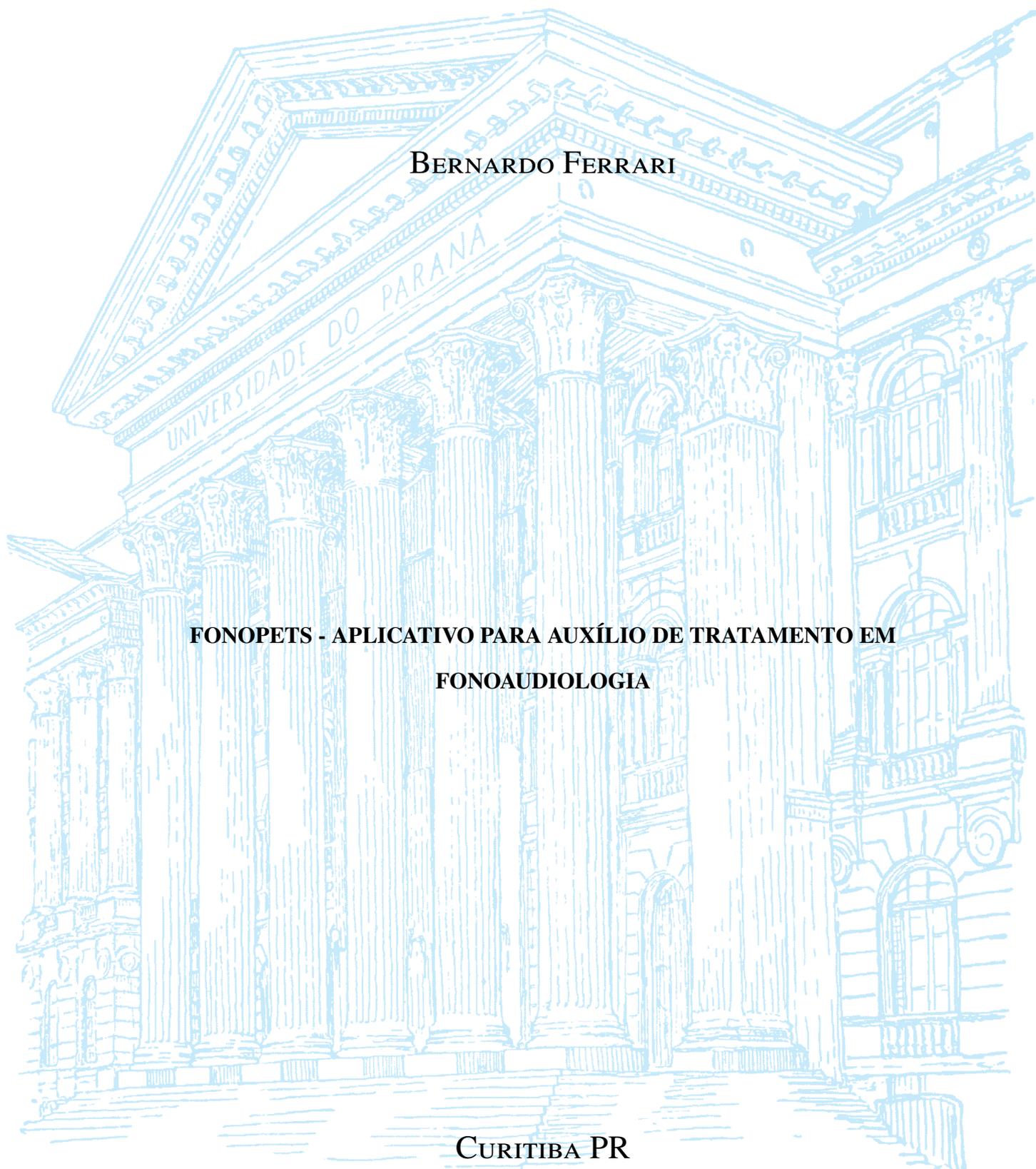


UNIVERSIDADE FEDERAL DO PARANÁ

BERNARDO FERRARI

**FONOPETS - APLICATIVO PARA AUXÍLIO DE TRATAMENTO EM
FONOAUDIOLOGIA**

CURITIBA PR
2018



BERNARDO FERRARI

**FONOPETS - APLICATIVO PARA AUXÍLIO DE TRATAMENTO EM
FONOAUDIOLOGIA**

Trabalho apresentado como requisito parcial à conclusão do Curso de Bacharelado em Ciência da Computação da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Prof. Dr. Roberto Pereira.

**CURITIBA PR
2018**

Agradecimentos

Agradeço a Deus pela minha vida, pela minha família, pelos meus amigos, pelo meu orientador, pelo grupo de IHC da UFPR e por todos que participaram de alguma forma na construção desse trabalho.

Resumo

Em tratamentos de fonoaudiologia com crianças é comum que os profissionais solicitem exercícios de fala que devem ser realizados fora do ambiente hospitalar (e.g., em casa), de modo a melhorar a dicção dos fonemas e facilitar o tratamento. No entanto, devido a diversos fatores socioeconômicos e culturais, como falta de apoio da família, é comum que esses exercícios não sejam realizados. Por se tratar de um público majoritariamente infantil, determinados métodos podem ser utilizados para incentivar a prática desses exercícios. O presente trabalho propõe um jogo baseado na plataforma Android que possa auxiliar no tratamento e acompanhamento de crianças. Utilizando o contexto do Hospital SOBRAPAR de Campinas/SP, foram identificados requisitos e avaliadas soluções existentes no mercado de aplicativos na área de fonoaudiologia, e verificou-se poucas opções gratuitas e capazes de atender às demandas práticas do Hospital. Por meio de uma sessão de Design Colaborativo com métodos do Design Socialmente Consciente, foi desenvolvida uma compreensão do contexto social e aspectos técnicos, tais como requisitos, as partes interessadas e os desafios existentes para os exercícios de fonoaudiologia. A partir disso, foi proposta uma solução que atenda aos requisitos advindos da sessão de Design Socialmente Consciente e tenha características relevantes não presentes nas soluções existentes, e.g. a possibilidade da fonoaudióloga adaptar exercícios para cada paciente. Foi então implementado um protótipo dessa solução proposta, inspirado no jogo Super Mario, em que o usuário precisa falar palavras para inverter a gravidade do personagem e poder cumprir os objetivos do jogo. Esse protótipo foi avaliado com 12 pessoas da área de informática e o resultado foi positivo, embora a tecnologia de reconhecimento de fala (PocketSphinx) tenha sido sensível a ruídos externos e tenha tido dificuldades em diferenciar a fala do usuário de ruído/silêncio.

Palavras-chave: Aplicativo, Fonoaudiologia, Android.

Abstract

In speech therapy with children, it is common for professionals to request speech exercises that should be performed outside of the hospital's environment (eg, at home), in order to improve phonemes' speech and facilitate treatment. However, due to diverse socioeconomic and cultural factors, like the lack of family support, it is common that these exercises are not performed. Since most of the stakeholders are children, certain methods can be used to encourage the practice of these exercises. This work proposes a game based on the Android platform that can aid in the treatment and monitoring of children. Using the context of the SOBRAPAR Hospital from Campinas/SP, requirements were identified and evaluated for existing solutions in the field of speech therapy applications, and there were few free options able to cope with the Hospital's demands and needs. Through a Collaborative Design session with methods of Socially Aware Design, an understanding of the social context and technical aspects, such as requirements, the stakeholders and the challenges for speech-language exercises were developed. From this, a solution was proposed that met the requirements of the Socially Aware Design session and could have relevant characteristics not present in the existing solutions. It was then implemented a prototype of this proposed solution, inspired by the game Super Mario, in which the user needs to speak words to reverse gravity and be able to fulfill the goals of the game. This prototype was evaluated with 12 people in the computer science area and the result was positive, although speech recognition technology (PocketSphinx) was sensitive to external noise and had difficulties in differentiating the user's speech from noise / silence.

Keywords: Android, Speech Therapy, Application.

Sumário

1	Introdução	9
2	Trabalhos Relacionados	11
2.1	Análise baseada em Valores	14
3	Metodologia	19
3.1	Relato da Sessão	19
3.2	Quadro de Prospecção de Valor.	23
3.3	Pontos em comum.	24
3.4	Solução proposta	25
4	Fonopets: Protótipo e Avaliação	28
4.1	Reconhecimento de voz.	30
4.2	Avaliação do protótipo	32
4.3	Resultados.	33
5	Conclusões e Trabalhos Futuros	39
	Referências	41
	Apêndice A: Informações coletadas durante avaliação do protótipo.	45
A.1	Respostas do questionário.	45
A.2	Desempenho no jogo	48

Lista de Figuras

2.1	Menu do aplicativo Lele Sílabas.	13
2.2	Menu do aplicativo Visual Reading.	14
3.1	Identificação dos Stakeholders.	20
3.2	Uma das folhas preenchidas durante o processo de <i>BrainWriting</i>	22
3.3	Quadro projetado durante a sessão de DSC.	24
3.4	Desenho da tela inicial da solução proposta. Na parte superior, há o progresso diário (representado por uma barra) e um animal, cuja emoção é refletida pelo progresso diário. Na parte inferior, há um menu para acessar cada jogo, representado por um animal diferente.	26
3.5	Imagem dos jogos <i>Flappy Bird</i> , <i>Super Mario Bros</i> , <i>Frogger</i> e do aplicativo <i>Yousician</i> [43], obtidas pela Internet [13, 4, 27, 42] e agrupadas pelo autor. . . .	27
4.1	Primeira fase. Apresenta o universo com blocos e cristais.	28
4.2	Segunda fase. Introduce bloco verde, que inverte a direção (horizontal) do movimento do personagem.	29
4.3	Terceira fase. Introduce bloco roxo, que teleporta o personagem 10 blocos para a frente.	29
4.4	Quarta fase. Introduce o bloco cinza, que teleporta o personagem 10 blocos para trás.	29
4.5	Quinta fase. Introduce o bloco azul claro, que pula o personagem.	29
4.6	Sexta fase. Mistura todas as mecânicas já apresentadas e aumenta o nível do desafio.	29
4.7	Sétima fase. Mistura todas as mecânicas anteriores, assim como a sexta fase. . . .	29
4.8	Início da fase 3 revela a introdução de uma nova mecânica e um aumento gradual na dificuldade. Blocos transparentes são renderizados como cristais.	30
4.9	Emoções usadas para representar o sentimento após cada fase.	33

Lista de Tabelas

2.1	Comparação entre os aplicativos analisados.	11
2.2	Comparação usando análise de valores sociais em cada aplicativo encontrado. . .	16
3.1	Problemas e soluções encontrados para cada parte interessada.	21
3.2	Tabela do framework semiótico.	23
4.1	Tabela de expressões regulares usadas para identificar palavras em português no protótipo.	31
4.2	Tempo gasto em cada fase do jogo por cada usuário..	34
4.3	Resultado do <i>SAM Assessment</i> , representando a emoção escolhida por cada participante para cada fase.	35
4.4	Tabela da relação entre pessoas com familiaridade com jogos, a média de cristais coletados, a média de tempo e a média da emoções sentidas nas fases 1,3,4 e 5. . .	36
4.5	Tabela da relação entre pessoas com familiaridade com <i>Super Mario</i> , a média de cristais coletados, a média de tempo e a média de emoções sentidas nas fases 1,3,4 e 5..	36
4.6	Tabela da relação entre pessoas que já cursaram alguma disciplina de IHC, a média de cristais coletados, a média de tempo e a média de emoções para as fases 1,3,4 e 5..	36
4.7	Tabela da média de erros no reconhecimento de fala em relação a cada fase. . . .	37
A.1	Resposta dos questionários com relação a idade, gênero, escolaridade e se a pessoa já fez alguma disciplina de IHC.	45
A.2	Resposta dos questionários com relação a familiaridade com jogos e familiaridade com o jogo <i>Super Mario</i>	45
A.3	Tabela com pontos positivos e negativos da interação.	46
A.4	Tabela com pontos positivos e negativos da interface.	47
A.5	Tabela com comentários, sugestões, dúvidas, problemas, etc.	48
A.6	Número de cristais coletados em relação ao número da fase. A última linha contém os valores máximos de cristais disponíveis por fase.	48
A.7	Número de palavras reconhecidas por usuário em relação ao número da fase e a palavra que devia ser falada. Não há um número máximo de acertos.	49
A.8	Número vezes em que uma palavra foi falada mas o aplicativo entendeu como silêncio ou ruído e ignorou.	49
A.9	Número vezes em que uma palavra foi falada e o aplicativo registrou, mas não encontrou definição no dicionário morfológico usado.	50

A.10	Porcentagem de acertos no reconhecimento de fala por usuário por fase.	50
A.11	Resultado do <i>SAM Assessment</i> agrupado e simplificado. No eixo vertical, estão as fases. No eixo horizontal, a emoção expressada pela pessoa ao jogar as fases. . .	50

1 Introdução

Dificuldades na fala podem repercutir negativamente na vida escolar e social de um indivíduo. De acordo com Schipor et al [37], 10% das crianças entre 4 e 7 anos possuem algum destes distúrbios. Esses problemas variam em grau, sendo mais severos em crianças com deformidades craniofaciais. Em geral, profissionais de fonoaudiologia requerem que seus pacientes realizem exercícios em casa, de preferência sete dias por semana [10].

O Hospital SOBRAPAR (“Sociedade Brasileira de Pesquisa e Assistência para Reabilitação Craniofacial”) é uma instituição privada que oferece tratamento cirúrgico e reabilitação de pacientes com deformidades craniofaciais congênitas ou adquiridas. O atendimento ocorre quinzenalmente, de forma gratuita pelo Sistema Único de Saúde (SUS), e os pacientes muitas vezes moram longe e precisam se deslocar por dezenas ou centenas de quilômetros para a consulta. Assim, a não realização dos exercícios regulares em domicílio pode prolongar o tratamento e aumentar o custo, prejudicando o resultado final esperado. Existe uma certa resistência por parte dos pacientes de comparecer à consulta, pois segundo Schneider [38], as crianças, em especial, podem encarar o ambiente hospitalar como um local aterrorizante. Assim, é benéfico que a execução em casa seja a mais efetiva possível.

Em conversa com duas fonoaudiólogas do Hospital SOBRAPAR [25], foi identificado que grande parte das famílias atendidas não possui internet em casa e a maioria dos softwares para auxílio aos exercícios ou à terapia são pagos e custam em média R\$400 cada. O desejo das profissionais é de que existisse um aplicativo que não dependesse de Internet, não fosse repetitivo e estivesse disponível gratuitamente para os pacientes [25]. Além disso, as fonoaudiólogas gostariam que o aplicativo fizesse gravação do exercício das crianças na forma de vídeo. A gravação pode ser útil para auxiliar, por exemplo, no tratamento de pacientes com o palato fissurado, que possuem uma percepção diferente do som por causa da anomalia [36]. A criança nasce com uma forma diferente de falar e para ela essa forma diferente de falar é a correta, então é necessário ajustar essa percepção até que esteja adequada. Uma forma de apoiar o tratamento é possibilitar que a criança ouça sua própria voz por meio da gravação e visualização.

A gravação de vídeo da criança realizando o exercício pode contribuir também para que o profissional possa acompanhar a evolução da movimentação da boca dos pacientes. Como muitas famílias não possuem acesso a internet, não é possível transmitir as gravações regularmente. Uma solução seria deixar as gravações armazenadas no próprio celular para que, quando a criança for para a consulta, o profissional tenha acesso a esses dados do celular, usando a rede do hospital.

Dentre os desafios para o exercício contínuo nos exercícios de fonoaudiologia estão a falta de motivação do paciente e seus familiares. Estratégias de gamificação e de jogos têm sido exploradas para promover mudanças de comportamento e motivar o engajamento de pessoas. Assim, com base no conhecimento do domínio do Hospital SOBRAPAR, este trabalho considera que um jogo pode ser capaz de promover uma maior interação com os exercícios de fonoaudiologia e engajar a criança mais facilmente, sendo uma opção viável no contexto das crianças em tratamento fonoaudiológico.

