avaliado através do uso de áudios de crianças pronunciando palavras isoladas. Esse conjunto de áudios contém aproximadamente 36 mil palavras lidas por 879 falantes distintos, todos com idades entre 6 e 9 anos. Cada criança leu no máximo 50 palavras. A base de dados cedida pelo *Centro de Políticas Públicas e Avaliação da Educação (CAEd/UFJF)* e foi gerada após a aplicação de teste de fluência em crianças de escolas públicas dos estados de Espírito Santo, Pará, Paraíba e Pernambuco. Os resultados de cada falante foram agrupados para simular a sua pontuação no jogo. Todas os áudios possuem uma avaliação manual e o resultado do jogo foi comparado com a avaliação manual de forma a verificar a acurácia do sistema.

Como resultado da avaliação, verificou-se que o jogo acerta a leitura da crianças em 86% das palavras (acurácia), conforme mostra a Figura 5. Segundo os valores encontrados, o ASR encontra a maior parte das palavras corretas lidas pelas crianças.

Matriz de confusão		Manual		
		Certo	Errado	Total
Automático	Certo	30618	1038	31656
	Errado	3792	484	4276
	Total	34410	1522	35932

Figura 5. Matriz de confusão gerada pelo serviço de ASR ao avaliar 36 mil leituras de áudio de uma base de dados.

V. CONCLUSÕES

Este trabalho apresentou a abordagem *Talk2Me*, uma plataforma modular para utilizar jogos digitais como meio de captura de áudios de leitura de crianças e avaliá-los através do reconhecimento automático de fala. Através do uso de jogos digitais espera-se que o processo de coleta

seja facilitado, aumentando consideravelmente o número de amostras de leitura que, em um segundo momento, poderão ser utilizadas pelos educadores e gestores para identificar problemas de fluência e projetar melhor ações de atuação.

Os resultados das avaliações individuais de cada módulo demonstram que o projeto está pronto para a condução de experimentos em salas de aula, para refinamento da dinâmica e método de coleta. Como trabalho futuro, o módulo de avaliação será desenvolvido em conjunto com os educadores antes do experimento de campo para gerar relatórios mais úteis em função da real necessidade das instituições de ensino.

AGRADECIMENTOS

Agradecemos ao CAEd/UFJF pelo fornecimento dos dados para treinamento do modelo de reconhecimento de fala, à UFJF pelo financiamento do projeto através das Pró-Reitorias de Extensão, Graduação e do Programa de Educação Tutorial e à CAPES pelo financiamento dos pesquisadores.

REFERÊNCIAS

- [1] N. M. Dias, C. B. R. Leão, T. d. C. B. Pazeto, G. L. L. Martins, A. P. P. Pereira, and A. G. Seabra, “Avaliação da leitura no Brasil: revisão da literatura no recorte 2009-2013,” *Psicologia: teoria e prática*, vol. 18, pp. 113 – 128, 04 2016.
- [2] MEC, “Base nacional comum curricular,” <http://basenacionalcomum.mec.gov.br/wp-content/uploads/2018/02/bncc-20dez-site.pdf>, 2018, acessado: 28-06-2019.
- [3] P. G. Celestino, “A oralidade infantil e desenvolvimento cognitivo à partir da prática docente,” *Revista Internacional de apoyo a la inclusión, logopedia, sociedad y multiculturalidad*, vol. 5, no. 1, 2019.
- [4] D. Yu and L. Deng, *AUTOMATIC SPEECH RECOGNITION*. Springer, 2016.
- [5] E. Soares, L. C. Carchedi, J. Gomes Jr, E. Barrère, and J. Souza, “Avaliação automática da fluência em leitura para crianças em fase de alfabetização,” in *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, vol. 29, no. 1, 2018, p. 11.
- [6] L. S. Fuchs, D. Fuchs, M. K. Hosp, and J. R. Jenkins, “Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis,” *Scientific studies of reading*, vol. 5, no. 3, pp. 239–256, 2001.
- [7] P. S. V. d. S. Rocha, *A importância do lúdico na educação infantil: uma análise a partir da concepção de professores*. Universidade Federal da Paraíba, 2018.
- [8] C. C. Abt, *Serious games*. Viking Press, 1970.
- [9] D. R. Michael and S. L. Chen, *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade, 2005.
- [10] J. Huizinga, *Homo ludens: o jogo como elemento da cultura*. Editora da Universidade de S. Paulo, Editora Perspectiva, 1971, vol. 4.
- [11] K. Salen, K. S. Tekinbaş, and E. Zimmerman, *Rules of play: Game design fundamentals*. MIT press, 2004.
- [12] D. Djaouti, J. Alvarez, J.-P. Jessel, and O. Rampnoux, “Origins of serious games,” in *Serious games and edutainment applications*. Springer, 2011, pp. 25–43.
- [13] C. I. Muntean, “Raising engagement in e-learning through gamification,” in *Proc. 6th International Conference on Virtual Learning ICVL*, vol. 1, 2011.
- [14] S. de Sousa Borges, V. H. Durelli, H. M. Reis, and S. Isotani, “A systematic mapping on gamification applied to education,” in *Proceedings of the 29th annual ACM symposium on applied computing*. ACM, 2014, pp. 216–222.
- [15] D. K. Ramos, B. S. Anastácio, G. A. da Silva, C. Venturieri, N. Stange, and M. E. Martins, “Jogos digitais, habilidades cognitivas e motivação: percepção das crianças no contexto escolar,” *Anais do Simpósio Brasileiro de Games e Entretenimento Digital-SBGames2018*, 2018.
- [16] R. S. Monclar, M. A. Silva, and G. Xexéo, “Jogos com propósito para o ensino de programação,” *Anais do Simpósio Brasileiro de Games e Entretenimento Digital-SBGames2018*, 2018.
- [17] F. R. De Paula and R. Fávero, “A gamificação da educação na compreensão dos profissionais da educação,” *Anais do Simpósio Brasileiro de Games e Entretenimento Digital-SBGames2016*, 2016.
- [18] A. Anwar, M. M. Rahman, S. Ferdous, S. A. Anik, and S. I. Ahmed, “A computer game based approach for increasing fluency in the speech of the autistic children,” in *Advanced Learning Technologies (ICALT), 2011 11th IEEE international conference on*. IEEE, 2011, pp. 17–18.
- [19] M. R. U. Saputra, “Lexipal: Design, implementation and evaluation of gamification on learning application for dyslexia,” *Int. J. Comput. Appl.*, vol. 131, pp. 37–43, 2015.
- [20] P. Luan, T. d. O. L. Silva, G. S. H. Alexandre Muller, F. S. Danielle, and V. C. Siqueira, “Popopó: Conceituando um jogo para apoio na fonoterapia infantil,” in *Proceedings of SBGames 2018 – Education Track*. SBC, 2018, pp. 1404–1407.
- [21] R. Raymer, “Gamification: using game mechanics to enhance elearning,” *ELearn*, vol. 2011, no. 9, p. 3, 2011.
- [22] BoardGameGeek, “Pick-up and Deliver,” <https://boardgamegeek.com/boardgamemechanic/2007/pick-and-deliver>, 2000.