Jogos de computador são uma forma de entretenimento com a proposta de engajar jogadores [11]. O apelo de jogos como 'Pac-Man', por exemplo, pode ser relacionado à forma como competência, controle e emoções pessoais são colocadas juntas [8]. Segundo Bowman [8], 'Pac-Man' é previsível no sentido de pontuação, em que todos os jogadores recebem a mesma recompensa ao alcançar o mesmo objetivo. Contudo, é imprevisível no sentido em que cada partida é diferente. Ainda, Bowman aponta que para os jogadores a percepção de sucesso é maior que o medo de fracasso: o jogo pode ser jogado quantas vezes quiser, sem sofrer consequências ou tirar notas ruins na escola.

Este trabalho propõe um jogo, na forma de aplicativo para a plataforma *Android*, para engajar o paciente diariamente na realização dos exercícios de fonoaudiologia em casa, e que permita um acompanhamento efetivo pelo profissional de fonoaudiologia. Aplicativos já existentes para as plataformas *Android* e *iOS* na área de fonoaudiologia foram pesquisados e explorados. Foi observado que, embora a área de fonoaudiologia e jogos seja ampla e possua potencial, existem poucas soluções que possam auxiliar no acompanhamento do tratamento. As melhores soluções encontradas na área possuem custo proibitivamente alto, na faixa de centenas ou milhares de reais por ano por profissional.

A partir das necessidades levantadas juntamente aos profissionais do hospital SOBRA-PAR em uma abordagem etnográfica [25], identificou-se um conjunto de necessidades básicas do domínio do problema. Essas necessidades foram utilizadas como entrada para uma sessão colaborativa de design utilizando práticas do *Design Socialmente Consciente* (DSC) [6], tais como prospecção de valores das partes interessadas e *Brainstorming* colaborativo.

Como resultado, foi proposto o jogo Fonopets: um conjunto de estratégias de jogos diferentes porém compartilhando um mesmo universo de temática animal. O jogo possui a voz como modo de interação, exigindo que o usuário pronuncie palavras ou fonemas específicos múltiplas vezes para jogar. O Fonopets reúne um conjunto de fases com inspiração nos jogos *Flappy Bird* [31], *Super Mario*, *Frogger* [21] e *Yousician* [43]. No jogo concebido, o personagem se move constantemente e cada palavra pronunciada corretamente faz com que a gravidade seja invertida. No escopo deste trabalho, foi implementado um protótipo contendo um jogo inspirado no *Super Mario*, com 7 fases diferentes com 4 blocos especiais que alteram o comportamento do jogo quando “encostados” pelo personagem.

Para a avaliação do protótipo do Fonopets, foi realizada uma sessão de avaliação com 12 pessoas que atuam na área de informática. Todos os avaliadores manifestaram sentimentos positivos em relação ao jogo na maioria das fases. No entanto, todos os avaliadores relataram alguma dificuldade com o reconhecedor de fala. Não foi detectada nenhuma relação entre a pontuação obtida no jogo com o fato das pessoas terem familiaridade com jogos ou terem jogado *Super Mario* anteriormente. Este aspecto é positivo, pois sugere que não há necessidade de experiência prévia para jogar. Portanto, o jogo apresenta indícios de ser uma alternativa para a realização de exercícios de fonoaudiologia, por utilizar da voz para realizar desafios e por levantar sentimentos positivos em relação ao jogo.

O restante desta monografia está organizado da seguinte forma: o Capítulo 2 apresenta uma síntese dos trabalhos relacionados, o Capítulo 3 descreve a metodologia aplicada para a concepção do jogo Fonopets, o Capítulo 4 apresenta o protótipo do jogo e a sua avaliação, e o Capítulo 5 as conclusões do trabalho.

2 Trabalhos Relacionados

Após consulta sobre o contexto do hospital SOBRAPAR, entre os dias 16 de agosto de 2018 e 4 de setembro de 2018 foi realizada uma busca nas lojas digitais *Play Store* e *App Store* por aplicativos que se enquadram no contexto de fonoaudiologia. Esta busca teve o propósito de conhecer quais são as opções existentes para o problema de fonoaudiologia, conhecendo aspectos positivos e oportunidades de soluções.

A busca foi realizada com os seguintes termos relacionados ao contexto de fonoaudiologia: Fonoaudiologia, Audiologia, Disfagia, Terapia de Fala, *Speech Therapy* e *Articulation*, todos utilizados de modo isolado. Como critério de seleção foram observados os primeiros 40 resultados de cada pesquisa e que eram relacionados ao tratamento fonoaudiológico para crianças – ver Tabela 2.1. Em ambas as lojas pesquisadas o critério do padrão de apresentação dos resultados de buscas é por “relevância”. Ao final do processo, foram retornados 240 resultados por loja – não foram retirados os aplicativos duplicados desse número. Desses, foram analisados 13 aplicativos mais relevantes na *Play Store* e 9 na *App Store*.

Alguns jogos estão disponíveis em vários idiomas e faixas de preço. O *GameFono*, por exemplo, possui versão gratuita e paga em duas linguagens diferentes, totalizando quatro versões. Como é o mesmo aplicativo, foi tratado como apenas um aplicativo nessa pesquisa. Além disso, na *App Store*, foram encontrados 10 aplicativos com preço igual ou superior a R\$ 37.90, sem versão gratuita ou para testes, o que inviabilizou uma análise detalhada. Durante a comparação, os jogos que são grátis para baixar mas possuem fases pagas ou funcionalidades extras pagas foram considerados como “freemium”, a mistura das palavras em inglês “free” com “premium”.

É importante ressaltar que a partir de um estudo bastante abrangente realizado em 2014 [14] no mercado de aplicativos fonoaudiológicos, percebeu-se uma evolução em termos de diversidade, tendo havido nesses últimos quatro anos um crescimento significativo na quantidade e qualidade de opções.

De acordo com os critérios estabelecidos acima, foram analisados os seguintes aplicativos:

Tabela 2.1: Comparação entre os aplicativos analisados.

ID	Nome	Plataforma, Downloads e Nota	Preço	Finalidade
1	Amigofono	<i>Android</i> : 5 mil, 3.7 (65) <i>iOS</i> : sem nota	<i>freemium</i>	Permitir o acompanhamento do paciente durante a semana.
2	Falar Melhor	<i>Android</i> : 500, 4.6 (11) <i>iOS</i> : sem nota	grátis	Auxiliar na realização de exercícios de ginástica articulatória e trava-línguas.
3	Tell a Story - Game to Train Speech & Logic	<i>Android</i> : 1 milhão, 3.9 (5391)	grátis	Auxiliar na lógica e desenvolvimento da fala por meio de uma sequência embaralhada de imagens, pede ao usuário para arrumar a ordem e contar uma história inspirada no contexto.

4	GameFono	<i>Android</i> : 10 mil, 4.7 (9.8) <i>iOS</i> : 4.4 (11)	<i>freemium</i>	Auxiliar na aprendizagem de fonemas, sons e imagens, a partir de exercícios visuais, sonoros e escritos.
5	LeLe Sílabas	<i>Android</i> : 100 mil, 4.6 (606) <i>iOS</i> : 5.0 (5)	grátis	Permitir que a criança aprenda a escrever e pronunciar o nome de objetos que ela já conhece através da de exercícios visuais e sonoros.
6	Falar a Brincar	<i>Android</i> : 10 mil, 4.3 (91)	grátis	Auxiliar o profissional durante a consulta por meio de exercícios que mostram uma palavra e pedem para a criança invertê-la ou modificar a sílaba inicial/final.
7	Fofuuu Kids	<i>Android</i> : 1000, 4.1 (8) <i>iOS</i> : sem nota	pago	Permitir que a criança melhore a fala a partir de jogos movidos pela voz.
8	singular/plural terapia da fala	<i>Android</i> : 100	grátis	Permitir que a criança aprenda a diferença entre singular e plural a partir de uma ou múltiplas imagens de um mesmo objeto.
9	Speech Essentials Therapy App	<i>Android</i> : 10 mil, 3.2 (159)	<i>freemium</i>	Auxiliar a aprendizagem de fonemas a partir de vários jogos menores, como jogo da memória, palavras cruzadas e “flashcards” para auxiliar no treino.
10	Articulation Essentials / Articulation Station	<i>Android</i> : 500, 1.8 (5) <i>iOS</i> : sem nota	<i>freemium</i>	Auxiliar a aprendizagem de fonemas a partir de “flashcards” altamente customizáveis. Permite selecionar o fonema, a categoria (palavras, frases ou sentenças) e se deseja treinar o fonema no começo, meio ou final da palavra.
11	Tongue twisters	<i>Android</i> : 100 mil, 4.5 (3140)	grátis	Auxiliar no desenvolvimento da fala a partir da geração de várias frases trava línguas, que auxiliam o profissional a sempre ter um exercício novo.
12	Speech Therapy Flashcards - S	<i>Android</i> : 10 mil, 4.2 (25)	grátis	Auxiliar a aprendizagem de fonemas a partir de “flashcards”
13	Visual Reading - special education	<i>Android</i> : 10 mil, 4.2 (58)	<i>freemium</i>	Auxiliar aprendizagem de fonemas e palavras através de exercícios altamente didáticos.
14	Talk Around It	<i>iOS</i> : sem nota	<i>freemium</i>	Auxiliar aprendizagem de palavras através de uma foto e várias dicas, como “primeiro som”, “primeira sílaba”, “soa assim”, “barulho” e “primeira letra”.
15	Conversation Therapy	<i>iOS</i> : sem nota	<i>freemium</i>	Auxiliar no desenvolvimento da fala através de uma imagem, com várias perguntas para o profissional fazer e poder conversar com os pacientes.
16	Foco X Cuca Fresca	<i>iOS</i> : 4.8 (10)	grátis	Auxiliar na alfabetização e no conhecimento de palavras através de um jogo em 3D com 3 fases e níveis diferentes de dificuldade.

A *Play Store*, para *Android*, mostra o número de downloads que um aplicativo teve nas seguintes categorias numéricas: 0, 1, 5, 10, 50, 100, 500, 1000, 5 mil, 10 mil, 50 mil, 100 mil, 500 mil, 1 milhão, 100 milhões, 500 milhões, 1 bilhão, 5 bilhões. Portanto, um aplicativo que recebeu “500 downloads” pode ter recebido até 999 downloads ou qualquer valor no intervalo 500-1000. A *App Store*, para *iOS*, não revela o número de downloads para nenhum aplicativo. Quanto às notas dos aplicativos, estas são globais para a *Play Store* e regionais para *iOS*, então é comum existir uma ordem de grandeza de diferença entre as notas de ambas as lojas. Para efeitos de pesquisa, foi considerada a loja brasileira para *iOS*.

Dentre os aplicativos analisados, os que mais se encaixam no escopo deste trabalho foram os aplicativos *Fofuuu Kids* [12] e *Amigofono* [3], que possuem custo bastante elevado. *Fofuuu*

Kids custa entre 288 e 1624 reais por ano, variando o número de pacientes e customizações para o profissional, e o *Amigofono* custa entre R\$ 39,99 e R\$ 69,90 por mês, por pacientes ilimitados, variando o tamanho da biblioteca de exercícios e tamanho dos vídeos que o profissional pode gravar. Ambos os aplicativos fazem uso de Internet, precisam de cadastro e periodicamente necessitam de utilização de conexão para o uso. *Fofuuu Kids* foi a única aplicação encontrada que faz uso de voz como entrada no jogo.

Existe uma clara distinção entre os aplicativos e jogos que foram projetados para profissionais da área usarem com pacientes e os que foram feitos especificamente para crianças. Os jogos projetados para profissionais muitas vezes necessitam de cadastro, Internet, possuem uma interface carregada em texto, e uma disposição dos menus e opções que não é fácil de entender. As Figuras 2.1 e 2.2 demonstram essa diferença – a Figura 2.1 demonstra a simplicidade do menu de um jogo projetado para crianças, enquanto a Figura 2.2 demonstra o menu de um jogo projetado para profissionais.



Figura 2.1: Menu do aplicativo Lele Sílabas.

Em vários jogos, o critério de pontuação depende do profissional e não do próprio jogo, logo o usuário não possui autonomia para jogar sozinho, sempre dependendo da ajuda e avaliação de outra pessoa.

Os aplicativos e jogos projetados para crianças são o oposto: extremamente simples, possuem menus grandes, coloridos, muitas vezes auxiliados por ícones, de forma a favorecer que até os menos alfabetizados consigam usar. São fáceis de entender e usar, possuem critério de pontuação objetivo, e não fazem uso da Internet.

Foi possível também perceber que boa parte das soluções existentes possuem grupos de características bem distintos. Estes grupos foram nomeados em categorias, descritas abaixo:

Flashcard: Estilo de aplicação em que o usuário seleciona o que deseja aprender, como um fonema em específico, uma categoria (palavras, frases ou sentenças) e a posição do fonema nas palavras. Por conta dessa diversidade de opções, é mais indicado para profissionais que estejam atendendo crianças e precisem de exercícios extremamente específicos. Também pode ser usado por pais ou responsáveis, sabendo da necessidade de tratamento dos filhos, para treinar os exercícios regularmente. Esse tipo de aplicação se encaixa menos como jogo e mais como ferramenta de auxílio à aprendizagem. Nessa categoria, o melhor exemplo encontrado foi o *Articulation Essentials* [26].

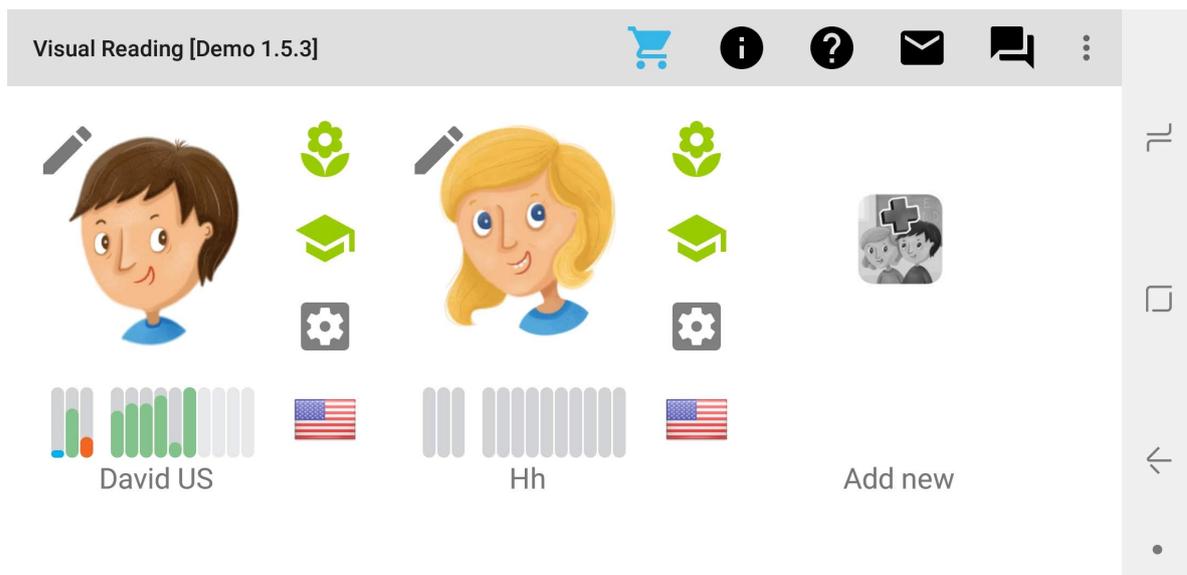


Figura 2.2: Menu do aplicativo Visual Reading.

Jogos que dependem de uma pessoa externa para garantir a pontuação adequada: Muitos dos jogos encontrados não possuem critério de pontuação adequado. São jogos que, por exemplo, permitem que o próprio usuário diga se acertou ou errou. Embora jogos como “Falar e Brincar” [20] possuam exercícios desafiadores e bem planejados, não há como saber se o usuário fez, e se acertou ou errou, sem o auxílio de uma pessoa externa.

Jogo que apresenta a imagem de um objeto, ou pronuncia uma palavra, e aguarda a identificação das sílabas pelo usuário: Nesta categoria se encaixam os jogos que instigam as crianças a passar um tempo jogando, e fazem com que elas exercitem a mente. São bem variados, e diferente da categoria anterior, a maioria não requer Internet, nem requer atenção de um adulto. O jogo também é capaz de pontuar corretamente, e são feitos para serem jogados repetidamente. Exemplos: *LeLe Sílabas* [29], *Visual Reading* [41], *Game Fono* [1].

Jogo movido ao som que a pessoa emite: Nesta categoria, foi encontrado apenas o *Fofuuu Kids* [12] e *Fofuuu Pro*, que são apenas diferentes em relação a quem vai usar (profissional ou criança). Eles contém 10 jogos com 82 fases no total, que podem ser aplicados em vários fonemas diferentes. Estes dois, para o contexto do problema do hospital SOBRAPAR, são os mais adequados devido à mecânica principal dos jogos ser baseada na fala do usuário. Dentre o escopo pesquisado, não há nenhuma solução gratuita. A própria solução oferecida pelo *Fofuuu*, em teste realizado pelo autor, possui problemas que podem ocasionar em várias dificuldades para os jogadores. Por exemplo, o reconhecedor de fala apresenta erros na precisão de acerto de falas: em uma das partidas realizadas pelo autor, o som que o jogo emitia era entendido pelo próprio jogo como a voz do usuário, fazendo com que movimentos fossem registrados involuntariamente ou o jogo entrasse em *loop* e continuasse pontuando mesmo com o jogador em silêncio. A interface do jogo é difícil de utilizar, com elementos pequenos e erros na mudança de orientação do celular.

2.1 Análise baseada em Valores

Utilizando como base a lista de valores em software social apresentada no artigo de Pereira et al. [35], cada aplicativo foi analisado pelo autor deste trabalho, e foram elencados abaixo os valores que mais se encaixam no escopo dos aplicativos avaliados e da solução

proposta. Embora considerados importantes para o contexto trabalhado, os valores de Identidade, Colaboração, Compartilhamento, Segurança, Personalização e Adaptação, Disponibilidade, Escalabilidade e Awareness não foram identificados ou não estão contemplados adequadamente nos aplicativos encontrados.

Identidade: possibilidade do usuário expressar a sua própria personalidade, como por meio de um personagem único [35]. Nenhuma aplicação pesquisada permite ao usuário se expressar ou se representar por meio de um perfil ou avatar. O *Amigofono* [3], no entanto, permite ao profissional criar um personagem único para ser representado dentro da aplicação. Para o Fonopets, é relevante que o usuário possa criar e personalizar um personagem, refletindo aspectos de sua identidade (e.g., nome, foto, dados). Isso pode aumentar o engajamento e a afinidade do usuário com o jogo, possibilitando, por exemplo, que os exercícios sejam feitos com maior frequência.

Colaboração: possibilidade de trabalhar com outros usuários para resolver um mesmo problema [35]. Entre todas as aplicações encontradas, apenas a *Conversation Therapy (iOS)* permite algum tipo de colaboração: organizar grupos de até quatro pessoas para fazerem o exercício ao vivo. A avaliação, no entanto, ocorre de modo individual para cada pessoa. O valor da colaboração passa a ser relevante ao considerar que a criança pode se sentir menos solitária ao fazer uma atividade com outra criança que tenha o mesmo distúrbio de fala. Não foi planejado nenhum mecanismo de colaboração para o Fonopets.

Compartilhamento: possibilidade dos usuários disponibilizarem objetos ou informações para outros usuários. Por exemplo: publicação de fotos em uma rede social [35]. Nenhum aplicativo encontrado permite qualquer tipo de compartilhamento. Como o Fonopets está em um contexto com alta limitação no uso da Internet, a transmissão de gravações entre o dispositivo do paciente a fonoaudióloga poderá ocorrer quando existir conexão *Wi-Fi*.

Segurança: engloba o modo como os dados armazenados no aplicativo serão guardados e transmitidos, e como a informação será mantida em caso de falhas (backup) [35]. Não é possível saber como os aplicativos estão transmitindo ou armazenando os dados. No *Android*, todas as informações de um aplicativo, por padrão, ficam guardadas em um diretório protegido pelo sistema e inacessível pelo usuário. Não há necessidade, portanto, de oferecer proteção maior que essa. No entanto, a questão de segurança é importante para este trabalho pois o Fonopets deverá conversar com um servidor, proteger a transmissão e, quando houver conexão, efetuar backup do material produzido pelo usuário.

Personalização e Adaptação: adaptar elementos da interface, configurar preferências do usuário, etc [35]. Nesse quesito, nenhuma aplicação pesquisada permite essas adaptações. No entanto, no escopo da solução do Fonopets, pode ser interessante mudar o tamanho, disposição e cores dos elementos da interface dependendo da idade da criança, de forma a proporcionar a maior harmonia possível.

Disponibilidade: capacidade da aplicação estar disponível para ser usada a qualquer momento, sem interrupções [35]. Levando em conta o contexto do SOB RAPAR, em que as pessoas ali atendidas são de baixa renda, é possível classificar a disponibilidade nas aplicações com relação à Internet entre “regularmente”, “apenas na instalação” ou “não faz uso”. A solução a ser proposta não deverá necessitar de Internet para uso, mas será útil utilizá-la ocasionalmente, pelo menos poder acessá-la no hospital, para o paciente poder sincronizar tudo o que foi realizado desde a última consulta.

Escalabilidade: capacidade de uma aplicação lidar corretamente com um crescente número de usuários [35]. Exemplo: a possibilidade de fazer *download* de mais fases/exercícios quando for necessário. Apenas o Amigofono e Fofuu permitem a criação de exercícios customizados a qualquer momento, mas o jogo *LeLe Sílabas* poderia ter a mesma mecânica

ampliada para mais fases. Para o Fonopets, a escalabilidade é relevante para que o aplicativo possa ser adequado desde a crianças mais novas até adultos, no começo até o final do tratamento.

Awareness: percepção de quem está disponível no sistema, de quem está fazendo o que [35]. Exemplo: o usuário é notificado do que mudou desde o último *login*. Nesse quesito, a maioria das aplicações são estáticas e somente o *Amigofono* possui vários mecanismos de gamificação e incentiva o usuário a abrir o aplicativo múltiplas vezes por semana. É uma oportunidade ter, no Fonopets, mecanismos que permitam que o paciente seja lembrado de fazer os exercícios, como por exemplo, notificações.

Métodos de interação: este não é um conceito de software social presente na lista de valores utilizada [35], mas é um elemento importante dos aplicativos pois se refere ao modo de interação dos usuários com as aplicações. Todos os aplicativos encontrados fazem uso da tela *touch*, por exemplo para interação com os menus. Quanto à funcionalidade principal, alguns são acionados por meio de toque de botões, outros arrastando-se objetos, mas apenas o *Fofuuu* usa a voz como método de entrada para os jogos: o usuário precisa emitir um som, ou fonema, diversas vezes, adequadamente, para ganhar mais pontos em diversos mini jogos. Isso é relevante para o contexto do Fonopets, pois o método de entrada para será a fala.

Os valores restantes abaixo, também baseados no artigo de Pereira et al. [35], apresentaram maior presença e variação nos aplicativos.

Estética: características relacionadas à beleza, harmonia visual, tamanho dos textos, imagens e ícones, além de botões [35]. Na Tabela 2.2, por análise e entendimento do autor, os aplicativos que possuem essas características foram classificados como “agradável” e os que não possuem como “desagradável”. É importante que o Fonopets tenha uma aparência agradável, pois até mesmo a harmonia das cores podem causar emoções diferentes em pessoas diferentes [30].

Autonomia: capacidade do usuário de decidir, planejar e agir de forma a atingir os próprios objetivos [35]. Nesse contexto, a aplicação deve ser intuitiva e o usuário deve entender o que é necessário fazer e como proceder, sem o auxílio de outra pessoa. Na Tabela 2.2, os aplicativos que permitem autonomia ao usuário foram classificados como “alta”, enquanto os que não são intuitivos ou precisam de uma segunda pessoa para ajudar, por exemplo, na pontuação, foram classificados como “baixa”. O Fonopets deve ser intuitivo e permitir que a criança pratique de forma independente.

Meta Communication: ferramentas que permitem ao aplicativo ser auto-explicativo para usuários iniciais [35], mostrando, por exemplo, tutorial ou ajudando o usuário inicial a navegar e entender a aplicação.

Privacidade: os usuários devem poder escolher quais informações a respeito deles estará disponível e quem terá acesso às informações [35]. Dos aplicativos pesquisados, alguns pedem acesso aos dados e recursos do celular. Muitas vezes assim que a primeira tela é aberta, é requisitado pelo aplicativo o acesso ao recurso de localização do usuário sem informar o motivo ou o que será feito com essa informação. Alguns aplicativos também obrigam o usuário a fazer um cadastro, o que permite recolher ainda mais informações. Na Tabela 2.2, foram classificados em “alta” os aplicativos que não fazem uso de informações privadas ou desnecessárias, como a localização, e “baixa” os aplicativos que oferecem pouco controle sobre isso ao usuário.

Os conceitos de Estética, Autonomia, *Meta Communication*, e Privacidade foram colocados em uma Tabela 2.2 a fim facilitar a comparação. Como elementos restantes da Tabela, tem-se o nome das aplicações, bem como o público a quem se destina.

Tabela 2.2: Comparação usando análise de valores sociais em cada aplicativo encontrado.

Nome	Público	Estética	Autonomia	Meta Communication	Privacidade
Amigofono	Profissional e paciente	Agradável	Alta	Alta	Alta
Falar Melhor	Criança	Desagradável	Alta	Alta	Baixa
Tell a Story - Game to Train Speech & Logic	Criança e familiar	Agradável	Baixa	Alta	Alta
GameFono	Criança	Agradável	Alta	Alta	Alta
LeLe Sílabas	Criança	Agradável	Alta	Alta	Alta
Falar a Brincar	Profissional	Desagradável	Baixa	Baixa	Alta
Fofuuu Pro/Kids	Profissional e criança	Desagradável	Alta	Média	Baixa
singular/plural terapia da fala	Profissional e familiar	Desagradável	Baixa	Baixa	Alta
Speech Essentials Therapy App	Profissional, criança e familiar	Agradável	Alta	Alta	Baixa
Articulation Essentials / Articulation Station	Profissional, criança e familiar	Agradável	Alta	Média	Alta
Tongue twisters	Profissional e familiar	Desagradável	Baixa	Baixa	Baixa
Speech Therapy Flashcards - S	Profissional, criança e familiar	Agradável	Baixa	Alta	Alta
Visual Reading - special education	Profissional, criança e familiar	Desagradável	Baixa	Média	Baixa
Talk Around It	Profissional, criança e familiar	Agradável	Alta	Alta	Baixa
Conversation Therapy	Profissional, familiar	Desagradável	Alta	Alta	Baixa
Foco X Cuca Fresca	Criança	Agradável	Alta	Alta	Alta

A partir da análise entre os aplicativos existentes na área e das necessidades apresentadas em conversa com duas fonoaudiólogas do Hospital SOBRAPAR [25], foram elaborados os seguintes requisitos que o aplicativo a ser proposto nesse trabalho deve ter.

- Funcionar em dispositivos *Android*.
- Não precisar de Internet para usar.
- Poder salvar o progresso da pessoa localmente.
- Permitir sincronização quando conectado em rede *Wi-Fi*.
- Ter uma estética agradável.
- Permitir o mínimo de customização, de forma a cada usuário pelo menos poder modificar as cores do aplicativo, ou criar um próprio personagem.
- Respeitar a privacidade.

- Possuir mecanismos de gamificação que incentive o usuário a usar o aplicativo regularmente.
- Permitir que o fonoaudiólogo crie ou customize exercícios para treinamento.
- Poder gravar o treino das crianças, idealmente na forma de vídeo, de modo que o movimento da boca possa ser visto, para que as crianças possam se ouvir e o fonoaudiólogo garantir que o treinamento está sendo feito ao ouvir a gravação.
- Dinâmica de jogo deve ser fluída.
- Aplicativo deve não exigir alta capacidade de processamento e armazenamento.
- Aplicativo deve funcionar em celulares posteriores à 2012.

Como resultado da análise, foi possível identificar que nenhum aplicativo existente é capaz de atender às demanda do contexto aplicado deste trabalho. Deste modo, com base nesses requisitos identificados foi realizada uma sessão de design colaborativo, descrita no Capítulo 3. Nessa sessão, cada parte interessada, requisito, problemas e possíveis soluções foram elencados e direcionados para a proposição da nova solução (i.e., o Fonopets). No Capítulo 4, relata-se como foi criado e avaliado um protótipo dessa solução.

3 Metodologia

Com o objetivo de propor um aplicativo no contexto do tratamento de fonoaudiologia para crianças pelo Hospital SOBRAPAR, foi realizada uma sessão com processo de design de acordo com os princípios do Design Socialmente Consciente (DSC) [6], em que os aspectos técnicos de uma solução de design dependem e afetam os aspectos formais e informais das organizações e da sociedade.

A sessão de design ocorreu no dia 11/10/2018, com duração de 4 horas e 20 minutos (entre as 08:50 e 13:10) no Auditório 2 do Departamento de Informática do Centro Politécnico, UFPR, Curitiba. Participaram no processo um estudante de graduação, quatro estudantes de mestrado, um mestre, três estudantes de doutorado e um professor da área de IHC. Com exceção do mestre, todos os outros integrantes fazem parte do Departamento de Informática da UFPR.

As atividades foram realizadas para apoiar o entendimento e a clarificação do problema, bem como ideação de uma solução. Sob os princípios do DSC, o processo de design de uma solução computacional deve envolver a perspectiva de diferentes partes interessadas, nos níveis informal, formal e técnico, favorecendo a identificação, articulação e formalização de aspectos relevantes do mundo social, possivelmente fazendo sentido para os usuários, pois seu design tende a refletir uma compreensão de seu mundo social.

Para a realização da sessão foram utilizados artefatos inspirados na Semiótica Organizacional e Design Participativo: Diagrama de Partes Interessadas (DPI); Quadro de Avaliação e Escada Semiótica. Para a ideação colaborativa, utilizou-se das técnicas de *Canvas Value* e *BrainWriting*. Ao final da sessão, os artefatos preenchidos (DPI, Quadro de Avaliação e Escada Semiótica) estavam preenchidos possibilitando identificar requisitos, e foram elaborados os cenários de elementos de um jogo, pensado colaborativamente, que abordasse atividades de fonoaudiologia, como exercícios, relatórios e avaliações.

A sessão seguinte apresenta o processo de design realizado, com descrição cronológica das atividades.

3.1 Relato da Sessão

A sessão teve início às 08:50 com a apresentação do tema de DSC, trazendo seu processo histórico desde a década de 70 com o Design Participativo e a teoria da Semiótica Organizacional, até Design Centrado no Usuário e Design Universal, importantes para o surgimento do DSC.

Após a introdução do tema, a teoria da Semiótica Organizacional [24], bem como os conceitos de signo, objeto e interpretante. Adentrou-se no conceito de normas, relacionando-as com os níveis informal, formal e técnico de Hall [15]. Os artefatos da semiótica organizacional foram discutidos, como a Cebola Semiótica e o DPI, na perspectiva do *Problem Articulation Method*, *Semantic Articulation Method* e *Norm Articulation Method*.

Com o entendimento das bases conceituais e dos artefatos a serem utilizados, iniciou-se (10:00) a apresentação do contexto do trabalho: o Hospital SOBRAPAR, os trabalhos existentes,

o tipo de atividade desenvolvida no hospital, a população que frequenta e as necessidades de fonoaudiologia. O conjunto de requisitos estabelecidos no final do Capítulo 2 foi então apresentado.

O preenchimento dos artefatos de maneira colaborativa foi iniciado (10:30) partindo do artefato DPI (ver Figura 3.1) para mapear os *stakeholders*, ou partes interessadas, que influenciavam ou eram influenciadas pelo projeto da solução computacional. Com as partes interessadas identificadas, adentrou-se (10:50) no preenchimento do Quadro de Avaliação (ver Tabela 3.1) com o propósito de antecipar os problemas que cada parte interessada no problema ou na sua solução poderia trazer para o contexto do design, assim como as possíveis ideias para se lidar com cada problema.

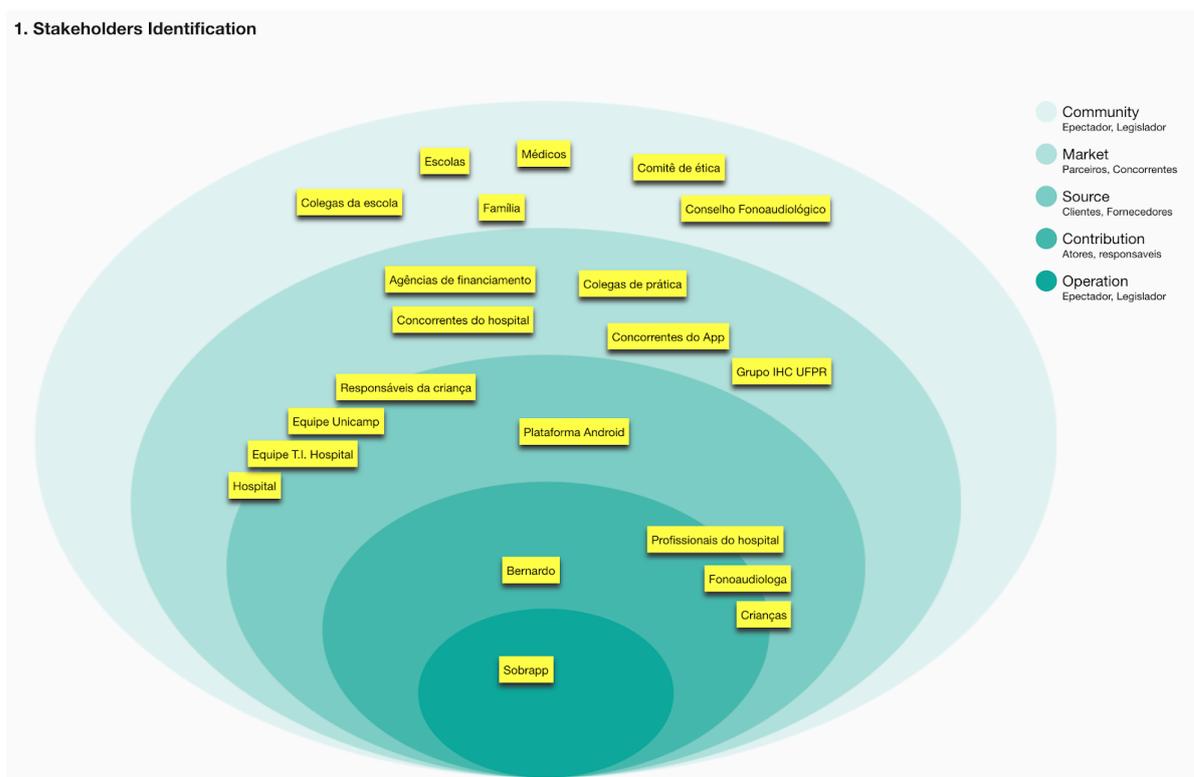


Figura 3.1: Identificação dos Stakeholders.

Após a clarificação das partes interessadas, os problemas e soluções (11:30), o *Canvas Value* foi preenchido, definindo os *stakeholders* mais relevantes para o projeto: a Criança (1^o *Stakeholder*) e Fonoaudióloga (2^a *Stakeholder*). Também foi destacado o valor agregado da solução, como a melhoria na habilidade da fala e a melhora na socialização com a família. O *Canvas Value* também possibilitou identificar os diferenciais que a solução proposta deve ter, como o armazenamento de dados, a criança poder propor um desafio, e a fonoaudióloga poder adaptar os exercícios.

Com a ideia da solução definida (11:50), o preenchimento da Escada Semiótica (ver Tabela 3.2) foi iniciado para se listar requisitos para a aplicação, considerando desde requisitos de infraestrutura técnica até requisitos do ambiente de uso e seus possíveis efeitos nas partes interessadas.

Com os requisitos listados (12:19), uma sessão de *BrainWriting* colaborativo foi conduzida com os mesmos participantes (ver Figura 3.2). Às 13:10, a atividade de *BrainWriting* foi finalizada, fez-se os agradecimentos aos participantes voluntários e encerrou-se a sessão.

Stakeholder	Problemas	Soluções
Bernardo	1 - O que é um bom benefício/recompensa? 2 - Avaliação do Processo? 3 - Como gamificar? 4 - Como implementar as práticas? 5 - Como deve ser o jogo?	1 - Deixar a fonoaudióloga decidir sobre a recompensa. 1 - Ganhar recompensas fora do aplicativo. 2 - Comparação entre desempenho. Atividades devem ser um modo de interação.
Fonoaudióloga	Praticidade	Desenvolvimento parte criança x Desenvolvimento parte fono. Praticidade; Poder escolher as atividades adequadas para cada criança.
Crianças	1 - Problema disponibilidade local x <i>Wi-Fi</i> . 2 - Criança não gostar do jogo, não ser interessante. 3 - Ter muitos exercícios realizados para subir no Drive. 4 - Motivação. Jogo ser natural para não sentir vergonha no uso.	Sincronização Assíncrona. Personalização de Criança para Criança; Personagem caracterizável. Atividades Colaborativas. Engajar Colegas/Família. App só subir as informações de Progresso quando conectado na <i>Wi-Fi</i> .
Profissionais do hospital	Médico tem relação próxima com criança/família. Dentistas, Médicos e Pedagogas têm interesse.	
Equipe Unicamp	Coleta de dados/informações. Interesse na aplicação	
Responsáveis da criança	1 - Como os pais vão acompanhar? 2 - Qual é a motivação para o responsável? 3 - Como não deixar pai/mãe esquecer das atividades?	1 - Mandar mensagem para responsável dizendo que criança não fez atividade. Engajar naturalmente pelo jogo; exemplo: pet está com fome. Apelo às crianças usando super heróis; resolver charada por exercícios fonoaudiológicos. Pessoa com maior relação com a criança é a mãe. Mãe sustenta casa e precisa acompanhar criança no ambiente hospitalar.
Equipe T.I. Hospital	1 -Infraestrutura para sincronização 2 - Infraestrutura para segurança 3. Transmissão e armazenamento	2. Criptografia
Concorrentes do hospital	Interesse na aplicação.	
Médicos	Acompanhamento da evolução	Feedback reativos. Dividir tipos de fases por falar, ler e ouvir.

Tabela 3.1: Problemas e soluções encontrados para cada parte interessada.

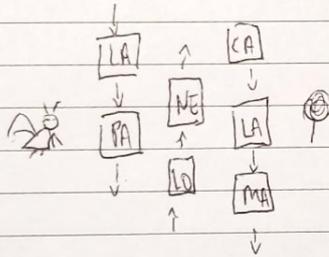
MATEOS

o Sobrapp é:

- A criança (paciente) ajuda os super-heróis a resolver problemas com sua voz.
- Os problemas são relacionados a enfrentar vilões e situações sociais.
- Ex: O super lobo precisa apagar um incêndio e a criança emite um som (shh) para ajudá-lo.
- Outra atividade pode ser ajudar o super lobo a cruzar um rio, ^{o ele está com fome.} para isso a criança precisa construir uma ponte. A ponte vai ser formada pelos silabos das palavras.



→ TAMBÉM PODE SER ADAPTADO O JOGO DO SAPO QUERENDO ATRAVESSAR O TRÂNSITO, MAS COM FONEMAS EM CADA CARRO



ALÉM DISSO, O SUPER LOBO PODE SER AUXILIADO POR outros pets com habilidades especiais, como a Patrôncia, ~~que~~ ~~constrói~~ uma vaca que amamenta o lobo com sua cauda potente. Os bichos devem ter nomes escolhidos pela criança e também posturi.

o jogador pode ir agrupando e montando seu grupo de super-heróis para enfrentar "N" fases de dificuldade dos vilões. Em alguns casos 2 ou mais ~~animais~~ animais podem se combinar para formar super poderes, (super silabas) como:

o Leão e a Hiena. o Sapo e a Cobra.
(LH) (SC) ...

Figura 3.2: Uma das folhas preenchidas durante o processo de BrainWriting.

Escada	Requisitos
Social World	Melhora na qualidade de vida da criança. Melhora na socialização da criança com o mundo. (preenchido posteriormente)
Pragmatic	Promover exercício da fala. Promover engajamento da criança. Promover exercícios fora do ambiente hospitalar. Temática Sobrapets: manter coerência; já está sendo utilizado no hospital; pode promover engajamento. Convívio familiar: promover socialização entre participantes da família. Engajamento e visibilidade. Excluvader: quer promover exclusão, pois é um canguru/T-rex que não consegue abraçar alguém.
Semantic	O que significa cada exercício? O que significam as funcionalidades?
Syntactic Empirical	Linguagem Android. Material Design. Padrão Sobrapets? Capacidade de Armazenamento: Nuvem, Drive (15GB), Dropbox (2GB), Backup. Uma vez feito backup da informação, liberar espaço.
Physical	Android. Smartphone e Tablet. Internet, <i>Wi-Fi</i> . SMS. SMTP (E-mail).

Tabela 3.2: Tabela do framework semiótico

3.2 Quadro de Prospecção de Valor

Na atividade de Prospecção de Valor, foi definido quem são as pessoas mais importantes para a solução a ser projetada, quais as necessidades dessas pessoas, quais dificuldades poderiam surgir com/para a solução, quais as contribuições da solução e quais as possibilidades de inovação (ver Figura 3.3). Como pessoa mais relevante (foco da solução), foi escolhido o paciente (especialmente a criança), e em segundo a fonoaudióloga. Embora ambos sejam importantes para a solução, o motivo da prioridade maior para o paciente foi o fato de que é ele quem deverá usar o aplicativo diariamente e ter o maior benefício com seu uso.

Em seguida, foram elencadas as necessidades das crianças/pacientes. Por exemplo, é necessário que as atividades no aplicativo sejam interessantes, façam sentido, sejam divertidas, fáceis de usar, engajadoras e desafiadoras. Também foi elencado que as atividades devem proporcionar sentimento de progresso.

A partir disso, foram especificadas as dificuldades que podem surgir para essas pessoas. Como a dinâmica de jogo é baseada em voz, é possível que algumas crianças tenham vergonha de falar ou possuam falta de “consciência” oral, o que pode prejudicar a realização de atividades orais. Além disso, pode ocorrer da família não incentivar ou não engajar-se no tratamento. Há a possibilidade, por exemplo, da família só possuir um aparelho celular e a criança querer usar o celular enquanto outra pessoa também quer. Portanto, a disponibilidade e habilidade no uso do dispositivo são pontos que podem ser difíceis, dependendo da família.

Com relação a aspectos que a solução pode melhorar na vida dessas pessoas, existe o desenvolvimento ou melhoria na capacidade de fala, o aumento de socialização na família e a promoção do círculo social. Quanto melhor a criança for capaz de falar, mais fácil ela ficará inserida na sociedade. O Fonopets também possui potencial de promover a interação colaborativa na família, engajando-os no jogo e no treinamento.

Finalmente, a solução pode surpreender e inovar em quesitos como a possibilidade da fonoaudióloga adaptar exercícios para cada paciente. A criança também pode vir a propor ou adaptar desafios no jogo. Além disso, o aplicativo pode ser cooperativo, para ser usado junto com outras crianças. A ideia de armazenar os dados no celular e transferir quando o paciente estiver na consulta é algo que nenhuma outra solução encontrada faz.

Prospecção de Valor

Considere que sua equipe já escolheu uma solução para desenvolver

Stakeholder central / Parte Interessada:
Quem são as pessoas mais importantes para a solução que a equipe projetará?

1. Criança / Paciente
2. Fonoaudióloga

O que essas pessoas realmente precisam?

→ Atividades que sejam interativas, sejam sentidas, divertidas
→ Fáceis de usar, integradas, desafiadas
→ Atividades que proporcionem progresso

Quais as dificuldades que essas pessoas enfrentam?
Quais medos, desafios, problemas que essas pessoas enfrentam no dia-a-dia?

→ Vergonha de falar, falta de "consciência"
→ Vergonha de falar em público, falta de escuta
→ Engajamento da família;
→ Disponibilidade e habilidade no uso do dispositivo...

Em quais aspectos a solução pode melhorar a vida dessas pessoas?
Na interação em suas casas e comunidades? No uso de seu tempo ou espaço? No aprendizado e na diversão? Na qualidade de vida? No exercício da cidadania?

- Desenvolvimento Melhorar das habilidades de fala
- Habilidades Sociais
- Socialização na família
→ Promover o círculo social

Como a solução pode surpreender e inovar?
Qual o "óbvio" que a solução deve evitar?

→ Não pode substituir os exercícios
→ Cooperativo, entre crianças
→ Crianças podem gerar um desafio
→ Anosamento de dados

Figura 3.3: Quadro projetado durante a sessão de DSC.

3.3 Pontos em comum

Para desenvolver de forma colaborativa as ideias levantadas, a sessão de *BrainWriting* foi conduzida para produzir propostas para a interface e interação no jogo. Na sessão, cada pessoa começava a escrever a descrição de um aplicativo ou jogo em uma folha e, após um tempo determinado (e.g., 1-2 minutos), cada pessoa passava a folha para a pessoa ao lado que continuaria a descrição e complementaria a ideia. A atividade se encerrou após cada pessoa ter comentado em quatro outras folhas. Finalizado o *BrainWriting*, foram consolidados os pontos em comum que foram escritos, gerando sugestões para o Fonopets:

- Temática de animais.
- Sons específicos para animais específicos (exemplo: shhh para cobra).
- Animais podem ter habilidades personalizadas/únicas.
- Possuir chefões: algumas fases mais difíceis (exemplo: super sílabas).
- Selecionar os gêneros dos personagens.
- Cenário (floresta, futebol, deserto, hospital).
- Desafio nos cenários, podendo ser medieval, Deserto do Saara, Floresta Negra e temáticas especiais, como dia do índio.
- Progresso no jogo pode alimentar o Pet, deixá-lo feliz, etc.

- Dinâmica envolvendo o hospital, exemplo, palavras escondidas nele.
- Várias opções de jogos para não ser repetitivo para criança. Os sons foram “capturados por um personagem” e agora é preciso libertá-los.
- Criança pode interagir com outras crianças.
- Fonoaudióloga pode customizar as atividades.

Parte destes pontos foram adaptados para a proposta deste trabalho, e tem potencial de ser utilizados em outras propostas.

3.4 Solução proposta

A partir dos pontos em comum listados anteriormente como resultado da sessão, o autor propôs um jogo com temática animal. Nessa sessão, será descrito como o jogo foi pensado e no Capítulo 4 como foi implementado um protótipo dele: o Fonopets. É necessário ressaltar que no Fonopets só foi implementado a segunda modalidade proposta, o jogo do *Super Mario*. Além disso, o Fonopets não possui os desafios de filmar a própria movimentação da boca, nem opções de customização – apenas a palavra a ser falada pode ser escolhida.

Na tela inicial do jogo proposto há um animal que fica feliz quando o usuário faz os exercícios diários e triste quando o usuário não os faz. O usuário pode selecionar qual animal ele gosta mais, permitindo assim um nível de identidade entre os dois. Também há, na tela inicial, uma barra de progresso que mostra quantos pontos o usuário precisa obter por dia (ver Figura 3.4). Essa meta pode ser customizada pela fonoaudióloga nas consultas de forma a adaptar corretamente para diferentes usuários. Ao completar a meta diária, o jogo incentiva o usuário a fazer um desafio: filmar o próprio movimento da boca pronunciando uma sílaba ou palavra difícil, em um vídeo de três, no máximo cinco segundos. Esse vídeo ficará armazenado no celular para ser transmitido quando existir uma conexão *Wi-Fi*. Caso não haja conexão em casa, o vídeo poderá ser transmitido durante a consulta no Hospital SOBRAPAR, que possui Internet. Existem vários métodos possíveis para o armazenamento e transmissão; foi pensado no uso de uma pasta compartilhada do *Google Drive*, de modo que a fonoaudióloga possa acessar e averiguar o progresso que ocorreu desde a última consulta.

Visto haver necessidade da criança usar o aplicativo diariamente, dificilmente uma única solução seria capaz de não se tornar repetitiva, previsível ou cansativa com o tempo. Inspirado na ideia do jogo *WarioWare*, que agrupa dezenas de jogos com mecânica simples, mas que têm alto valor de repetitividade e engajamento, foram propostos, portanto, quatro modalidades diferentes para o Fonopets. Cada modalidade possui um animal diferente como personagem representado: pássaro, coelho, sapo e cobra.

A primeira modalidade é baseada no jogo *Flappy Bird* [31] (ver Figura 3.5), no qual um pássaro precisa atravessar diversos canos com diversas alturas. Os canos se movem horizontalmente com velocidade constante, enquanto o pássaro verticalmente, com a força da gravidade. Diferente do jogo original, a mecânica será através da voz, não do toque. Por causa disso, para evitar que o jogo se torne punitivo, a força da gravidade deve ser reduzida e a velocidade com que os canos se movem também.

A segunda modalidade é baseada no jogo *Super Mario* (ver Figura 3.5). Um estudo [22] revelou que jogar *Super Mario* por pelo menos trinta minutos por dia, durante dois meses, provocou mudanças positivas na área cerebral de um grupo de pessoas, o que pode reduzir esquizofrenia e *stress* pós traumático. A proposta é que essa modalidade tenha a voz como forma

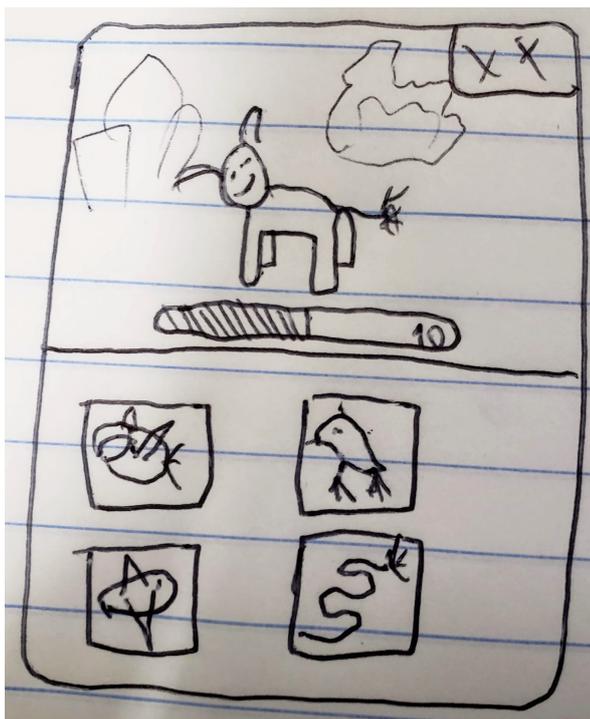
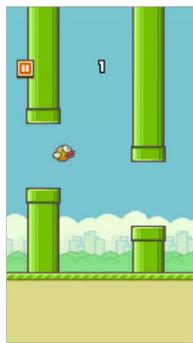


Figura 3.4: Desenho da tela inicial da solução proposta. Na parte superior, há o progresso diário (representado por uma barra) e um animal, cuja emoção é refletida pelo progresso diário. Na parte inferior, há um menu para acessar cada jogo, representado por um animal diferente.

de controle, assim o botão “A”, que no jogo original aciona o pulo. No protótipo implementado, não foi permitido ao usuário controlar o movimento horizontal do personagem, apenas se a gravidade estará para cima ou para baixo. Além disso, no protótipo foram usados blocos especiais que, ao serem encostados, geram uma situação diferente, como inverter a velocidade ou teleportar o personagem. Em uma implementação com maior número de fases ou níveis de dificuldades, poderia ser usada uma mecânica mista, com mais controles, em que o usuário pudesse controlar o movimento horizontal do personagem pelo toque na tela, além do movimento vertical pela voz.

A terceira modalidade proposta é inspirada no jogo *Frogger* [21], no qual um sapo tenta atravessar ruas e rios sem ser atropelado pelo trânsito ou se afogar na água (ver Figura 3.5). A proposta é que essa modalidade use a voz como a forma de mover o personagem para frente, verticalmente. Como alguns objetos estão em constante movimento horizontal, como troncos, o jogador precisa manter a concentração e jogar precisamente. Essa modalidade, portanto, depende menos da quantidade de vezes em que o usuário pronuncia uma palavra e mais da precisão, do momento certo.

A quarta modalidade proposta é inspirada em jogos como *Guitar Hero* [28] e *Piano Tiles* [9], e na aplicação *Yousician* [43] (ver Figura 3.5). O objetivo é que o usuário fale fonemas ou palavras no tempo certo, seguindo indicação visual na tela. Podem existir modos em que existam várias palavras diferentes e cada uma tenha que ser dita na hora certa. Também pode ser escolhida uma palavra e falada no ritmo certo.



Flappy Bird



Super Mario Bros



Frogger



Yousician

Figura 3.5: Imagem dos jogos *Flappy Bird*, *Super Mario Bros*, *Frogger* e do aplicativo *Yousician* [43], obtidas pela Internet [13, 4, 27, 42] e agrupadas pelo autor.

4 Fonopets: Protótipo e Avaliação

Este capítulo descreve o processo de desenvolvimento e avaliação de um protótipo para o aplicativo proposto na Sessão 3.4. Toda a aplicação foi desenvolvida em aproximadamente um mês e a avaliação foi realizada durante dois dias.

O aplicativo foi proposto para a plataforma *Android*, usando como base um *framework* para jogos chamado *LibGdx*. Segundo Ahmed et al.[2], embora o *LibGdx* possua complexidade elevada, a performance e facilidade de migrar para outras plataformas, como *desktop*, sem a necessidade de depender de um aparelho com *Android* ou de um emulador que seja mais lento, fazem dela mais promissora e rápida para desenvolvimento, já que é possível testar o jogo diretamente no computador.

Inicialmente, foi desenvolvido um jogo inspirado no *Flappy Bird* [31], em que um pássaro se move verticalmente e canos, com abertura entre eles, se movem horizontalmente. A missão é não encostar nos canos, ou no solo, pelo maior tempo possível; cada cano atravessado conta um ponto. Imediatamente, foi percebido que mesmo reduzindo a força de aceleração da gravidade no pássaro, ou reduzindo a velocidade com que os canos se movem, o jogo não funciona bem com a voz como método de entrada, já que requer várias repetições em curto espaço de tempo e qualquer *delay* no processamento da voz pode levar a perder o jogo.

A partir disso, foi implementado a segunda solução proposta, um jogo inspirado no *Super Mario* [40], em que o personagem se move constantemente na horizontal e, em cada interação do usuário, a gravidade (força vertical) é invertida. Foram desenvolvidas sete fases, sendo que cada uma das cinco primeiras fases introduz uma mecânica nova por meio de blocos diferenciados.

A primeira fase (Figura 4.1) introduz os blocos normais, que podem ser usados como chão, teto ou barreira, e os cristais coletáveis, que são usados como método de pontuação. A segunda fase (Figura 4.2) introduz blocos verdes que, quando encostados, invertem a direção horizontal em que o personagem está movendo. A terceira (Figura 4.3) introduz blocos roxos que, quando encostados, teleportam o personagem dez blocos para frente. A quarta (Figura 4.4) introduz blocos cinzas, que quando encostados, teleportam o personagem dez blocos para trás. A quinta (Figura 4.5) introduz blocos da cor azul claro, que quando encostados, fazem o personagem pular. A sexta e sétima fase (Figuras 4.6 e 4.7) misturam todas as mecânicas descritas anteriormente, de forma a aumentar o nível de dificuldade, já que elas não precisam mais serem introduzidas.



Figura 4.1: Primeira fase. Apresenta o universo com blocos e cristais.

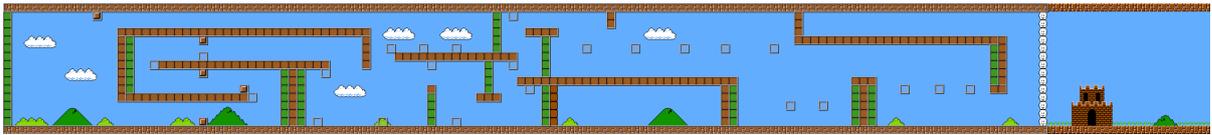


Figura 4.2: Segunda fase. Introduce bloco verde, que inverte a direção (horizontal) do movimento do personagem.

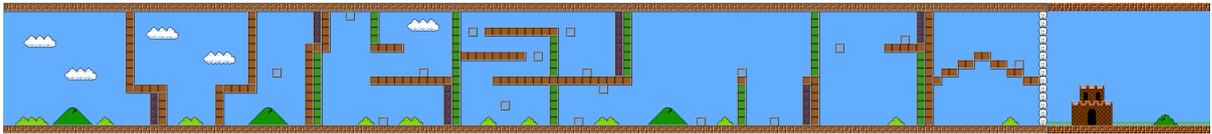


Figura 4.3: Terceira fase. Introduce bloco roxo, que teleporta o personagem 10 blocos para a frente.



Figura 4.4: Quarta fase. Introduce o bloco cinza, que teleporta o personagem 10 blocos para trás.



Figura 4.5: Quinta fase. Introduce o bloco azul claro, que pula o personagem.



Figura 4.6: Sexta fase. Mistura todas as mecânicas já apresentadas e aumenta o nível do desafio.



Figura 4.7: Sétima fase. Mistura todas as mecânicas anteriores, assim como a sexta fase.

A ideia por trás de cada mecânica nova foi de poder diversificar o jogo e ensinar ao usuário novas estratégias. De acordo com Andersen et al. [5], não há vantagens evidentes no uso de tutoriais em jogos cuja mecânica possa ser descoberta experimentalmente. Devido a isso, cada uma das cinco fases iniciais começa em um ambiente seguro, em que o usuário pode encostar no novo bloco e descobrir o que acontece, sem punição ou sendo recompensado com cristais. A partir disso, há um cenário levemente mais desafiador e, ao final da fase, um desafio maior para a pessoa provar que aprendeu a mecânica nova. Esse aumento gradual de dificuldade pode ser visto na Figura 4.8. Um jogo como *Super Mario* faz uso dessa estratégia diversas vezes, por exemplo nos aspectos:

1. Introduce uma plataforma que se move no ar, com proteção em baixo caso o personagem caia.
2. Coloca a mesma plataforma, mas sem a proteção em baixo.

3. Coloca duas plataformas juntas, sincronizadas, para o usuário pular entre elas.
4. Adiciona um obstáculo entre as duas plataformas, como um inimigo.
5. Aumenta o espaçamento entre as plataformas e faz com que elas se movam independentemente.

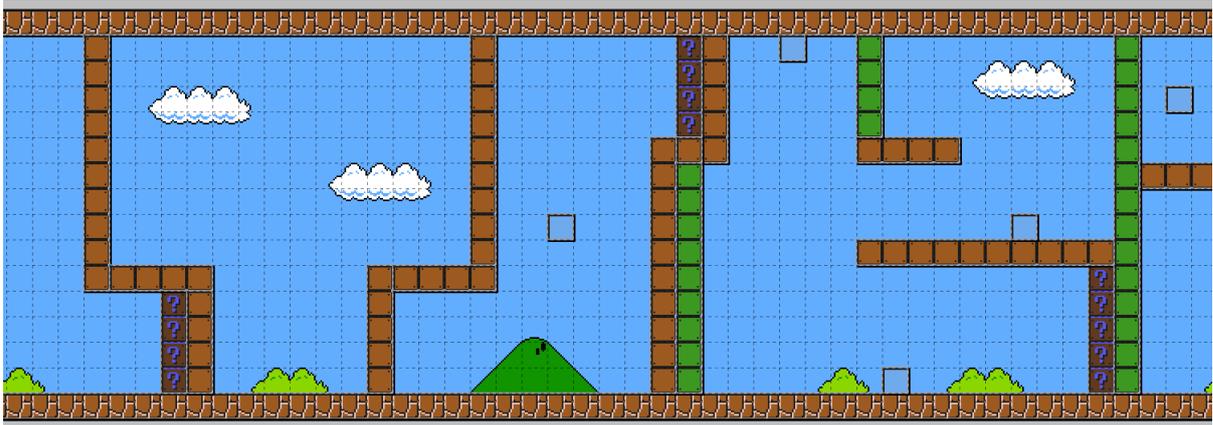


Figura 4.8: Início da fase 3 revela a introdução de uma nova mecânica e um aumento gradual na dificuldade. Blocos transparentes são renderizados como cristais.

Todas as fases foram criadas com mesmo tamanho (148 blocos de largura e 16 blocos de altura, onde cada bloco possui 16 *pixels*) e possuem duração média aproximada de um minuto. As imagens estáticas que formam o visual do jogo foram obtidas gratuitamente do site *The Spriters Resource* [40].

4.1 Reconhecimento de voz

Para reconhecimento de voz foi utilizada a biblioteca *PocketSphinx* [18]. Durante a execução deste trabalho não foi encontrada API (*Application Programming Interface*) para o sistema *Android* que permita o reconhecimento de fala sem depender de conexão com a Internet. Devido a limitação de acesso à Internet no contexto dos pacientes do SOBRAPAR, o *PocketSphinx* (escrito na linguagem C), variação do *Sphinx* (escrito na linguagem Java) focado em portabilidade e agilidade, se mostrou adequado para os requisitos do projeto.

O *Sphinx* [23] é o resultado de pesquisa da *Carnegie Mellon University*, sendo uma das poucas soluções de código aberto disponível para o reconhecimento de fala. No Brasil, existe um grupo de pesquisa chamado FalaBrasil, criado pelo Laboratório de Processamento de Sinais (LaPS) da Universidade Federal do Pará (UFPA), com foco em desenvolver recursos para sistemas de reconhecimento automático e síntese de voz para português brasileiro. O FalaBrasil é responsável por criar modelos de linguagem, scripts de treino, modelos acústicos, dicionários fonéticos e base de dados de voz e textos em PT-BR [32]. Apesar do esforço do FalaBrasil, para este trabalho o autor realizou testes com o reconhecimento de diversas palavras, como identificação de números de zero a nove, e concluiu que o grau de precisão do modelo brasileiro ainda é bastante inferior ao modelo em inglês. Esse resultado foi semelhante ao de outros trabalhos [33, 7], que identificaram o mesmo problema com a precisão. Além disso, há uma notável diferença entre o tamanho dos modelos em português e inglês: o modelo padrão do *PocketSphinx* para inglês pesa 6.6 MegaBytes. Em compensação, um modelo em português,

treinado utilizando os arquivos do FalaBrasil, pesa 23.6 MegaBytes, mais do que o triplo do modelo em inglês. Devido a diferenças de tamanho e precisão entre os modelos, para o protótipo foi utilizado o modelo em inglês. Na ideia inicial do Fonopets, o usuário deveria pronunciar fonemas isolados repetidas vezes, como “la”, “ca” e “me”. No entanto, durante o estágio de desenvolvimento foi identificado que a biblioteca de reconhecimento muitas vezes ignora ou trata como ruído uma sílaba isolada (e.g., “la”). Para contornar essa limitação, no escopo do protótipo desenvolvido para o Fonopets foram usadas palavras inteiras (e.g., “verde”).

Para treinar o modelo, o autor pronunciou diversas vezes a mesma palavra com entonações diferentes e anotou como a biblioteca registrou cada uma delas. Isso serviu para identificar possíveis variações na fala que possam ocorrer quando os usuários do Fonopets forem jogar.

Em inglês, por exemplo, no dicionário morfológico fornecido pela *Carnegie Mellon University* com mais de trinta mil definições, a palavra “round” é representada como “R AW N D” e “accessibility” como “AE K S EH S AH B IH L IH T IY”. Algumas palavras, como “authentic” e “associate” possuem quatro variações. Na prática, a biblioteca faz um modelo estatístico e compara o que foi falado com os valores existentes no dicionário para encontrar a palavra mais semelhante à que foi dita. Em português, devido a variações linguísticas, o modelo criado possui aproximadamente oito variações possíveis para cada palavra. Por exemplo, a palavra “lula” pode ser composta por “UW” ou “OW”, dependendo de como o usuário pronunciar; a palavra “azul” pode ser entendida como “Z UW”, “Z NG” ou “AW Z”. Essa estratégia usada para o jogo, portanto, só consegue reconhecer uma palavra por vez: é possível verificar se o usuário acabou de falar uma palavra específica, mas não é possível pedir para o usuário falar qualquer palavra e adivinhar qual foi. Devido a isso, não foi implementado uma forma de detecção “não fonética”: existem pessoas que ainda não conseguem emitir palavras, apenas sons. Além disso, a solução proposta não leva em conta sons de longa duração, apenas palavras que possam ser faladas rapidamente.

Na tentativa de maximizar a taxa de acertos, foram utilizadas expressões regulares, de modo que qualquer identificação entre o que foi falado e a respectiva expressão gera um acerto. A Tabela 4.1 foi projetada com todas as variações usadas no protótipo. Na seção 4.3 é avaliado como foi a experiência dos usuários com esse modelo em termos de precisão e experiência de uso durante a exploração do protótipo funcional do Fonopets.

Palavra	Expressão Regular
sabão	.*(S Z N CH F AH NG B TH).*
laranja	.*(NG AW L AY ZH TH OW AY OW AY AH ER Y DH).*
azul	.*(Z NG UW TH AH AY Z T AE AW).*
verde	.*(IY EY TH ZH V T IH JH G V EY ER).*
vermelho	.*(IY Z IH Z EY V EY G V IH V EY JH IH UW TH ER).*
caixa	.*(TH ZH AY D AH SH AW AE EY T).*
gato	.*(D EH NG CH AW F NG D AE D B EH M IH Z D).*
cavalo	.*(AW TH D AE IY K AO T D EH OW).*
violeta	.*(IY UH IH TH D Z Y OW V Y IL JH UW).*
lula	.*(UW AH N L TH IY OW Z L).*

Tabela 4.1: Tabela de expressões regulares usadas para identificar palavras em português no protótipo.

4.2 Avaliação do protótipo

Para avaliar o jogo em termos de suas funcionalidades, capacidade de reconhecimento da fala e experiência do jogador, foi conduzido um experimento para medir o número de acertos e erros no processo de reconhecimento de fala, e a adequação do jogo para o cenário proposto.

O propósito do experimento foi obter *feedback* para o protótipo construído, sendo exploradas as fases 1 à 5 do jogo. De modo a evitar que cada sessão se tornasse longa ou cansativa, as fases 6 e 7 não estiveram presentes na avaliação. A missão de cada pessoa foi jogar as cinco fases, sequencialmente, a partir da primeira. Como método de entrada, foram definidas as palavras “lula”, “gato”, “verde”, “sabão” e “caixa” para, respectivamente, cada uma das fases. O critério de seleção das palavras foi que fizessem uso de fonemas diferentes. Nenhuma delas possui a mesma sílaba, todas são completamente distintas.

Também foi realizada uma avaliação baseada no modelo *Self-Assessment Manikin* e inspirada em Hayashi et al. [16], usando cinco emoções diferentes (ver Figura 4.9) para a pessoa sumarizar o modo como se sentiu com a experiência de jogar cada uma das fases do Fonopets. Após terminar de jogar as cinco fases, a pessoa foi convidada a responder um questionário de avaliação baseado no *Technology Acceptance Model*. Esse questionário foi desenvolvido com as seguintes perguntas:

1. Qual é a sua idade?
2. Qual é seu gênero?
3. Qual é sua escolaridade? (Estudante de graduação; Estudante de mestrado; Estudante de doutorado; Outro)
4. Já fez alguma disciplina na área de IHC? (Sim, na graduação; Sim, na pós graduação; Não)
5. Eu tenho familiaridade com jogos (Concordo fortemente; Concordo; Neutro; Discordo; Discordo fortemente)
6. Qual é a sua familiaridade com o jogo Super Mario? (Altíssima: ainda hoje jogo muito; Alta: jogava bastante no passado; Média: joguei algumas vezes; Baixa: Já ouvi falar, mas nunca joguei; Nenhuma: Não sei o que é isso)
7. Quais pontos NEGATIVOS da interação com o jogo você identificou?
8. Quais pontos POSITIVOS da interação com o jogo você identificou?
9. Quais pontos NEGATIVOS da interface do jogo você identificou?
10. Quais pontos POSITIVOS da interface do jogo você identificou?
11. Deixe comentários, sugestões, dúvidas, problemas, etc.

O questionário buscava identificar o conhecimento do usuário tanto pela sua familiaridade de jogos, quanto de conhecimentos de Interação Humano-Computador, a fim de caracterizar sua experiência prévia e base em interface e interação. O questionário também verificava o sentimento do usuários sobre aspectos positivos e negativos da interação e interface do jogo, de modo a identificar aspectos gerais da experiência do usuário.

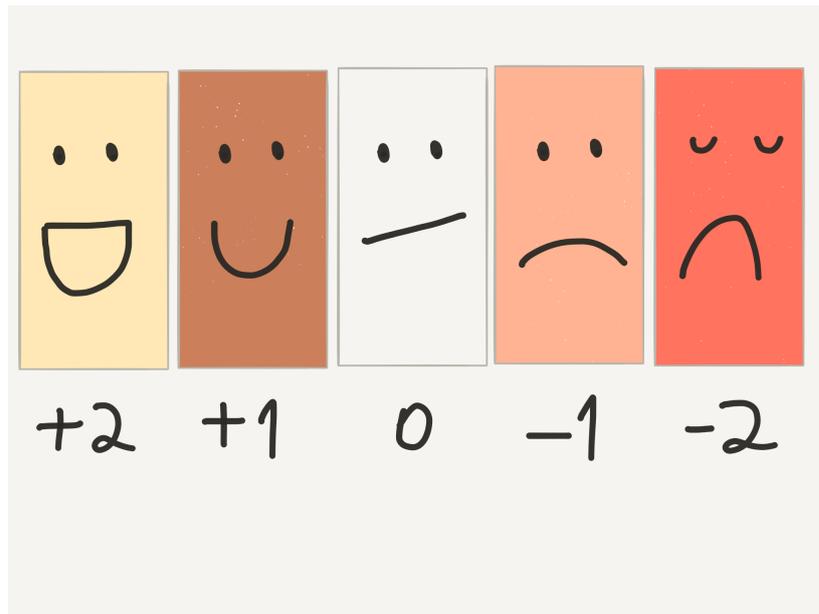


Figura 4.9: Emoções usadas para representar o sentimento após cada fase

Para 11 dos 12 participantes, a avaliação ocorreu nas dependências dos laboratórios de pós-graduação do Departamento de Informática da UFPR: oito pessoas participaram da avaliação no Laboratório de Interação Humano Computador, duas no laboratório de Inteligência Artificial e Métodos Formais, uma no Laboratório de Banco de Dados de Grande Escala. Para um participante, a avaliação ocorreu na própria residência dele. Em todos os ambientes, o silêncio foi predominante, embora em duas situações houve ruído no ambiente, o que forçou o usuário a falar mais alto.

Os participantes foram recrutados por email, redes sociais e pessoalmente. Não houve nenhum critério além da disponibilidade das pessoas. Em todas as avaliações foi utilizado o celular do autor, um Samsung Galaxy S9, enquanto o registro das emoções entre as fases e o preenchimento do formulário final pelo usuário eram realizados em um iPad.

Cada sessão começava com uma explicação do autor sobre os objetivos principais do trabalho e como a mecânica principal era a voz. Antes de iniciar cada fase, o autor dizia a palavra que a pessoa deveria pronunciar para jogar. O autor se encarregava então de iniciar a fase, em ordem crescente, da primeira até a quinta. Não foi dito nada a respeito dos cristais coletáveis, era dado a cada pessoa a liberdade e criatividade para coletá-los ou não. Ao final de cada fase, era registrado qual emoção a pessoa professou ao realizar a fase, de acordo com o modelo *Self-Assessment Manikin* adaptado para cinco emoções [16] (ver Figura 4.9). Ao final das cinco fases, a pessoa era convidada a preencher o formulário.

Não houve uma duração pré-definida para as sessões de jogo, mas elas duraram aproximadamente 10 minutos com cada pessoa. Em cada sessão foi gravado a tela do celular, usando o aplicativo *DU Recorder* [39], e o áudio, no *iPad*, usando o aplicativo *Just Press Record* [34]. A gravação foi útil para registrar, após a sessão, o tempo gasto em cada fase e a taxa de acertos de cada pessoa com o reconhecedor de fala.

4.3 Resultados

No total, doze pessoas testaram o jogo e preencheram o formulário. Dessas, oito se declararam do gênero masculino e quatro do gênero feminino. Com relação à escolaridade,

quatro pessoas disseram ser estudantes de Doutorado, cinco estudantes de Mestrado, dois de Graduação, e um professor, todos da área de Computação.

Durante a execução do experimento foi identificada limitação para o reconhecimento da palavra “gato” na segunda fase do jogo. Esta limitação transformou a fase em extremamente restritiva e o cenário utilizado se tornou muito punitivo: apenas duas pessoas conseguiram finalizar essa fase. Dependendo do momento em que a interação ocorria, o personagem era obrigado a caminhar longas distâncias até poder reverter para a direção correta. Devido a esse problema, se o autor identificava que a alta quantidade de erros (acima de 50% de palavras não identificadas) tornava a jogabilidade impraticável, o usuário era convidado a ir para a próxima fase. Apesar disso, a menor duração na segunda fase foi de trinta e seis segundos (36s) e a média foi semelhante a outras fases, aproximadamente de um minuto e catorze segundos (1:14min). Estas informações estão na Tabela 4.2.

ID	1 (lula)	2 (gato)	3 (verde)	4 (sabão)	5 (caixa)	Média
1	0:59	1:38	1:31	1:50	0:49	1:21
2	1:17	0:42	1:38	1:41	0:54	1:14
3	1:24	0:52	1:54	1:21	1:42	1:26
4	1:04	1:05	1:54	2:12	0:48	1:24
5	1:08	2:52	1:52	1:54	0:52	1:43
6	1:15	0:43	2:10	1:15	1:03	1:17
7	1:15	1:56	1:20	2:27	0:41	1:31
8	1:18	0:48	2:10	1:24	1:05	1:21
9	1:11	1:23	1:41	1:49	0:35	1:19
10	1:14	0:36	1:57	2:33	1:04	1:28
11	1:05	0:41	1:47	2:24	1:12	1:25
12	1:14	1:33	1:51	1:18	1:11	1:25
Média	1:12	1:14	1:48	1:50	0:59	1:25

Tabela 4.2: Tempo gasto em cada fase do jogo por cada usuário.

Após o término de cada fase, foi realizado um quadro inspirado no trabalho de Hayashi et al.[16] demonstrando 5 emoções, e apresentado na Figura 4.9. Questionou-se para os participantes sobre com qual emoção eles se identificaram ao jogar a fase. Os resultados detalhados estão disponíveis na Tabela 4.3 e na Tabela A.11, na seção de Apêndices.

A Tabela 4.3 contém a representação das respostas de cada pessoa para cada fase, indicando a emoção que os participantes sentiram ao jogar cada uma delas. Os valores foram representados como “:D” (muito feliz), “:)” (feliz), “:|” (neutro), “:(” (triste) e “D:” (muito triste).

ID	1 (lula)	2 (gato)	3 (verde)	4 (sabão)	5 (caixa)
1	:)	:	:	:)	:)
2	:)	:(:D	:)	:D
3	:D	:	:D	:D	:)
4	:D	:)	:D	:D	:)
5	:D	:)	:)	:	:D
6	:D	:(:D	:)	:)
7	:)	:)	:)	:)	:)
8	:D	:	:)	:)	:)
9	:)	:	:)	:)	:D
10	:)	:	:)	:D	:D
11	:)	:(:)	:	:(
12	:)	:	:)	:)	:D

Tabela 4.3: Resultado do *SAM Assesment*, representando a emoção escolhida por cada participante para cada fase.

É possível perceber que a reação dos avaliadores foi majoritariamente positiva. Mesmo na segunda fase, com dificuldades na interação, a maioria das pessoas (n=6) respondeu neutro e a quantidade de pessoas que respondeu triste (n=3) foi a mesma que respondeu feliz (n=3): algumas se sentiram extremamente desafiadas e gostaram da dificuldade extra com relação ao reconhecedor de voz, outras tiveram a reação oposta. Isso tanto indica que o protótipo foi capaz de proporcionar experiências positivas quanto revela que o jogo tem potencial de ampliação e melhoria.

A primeira fase foi a que mais fez as pessoas se sentirem bem: houve cinco pessoas que se sentiram muito feliz, enquanto as outras sete se sentiram feliz. Com exceção da segunda fase, devido aos problemas anteriormente mencionados, todas as outras fases provocaram um sentimento feliz. Não foi detectada, portanto, uma variação de emoção geral entre as primeiras e últimas fases. Apesar disso, mesmo ignorando os resultados da segunda fase, apenas um avaliador manteve as mesmas reações em todas as fases, todos os outros tiveram variações na emoção sentida. Três avaliadores tiveram uma reação mais positiva em uma fase do que em todas as outras, enquanto outros três tiveram o oposto, uma reação mais negativa.

Foi questionado no formulário sobre o nível de familiaridade com jogos e com o *Super Mario*. A intenção da pergunta era para averiguar possíveis variações no desempenho e na pontuação devido a diferença de familiaridade. Nenhuma pessoa discordou sobre ter familiaridade com jogos: seis pessoas concordaram fortemente, cinco concordaram e uma respondeu neutro. Com relação ao jogo do Mario, apenas uma pessoa respondeu nunca ter jogado, duas pessoas disseram jogar até hoje, cinco disseram ter jogado bastante no passado e quatro disseram ter jogado algumas vezes. Essas respostas indicam que os avaliadores possuem conhecimento do domínio de jogos, reduzindo as chances de haver viés pela falta de experiência do avaliador.

A partir dessas informações, com a finalidade de averiguar se existem co-relações entre familiaridade com jogos, *Super Mario* ou IHC e o desempenho obtido ou a emoção que o avaliador sentiu, foram produzidas e agrupadas as seguintes médias para as fases 1,3,4 e 5: média do tempo gasto por cada avaliador nelas, média entre o número de cristais coletados em relação à quantidade total de cristais disponíveis e média de emoção positiva que cada avaliador sentiu. As emoções foram convertidas em uma escala de -2 (muito triste) a +2 (muito feliz) e normalizadas entre 0% e 100%, onde 0% significa “muito triste” e 100% significa “muito feliz”. O resultado está presente nas Tabelas 4.4, 4.5 e 4.6

É possível perceber que a média de tempo gasto nas fases, assim como a emoção positiva que o avaliador sentiu, foram proporcionais à familiaridade com jogos e à familiaridade com *Super Mario*. Houve uma exceção, em que o avaliador respondeu ter neutra familiaridade com jogos e baixa familiaridade com *Super Mario*, mas teve desempenho maior que a maioria dos outros avaliadores. Apesar disso, para os outros 11 avaliadores, houve essa notável relação entre familiaridade, desempenho e emoção.

Também é possível perceber que avaliadores que cursaram disciplina na área de IHC tiveram a média de tempo e o número de cristais coletados inferior aos que nunca cursaram. Em compensação, ao avaliar os pontos negativos e positivos da interação, foi detectada uma variação na quantidade de pontos que cada avaliador mencionou: os que cursaram disciplina na área de IHC escreveram 36% mais pontos que as que nunca cursaram – média de 1.875 ponto por pessoa contra 1.375.

Tenho familiaridade com jogos	média cristais	média tempo	média emoção
Concordo Fortemente (6)	63.40%	1:32	82.29%
Concordo (5)	59.21%	1:23	78.75%
Neutro (1)	72.36%	1:25	87.50%

Tabela 4.4: Tabela da relação entre pessoas com familiaridade com jogos, a média de cristais coletados, a média de tempo e a média da emoções sentidas nas fases 1,3,4 e 5.

Familiaridade com <i>Super Mario</i>	média cristais	média tempo	média emoção
Altíssima (2)	60.07%	1:32	93.75%
Alta (5)	63.71%	1:28	82.50%
Média (4)	59.44%	1:24	67.88%
Baixa (1)	72.36%	1:25	87.50%

Tabela 4.5: Tabela da relação entre pessoas com familiaridade com *Super Mario*, a média de cristais coletados, a média de tempo e a média de emoções sentidas nas fases 1,3,4 e 5.

Cursou disciplina em IHC	média cristais	média tempo	média emoção
Sim, na graduação (2)	58.68%	1:19	78.13%
Sim, na pós graduação (6)	59.63%	1:25	82.29%
Não (4)	68.42%	1:35	81.25%

Tabela 4.6: Tabela da relação entre pessoas que já cursaram alguma disciplina de IHC, a média de cristais coletados, a média de tempo e a média de emoções para as fases 1,3,4 e 5.

Como pontos negativos da interação (Tabela A.3, no Apêndice A), onze das doze pessoas relataram alguma dificuldade na identificação de fala. Duas pessoas relataram que o identificador não funciona bem em ambientes com ruído. Durante os experimentos, uma pessoa chegou a espirrar e o jogo interpretou como se o jogador tivesse falado. Um avaliador também identificou que o *delay* entre a fala e o movimento do personagem é perceptível. Outro avaliador relatou que o jogo não funciona bem com sotaques diferentes. Um avaliador também relatou como negativo a introdução surpresa de novos elementos.

Para analisar a questão da precisão, a Tabela 4.7 apresenta a média de erros em relação a cada fase. Cada vez que o jogador falava, era anotado o comportamento do aplicativo com relação ao reconhecimento da fala. Caso fosse encontrado uma correspondência entre o que foi falado e as expressões regulares da Tabela 4.1, era contabilizado um acerto. Caso o jogador falasse, por exemplo, 3 vezes, e o aplicativo só registrasse uma, era contabilizado um erro por não registrar. Ainda, caso o jogador falasse mas o aplicativo interpretasse como silêncio ou barulho, também era contabilizado um erro por não registrar. Caso fosse falado e o aplicativo não encontrasse correspondência com as expressões regulares da Tabela 4.1, era contabilizado um erro por registrar errado.

	1	2	3	4	5	Média
Erro por não registrar	21.52%	21.93%	19.35%	19.35%	24.02%	24.02%
Erro por registrar errado	8.28%	33.77%	10.39%	14.22%	12.25%	15.06%
Erro total	29.80%	55.70%	29.75%	38.24%	36.27%	37.30%

Tabela 4.7: Tabela da média de erros no reconhecimento de fala em relação a cada fase.

Se considerarmos a fase número 2 como um *outlier* e a removermos da média, obtemos uma média de 22.30% de erro por não registrar a voz e 11.48% de erro por registrar errado, totalizando 33.78% de erro, ou 66.22% de acerto. No total, foram observados 891 acertos, 316 erros por não registrar a palavra e 214 erros por registrar errado. Desconsiderando a segunda fase, foram 790 acertos, 266 erros por não registrar a palavra e 137 erros por registrar errado.

Esses números indicam que, embora seja necessário melhorar a precisão do reconhecimento de voz, a maior limitação encontrada foi na dificuldade da biblioteca usada ao diferenciar palavras de silêncio.

Reconhecedores de voz existentes normalmente não são pensados para serem usados em jogos, então várias otimizações na hora de construir uma biblioteca ocorrem em pontos que podem ser diferentes do desejado no contexto de jogos. Por exemplo, se uma pessoa fala a mesma palavra duas vezes rapidamente, o *PocketSphinx* leva mais tempo para processar e interpreta como se fosse uma palavra mais longa, onde o desejado é que fosse interpretado duas palavras separadas.

Nos testes com usuários, houve pelo menos dois casos de pessoas que falavam duas vezes para “garantir” por medo, causando confusão para a biblioteca. Em outros casos, houve pelo menos três pessoas que, enquanto jogavam, produziram reações auditivas como “haha”, “que legal” e “ae!”. No entanto, a biblioteca mais uma vez entendeu como uma palavra ou frase maior e aumentou o tempo de resposta, conseqüentemente prejudicando a jogabilidade.

Também foi notado por uma pessoa como ponto negativo a ausência de uma trilha sonora. Diversos jogos e filmes são marcados por possuírem uma boa trilha sonora ou efeitos sonoros. O Super Mario, por exemplo, inspiração para este protótipo, possui uma trilha sonora característica, mesmo por quem não joga há muito tempo. Apesar disso, este protótipo não produz nenhum som.

Visto todas as limitações da biblioteca de reconhecimento de som descritas acima, a inclusão de música e efeitos sonoros dificultaria ainda mais no foco principal do jogo, que é a fala. Existe, por exemplo, o risco que o som transmitido pelo jogo seja identificado como voz do usuário pelo próprio jogo, causando um *loop* na interação. Também há a dificuldade de que cada celular possui um posicionamento diferente do sistema de auto falante e microfone, então mesmo treinando exaustivamente, resultados obtidos em um celular poderiam ser completamente diferente em outro modelo.

No formulário de avaliação, as pessoas foram questionadas sobre os pontos positivos e negativos da interface e da interação com o jogo, e foram convidadas a deixar sugestões de melhorias e opiniões. As respostas, no geral, foram sucintas e citaram problemas específicos, normalmente sem detalhar ou refinar as sugestões. Como pontos positivos da interação foi relatado que o uso da voz como entrada é melhor do que toque em telas pequenas, a jogabilidade é inovadora, os níveis de dificuldade são interessantes e as fases são complexas a ponto de motivar o jogo, sem frustrar a experiência. Além disso, foi elogiado que o protótipo não parece ser um jogo de criança, no sentido em que ela fique limitada a poucas opções ou simplificações demais, mas é apresentado um mundo e a criança tem liberdade para encarar da forma que quiser, pegando mais cristais quando se sentir mais livre. Por fim, foi dito que o protótipo permite realizar o treinamento de forma lúdica e divertida, em vez de mecanizada e sem graça, como o tratamento normal deve ocorrer, apenas com repetições. As respostas completas dos avaliadores podem ser visualizadas na Tabela A.3, no Apêndice A).

Como pontos negativos da interface do jogo (Tabela A.4, no Apêndice A) um avaliador identificou que o tamanho do personagem é muito pequeno, se comparado ao tamanho da tela. Outro avaliador identificou, quanto à tela inicial, que a barra de progresso diária é pouco intuitiva: após coletar 51 cristais da meta diária de 50, estranhou o progresso estar 51/50 (mais de 100%), pensando que até poderia ser um erro no jogo. Também foi dito por outro avaliador que a mecânica do bloco de pulo confunde na primeira vez em que aparece.

Como pontos positivos da interface do jogo (Tabela A.4, no Apêndice A), quatro avaliadores relataram que a interface é simples. Além disso, um avaliador relatou que a semelhança com Mario é uma coisa boa, visto que as pessoas já conhecem e sabem o que esperar. Outro avaliador elogiou o grau de evolução e o acesso rápido às fases.

Como comentários, sugestões, dúvidas e problemas (Tabela A.5, no Apêndice A), um avaliador sugeriu que o jogo poderia possibilitar às pessoas criarem do jogo. Outro avaliador relatou que o jogo poderia permitir ao fonoaudiólogo inserir as palavras a serem treinadas (incluindo variação de sotaque). Também foi sugerido, pela mesma pessoa, a possibilidade de testar a aplicação com usuários surdos, que também possuam dificuldade em aprender fonemas. Um avaliador relatou que a identificação da fala poderia melhorar por meio de um processo de calibração – dessa forma, o jogo poderia funcionar bem em ambientes com bastante ruído. Outro avaliador também identificou que poderia existir uma fase de treinamento, ou primeira fase guiada, para explicar todas as mecânicas e como funciona o jogo. Um avaliador relatou que o jogo poderia ter mais animações, reações ou possibilidades de ações para ficar ainda mais interessante.

A questão de animações e reações é importante, pois o usuário só sabe se o jogo entendeu a fala quando a gravidade do personagem é invertida. Poderia haver uma melhor sinalização para o usuário saber quando pode falar ou quando o jogo está processando, dessa forma sendo o mais transparente possível e reduzindo o *stress* que pode gerar na falta de *feedback*.

5 Conclusões e Trabalhos Futuros

Nesse trabalho foi proposto, desenvolvido e avaliado um protótipo de jogo para auxiliar na prática de exercícios de fonoaudiologia. Considerando o contexto do Hospital SOBRAPAR, de Campinas/SP, como motivação, foi considerado o público-alvo como sendo crianças de baixa renda, que frequentam o hospital para consultas periódicas e que precisam de apoio para realizar os exercícios de fonoaudiologia no tempo entre uma consulta e outra. A prática dos exercícios fora do ambiente hospitalar (e.g., em casa) é, portanto, crítica para apoiar a evolução da criança no tratamento, e importante o profissional de fonoaudiologia poder acompanhar essa evolução no tempo entre uma consulta e outra.

Na etapa inicial de investigação de soluções existentes, foram pesquisados aplicativos tanto para *Android* quanto para *iOS*, e observado a ausência de uma ferramenta gratuita, funcional, que consumisse pouco poder de processamento e que pudesse ser utilizada como apoio em um hospital com contexto similar ao do SOBRAPAR. A análise de soluções existentes possibilitou identificar exemplos de funcionalidades e atributos que uma aplicação deveria oferecer, e também de problemas a serem evitados.

Com os resultados da análise de soluções existentes, um processo de design colaborativo inspirado nas práticas de Design Social Consciente permitiu considerar aspectos relevantes de um contexto complexo como o hospital e suas partes interessadas. Deste processo surgiram diversas ideias de solução para o problema situadas no contexto da problemática da fonoaudiologia e da realidade social dos indivíduos relacionados, como a falta de Internet, a ida ao hospital pela criança e seu responsável, a necessidade de motivar a criança a fazer os exercícios, e a demanda por recursos que apoiem o profissional da fonoaudiologia em suas práticas. Os resultados do design colaborativo foram utilizados como insumos para a criação do protótipo do jogo Fonopets.

O protótipo do jogo teve cinco fases avaliadas por doze pessoas em uma sessão exploratória. Em relação ao reconhecimento da fala, todas as pessoas experimentaram alguma dificuldade ou reclamação. Ao observar com mais detalhes, foi identificado problemas com o reconhecimento por a ferramenta *PocketSphinx* reconhecer silêncio ou barulho do ambiente como falas válidas. Os resultados mostraram que tanto a ferramenta quanto o modelo utilizado podem melhorar de modo a aumentar a precisão do reconhecimento. A ausência de modelos disponíveis na língua portuguesa que funcionem bem e estejam otimizados para o armazenamento também dificultam. No idioma inglês é possível identificar até frases completas com precisão alta.

Deste modo, se for possível identificar e distinguir palavras em português, o jogo poderá oferecer mecânicas diferentes e recursos de interação mais completos e completos, alinhados com exercícios que os profissionais de fonoaudiologia desejassem elaborar. Por exemplo, a mesma fase poderia funcionar com um grupo de palavras, em vez de uma única palavra pré-selecionada. Também poderia ser desenvolvida uma solução baseada em redes neurais, como proposto em [17]. Um desafio que surge, no entanto, é a necessidade de possuir centenas de horas de material para poder treinar uma rede que produza resultados satisfatórios, como proposto em [19].

Além disso, poderia ser usado processamento de sinais para reconhecimento de sons, permitindo assim sons de longa duração, detecção de sons que não fossem palavras e a

possibilidade de evitar trapaça, que pode prejudicar o próprio tratamento do indivíduo. O jogo também poderia ser avaliado por fonoaudiólogos e até mesmo crianças em tratamento, para averiguar a viabilidade dele em cenários reais.

Os resultados obtidos ao avaliar o protótipo da aplicação com pessoas da área de informática revelaram que, mesmo com o número limitado de participantes para avaliar o jogo, existem diversas situações e questões que devem melhoradas antes de disponibilizar o jogo para uso pelos profissionais e seus pacientes. Ao mesmo tempo, os resultados indicam a viabilidade da ideia, de modo que o conceito do jogo é promissor para motivar os pacientes na execução dos exercícios, para contribuir na evolução do quadro do paciente em seu tratamento, e para apoiar o trabalho do profissional de fonoaudiologia com a criação e adaptação de exercícios personalizados aos seus pacientes, assim como o acompanhamento da evolução do paciente no decorrer do tempo e entre as consultas formais.

Referências

- [1] Abilio Gambim Parada. GameFono. <https://play.google.com/store/apps/details?id=com.abilio.gamefono>, 2018. Online; Acessado em novembro de 2018.
- [2] Rakiv Ahmed and Jonas Aule. An Evaluation of the Framework Libgdx when developing a Game prototype for Android Devices, 2011.
- [3] Amigofono. Amigofono para Pacientes. <https://play.google.com/store/apps/details?id=com.amigofono.app>, 2018. Online; Acessado em novembro de 2018.
- [4] Amit Agarwal. Super Mario Bros Game. <https://www.flickr.com/photos/amit-agarwal/4566449122/>, 2018. Online; Acessado em dezembro de 2018.
- [5] Erik Andersen, Eleanor O'Rourke, Yun-En Liu, Rich Snider, Jeff Lowdermilk, David Truong, Seth Cooper, and Zoran Popovic. The impact of tutorials on games of varying complexity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 59–68. ACM, 2012.
- [6] M Cecilia C Baranauskas. Socially aware computing. In *Proceedings of International Conference on Engineering and Computer Education*, volume 6, 2014.
- [7] Cassio Batista, Thiago Coelho, Bruno Haick, Nelson Neto, and Aldebaro Klautau. Desenvolvimento e Comparação De Reconhedores De Fala Embarcados e Distribuídos Para Android. 2013.
- [8] Richard F Bowman. A "Pac-Man" theory of motivation: Tactical implications for classroom instruction. *Educational Technology*, 22(9):14–16, 1982.
- [9] Cheetah Games. Piano Tiles 2™. <https://play.google.com/store/apps/details?id=com.cmpplay.tiles2>, 2018. Online; Acessado em novembro de 2018.
- [10] Grazielle Costa Coutrin, Luciana Ulhôa Guedes, and Andréa Rodrigues Motta. Treinamento muscular na face: a prática dos fonoaudiólogos de Belo Horizonte. *Revista da Sociedade Brasileira de Fonoaudiologia*, 2008.
- [11] Michele D Dickey. Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development*, 53(2):67–83, 2005.
- [12] Fofuuu - Saúde e Desenvolvimento Infantil. Fofuuu - Para Pais. <https://play.google.com/store/apps/details?id=com.fofuuu.apps.kids>, 2018. Online; Acessado em novembro de 2018.

- [13] G1. Criador do 'flappy bird' cumpre promessa e retira game do ar. <http://g1.globo.com/tecnologia/games/noticia/2014/02/criador-do-flappy-bird-cumpre-promessa-e-retira-game-do-ar.html>, 2018. Online; Acessado em dezembro de 2018.
- [14] Camilla Guarnieri, Bruno Felipe Arndt, and Simone Aparecida Lopes-Herrera. Levantamento dos aplicativos mobile na área de fonoaudiologia. In *Congresso Brasileiro de Fonoaudiologia*. Sociedade Brasileira de Fonoaudiologia, 2014.
- [15] Edward Twitchell Hall et al. *The silent language*, volume 3. Doubleday New York, 1959.
- [16] Elaine Hayashi, Julián E Gutiérrez Posada, Vanessa RML Maike, and M Cecília C Baranauskas. Exploring new formats of the Self-Assessment Manikin in the design with children. In *Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems*, page 27. ACM, 2016.
- [17] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6):82–97, 2012.
- [18] David Huggins-Daines, Mohit Kumar, Arthur Chan, Alan W Black, Mosur Ravishankar, and Alexander I Rudnicky. Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 1, pages I–I. IEEE, 2006.
- [19] Navdeep Jaitly, Patrick Nguyen, Andrew Senior, and Vincent Vanhoucke. Application of pretrained deep neural networks to large vocabulary speech recognition. In *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.
- [20] JJMR Productions. Falar a Brincar. <https://play.google.com/store/apps/details?id=com.productionsjjmrgmail.conscienciafonologica>, 2018. Online; Acessado em novembro de 2018.
- [21] Frogger Konami. Game [arcade], 1981.
- [22] Simone Kühn, Tobias Gleich, Robert C Lorenz, Ulman Lindenberger, and Jürgen Gallinat. Playing super mario induces structural brain plasticity: gray matter changes resulting from training with a commercial video game. *Molecular psychiatry*, 19(2):265, 2014.
- [23] Kai-Fu Lee, Hsiao-Wuen Hon, and Raj Reddy. An overview of the SPHINX speech recognition system. In *Readings in speech Recognition*, pages 600–610. Elsevier, 1990.
- [24] Kecheng Liu. *Semiotics in information systems engineering*. Cambridge University Press, 2000.
- [25] Mayra Souza Resende. Notas pessoais da pesquisadora em visita ao Hospital SOBRAPAR, obtidas via solicitação direta à pesquisadora, 4 2018.
- [26] MeshTech Solutions. Articulation essentials 2.0 - lite. <https://play.google.com/store/apps/details?id=com.mts.articulationessentials.pro.lite>, 2018. Online; Acessado em novembro de 2018.

- [27] MethodShop. Frogger. <https://www.flickr.com/photos/methodshop/2610300581>, 2018. Online; Acessado em dezembro de 2018.
- [28] Kiri Miller. Schizophonic performance: Guitar hero, rock band, and virtual virtuosity. *Journal of the Society for American Music*, 3(4):395–429, 2009.
- [29] Morgade Soft. LeLe Sílabas. <https://play.google.com/store/apps/details?id=com.morgade.lele>, 2018. Online; Acessado em novembro de 2018.
- [30] KAYA NAz and Helena Epps. Relationship between color and emotion: A study of college students. *College Student J*, 38(3):396, 2004.
- [31] Dong Nguyen. Flappy bird. *Apple App Store*, 2013.
- [32] Rafael Oliveira, Pedro Batista, Nelson Neto, and Aldebaro Klautau. Baseline acoustic models for brazilian portuguese using CMU sphinx tools. In *International Conference on Computational Processing of the Portuguese Language*, pages 375–380. Springer, 2012.
- [33] RAFAEL SANTANA OLIVEIRA. DESENVOLVIMENTO DE RECURSOS E FERRAMENTAS PARA RECONHECIMENTO DE VOZ EM PORTUGUÊS BRASILEIRO PARA DESKTOP E SISTEMAS EMBARCADOS. *Mestrado em Ciência da Computação*, page 69, 2012.
- [34] Open Planet Software. Just Press Record. <https://itunes.apple.com/br/app/just-press-record/id1033342465>, 2018. Online; Acessado em dezembro de 2018.
- [35] Roberto Pereira, M Cecília C Baranauskas, and Sergio Roberto P da Silva. Social software and educational technology: informal, formal and technical values. *Journal of Educational Technology & Society*, 16(1):4, 2013.
- [36] Maria Isabel Ramos do Amaral, José Eduardo Martins, and Maria Francisca Colella dos Santos. Estudo da audição em crianças com fissura labiopalatina não-sindrômica. *Brazilian Journal of otorhinolaryngology*, 76(2), 2010.
- [37] Ovidiu Andrei Schipor, Stefan Gheorghe Pentiu, and Maria Doina Schipor. Improving computer based speech therapy using a fuzzy expert system. *Computing and Informatics*, 29(2):303–318, 2012.
- [38] Carine Marlene Schneider and Letícia Galery Medeiros. Criança hospitalizada e o impacto emocional gerado nos pais. *Unoesc & Ciência-ACHS*, 2(2):140–155, 2012.
- [39] Screen Recorder & Video Editor. DU Recorder – Screen Recorder, Video Editor, Live. <https://play.google.com/store/apps/details?id=com.duapps.recorder>, 2018. Online; Acessado em dezembro de 2018.
- [40] The Spriters Resource - Super Mario Bros. <https://www.spriters-resource.com/nes/supermariobros/sheet/52571>, 2018. Online; Acessado em novembro de 2018.
- [41] Visual Reading j.s.a. Visual Reading – special education. <https://play.google.com/store/apps/details?id=com.normex.visualreadingenginapp>, 2018. Online; Acessado em novembro de 2018.

- [42] Yousician Ltd. **Yousician**. https://play.google.com/store/apps/details?id=com.yousician.yousician&hl=pt_br/, 2018. Online; Acessado em dezembro de 2018.
- [43] Yousician Ltd. **Yousician - Learn Guitar, Piano, Bass & Ukulele**. <https://play.google.com/store/apps/details?id=com.yousician.yousician>, 2018. Online; Acessado em novembro de 2018.

Apêndice A: Informações coletadas durante avaliação do protótipo

A.1 Respostas do questionário

ID	1. Idade	2. Gênero	3. Escolaridade	4. Fez disciplina de IHC?
1	23	Masculino	Estudante de mestrado	Sim, na graduação
2	27	Feminino	Estudante de mestrado	Sim, na graduação
3	28	Masculino	Professor	Não
4	31	Masculino	Estudante de doutorado	Sim, na pós graduação
5	22	Masculino	Estudante de mestrado	Sim, na pós graduação
6	45	Masculino	Estudante de doutorado	Não
7	25	Feminino	Estudante de doutorado	Sim, na pós graduação
8	23	Feminino	Estudante de mestrado	Sim, na pós graduação
9	25	Masculino	Estudante de mestrado	Sim, na pós graduação
10	22	Masculino	Estudante de graduação	Não
11	21	Masculino	Estudante de graduação	Não
12	39	Feminino	Estudante de doutorado	Sim, na pós graduação

Tabela A.1: Resposta dos questionários com relação a idade, gênero, escolaridade e se a pessoa já fez alguma disciplina de IHC.

ID	5. Tenho familiaridade com jogos	6. Familiaridade com Super Mario
1	Concordo	Média: joguei algumas vezes
2	Concordo	Alta: jogava bastante no passado
3	Concordo fortemente	Altíssima: ainda hoje jogo muito
4	Concordo fortemente	Altíssima: ainda hoje jogo muito
5	Concordo	Média: joguei algumas vezes
6	Neutro	Baixa: Já ouvi falar, mas nunca joguei
7	Concordo	Alta: jogava bastante no passado
8	Concordo fortemente	Alta: jogava bastante no passado
9	Concordo fortemente	Média: joguei algumas vezes
10	Concordo fortemente	Alta: jogava bastante no passado
11	Concordo fortemente	Média: joguei algumas vezes
12	Concordo	Alta: jogava bastante no passado

Tabela A.2: Resposta dos questionários com relação a familiaridade com jogos e familiaridade com o jogo Super Mario.

ID	7. Pontos NEGATIVOS da interação	8. Pontos POSITIVOS da interação
1	Reconhecimento de voz falha às vezes, depende da posição do microfone; introdução surpresa de novos elementos	Melhor que toque em telas pequenas
2	O jogo não entende a afobação do jogador	A ideia do jogo é muito boa
3	As vezes nao reconhece	Mecanica simples, divertido
4	Tempo baixo de resposta do bichinho; faltou um tutorial ou dicas de jogo na tela; bloqueio de toques na tela, talvez.	Divertido; desafiador; construtivo;
5	As vezes nao captava a voz corretamente, nao surtindo efeito no jogo	Inovação
6	Delay e suscetivel a ruidos externos	Possibilidade de jogar apenas com a voz
7	O pedido da fase era falar uma palavra, porém, não é necessário falar a palavra inteira, apenas o fonema. Poderia ter uma fase de teste para treinamento e explicação, tipo um tutorial.	O jogo é legal de jogar, as fases são desafiadores. Aos poucos vai aprendendo o funcionamento e vai se tornando mais dinâmica a interação.
8	Nao ficou muito claro certos blocos e seu funcionamento	Jogabilidade inovadora, nivel de dificuldades interessantes
9	A identificação da fala é simples e não funciona bem em ambientes com ruído	Ele é bonito e as fases são complexas a ponto de motivar o jogo, sem frustrar a experiência
10	Nao reconhecimento da palavra em certos momentos, assim como o reconhecimento em momentos errados	Algumas palavras, como sabao d caixa, funcionaram muito bem. O jogo se mostrou divertido nas fases em que a interação funcionou.
11	A identificação da voz não é muito precisa	A ideia de usar a voz para conduzir o personagem
12	Não reconheceu algumas palavras, algumas devido ao meu sotaque, então usei o sotaque local. Mensagens em inglês em vez de Português. Não entendi o propósito da parede com interrogação em relação aos espaços abertos sem interrogação.	Pareceu bem util para o treinamento de fonemas. Permite realizar o treinamento de forma lúdica e divertida, em vez de mecanizada e sem graça que deve ser essa atividade só com repetição.

Tabela A.3: Tabela com pontos positivos e negativos da interação.

ID	9. Pontos POSITIVOS da interface	10. Pontos NEGATIVOS da interface
1		Precisa de identificação mais clara da meta diária (51/50 parece um bug)
2	Simples, porem bem elaborada e bonita	Nenhuma
3	Simples, nostalgico	A mecanica do bloco de pulo confunde na primeira vez que aparece
4	Simples; relativamente funcional mas sem feedback	Sem feedback;
5	Bem parecida com a do Mario, um jogo bem popular e que muitas pessoas gostam de jogar	Na interface nenhuma, porem a captacao por voz tem que ser precisa para nao prender o jogo entre paredes
6	Grau de evolucao e acesso rápido às fases	Pouca clareza das funcionalidades
7	Pelo cenário ser parecido com o do jogo Mario, a interação é muito intuitiva.	Alguns pontos que o personagem passa direto é muito rápido, deveria ser um pouco mais lento, principalmente por ser destinado as crianças. O personagem poderia ser maior, ele é muito pequeno para o tamanho da tela.
8	Nostalgia com o mario, interface simples e intuitiva	Não é muito diferente do que ja tem disponível
9	A interface é simples e cumpre seu propósito	A identificação da fala não é precisa e complica a jogabilidade
10	Menu inicial de escolha de fases bonito. Fluidez do cenario boa	Nenhuma
11	Facilidade para jogar e entender o objetivo	Só um bug na ultima fase onde o personagem ficou preso.
12		

Tabela A.4: Tabela com pontos positivos e negativos da interface.

ID	11. Deixe comentários, sugestões, dúvidas, problemas, etc.
1	Trilha-sonora de fundo (para ambientação)
2	
3	
4	Legal;
5	Melhorar a captação por voz
6	Verificar compatibilidade de recursos de hardware para tratar delay
7	Fase de treinamento, ou uma primeira fase guiada.
8	Adorei o jogo, seria legal no futuro permitir que as pessoas criem fases do jogo
9	Melhorar a identificação da fala, calibrar, algo nesse sentido. O jogo em si é interessante e com certeza motivará as crianças
10	
11	Ter mais animações ou possibilidades de ações no jogo pode deixar o jogo mais interessante
12	Testar aplicação com usuarios surdos para aprender fonemas. Permitir que fono-audiólogo insira as palavras a serem treinadas pelos pacientes de acordo com os fonemas que eles estão aprendendo naquela fase do tratamento.

Tabela A.5: Tabela com comentários, sugestões, dúvidas, problemas, etc.

A.2 Desempenho no jogo

ID	1	2	3	4	5
1	13	5	9	12	12
2	11	2	13	13	10
3	17	3	13	10	12
4	11	2	10	15	8
5	15	19	11	12	6
6	14	1	17	14	12
7	11	15	12	14	10
8	13	1	15	15	11
9	7	8	12	13	11
10	15	1	8	17	13
11	19	1	12	16	10
12	17	7	13	10	12
n° máximo	24	25	18	20	18

Tabela A.6: Número de cristais coletados em relação ao número da fase. A última linha contém os valores máximos de cristais disponíveis por fase.

ID	1 (lula)	2 (gato)	3 (verde)	4 (sabão)	5 (caixa)
1	15	11	7	20	8
2	19	3	16	19	14
3	17	1	8	11	10
4	20	11	29	34	10
5	19	21	15	19	9
6	15	5	18	23	10
7	18	21	11	29	10
8	19	2	21	21	14
9	16	9	21	16	8
10	18	4	24	18	12
11	15	2	16	28	15
12	21	11	10	14	10

Tabela A.7: Número de palavras reconhecidas por usuário em relação ao número da fase e a palavra que devia ser falada. Não há um número máximo de acertos.

ID	1 (lula)	2 (gato)	3 (verde)	4 (sabão)	5 (caixa)
1	5	3	10	13	6
2	14	5	3	8	6
3	3	2	1	3	3
4	5	6	6	7	2
5	4	6	3	8	3
6	4	5	3	6	6
7	4	4	3	5	2
8	4	2	4	4	6
9	6	5	8	11	2
10	9	3	6	14	7
11	3	3	4	14	6
12	4	6	3	5	0

Tabela A.8: Número vezes em que uma palavra foi falada mas o aplicativo entendeu como silêncio ou ruído e ignorou.

ID	1 (lula)	2 (gato)	3 (verde)	4 (sabão)	5 (caixa)
1	0	9	2	4	0
2	2	9	0	7	2
3	0	4	2	1	3
4	1	6	3	6	4
5	3	4	3	6	3
6	5	8	4	4	2
7	3	5	0	6	1
8	4	8	1	3	2
9	2	6	4	7	3
10	3	6	3	7	2
11	1	6	1	4	3
12	1	6	6	3	0

Tabela A.9: Número vezes em que uma palavra foi falada e o aplicativo registrou, mas não encontrou definição no dicionário morfológico usado.

ID	1 (lula)	2 (gato)	3 (verde)	4 (sabão)	5 (caixa)
1	75.00%	47.83%	36.84%	54.05%	57.14%
2	54.29%	17.65%	84.21%	55.88%	63.64%
3	85.00%	14.29%	72.73%	73.33%	62.50%
4	76.92%	47.83%	76.32%	72.34%	62.50%
5	73.08%	67.74%	71.43%	57.58%	60.00%
6	62.50%	27.78%	72.00%	69.70%	55.56%
7	72.00%	70.00%	78.57%	72.50%	76.92%
8	70.37%	16.67%	80.77%	75.00%	63.64%
9	66.67%	45.00%	63.64%	47.06%	61.54%
10	60.00%	30.77%	72.73%	46.15%	57.14%
11	78.95%	18.18%	76.19%	60.87%	62.50%
12	80.77%	47.83%	52.63%	63.64%	100.00%
Média	71.29%	37.63%	69.84%	62.34%	65.26%

Tabela A.10: Porcentagem de acertos no reconhecimento de fala por usuário por fase.

	+2	+1	0	-1	-2
1 (lula)	5	7	0	0	0
2 (gato)	0	3	6	3	0
3 (verde)	4	7	1	0	0
4 (sabão)	3	7	2	0	0
5 (caixa)	5	6	0	1	0

Tabela A.11: Resultado do SAM Assesment agrupado e simplificado. No eixo vertical, estão as fases. No eixo horizontal, a emoção expressada pela pessoa ao jogar as fases.