



## **EasyTEFA Solução para apoio à terapia da fala**

**RÚBEN MAGINA PEREIRA VIEIRA RIBEIRO**

Outubro de 2019

# **EasyTEFA – Solução para apoio à terapia da fala**

**Rúben Magina Pereira Vieira Ribeiro**

**Dissertação para obtenção do Grau de Mestre em  
Engenharia Informática, Área de Especialização em  
Sistemas Gráficos e Multimédia**

**Orientador: Prof. Doutor António Vieira de Castro**

**Supervisores: Prof. Doutora Maria João Cunha**

**Prof. Doutor André Araújo**



*Aos meus pais, irmão e avós por tudo o que me proporcionaram  
ao longo da vida.*

*À minha namorada pelo apoio constante.*



# Resumo

A terapia da fala é uma ciência da área da saúde que trata distúrbios no campo da linguagem. Foca-se essencialmente em situações de atraso no desenvolvimento e de perturbações específicas da linguagem falada e escrita, mas também nas perturbações articulatórias, no caso da voz, e mesmo em situações de perturbação da fala, quer resultantes de lesão neurológica quer na dificuldade em executar os movimentos necessários à sua produção.

O terapeuta da fala é o profissional que tem como função o tratamento e estudo científico da comunicação e por consequência identificar e tentar tratar os distúrbios e perturbações a ela associados.

Verificamos que existe uma lacuna ao nível de soluções tecnológicas que ajudem o terapeuta da fala na sua atividade pelo que definimos que o objetivo principal da presente dissertação seria o de tentar perceber e analisar tecnologias que pudessem proporcionar o desenvolvimento de uma solução com potencial para ajudar estes profissionais.

Para o efeito entramos em contato com terapeutas da fala que prontamente se disponibilizaram a supervisionar este trabalho sob o ponto de vista dos procedimentos necessários para a implementação da solução a desenvolver. Estas reuniões e sugestões proporcionaram o planeamento do EasyTEFA tendo sido possível adaptar os requisitos ao seu desenvolvimento. Após terem sido analisadas diferentes tecnologias potenciais de reconhecimento de voz a ser utilizadas para o desenvolvimento do protótipo demos início ao processo de planeamento do desenvolvimento do EasyTEFA.

Com os requisitos definidos e considerando a existência de algumas limitações nas tecnologias analisadas, construiu-se, desenhou-se e implementou-se o protótipo que foi testado e avaliado nas vertentes de usabilidade e funcionalidade tendo sido analisados e explorados os resultados obtidos.

Conclui-se que com esta solução é apresentada uma mais valia no âmbito do tratamento de problemas da fala e que pode contribuir para uma análise e até tratamentos mais eficazes, sobretudo com crianças. O trabalho termina com a apresentação de algumas conclusões.

**Palavras-chave:** Terapia da fala, APIs áudio, EasyTEFA, *Speech Recognition*, Saúde



# Abstract

Speech therapy is a health science that treats language disorders. It focuses mainly on situations of developmental delay and specific disorders of spoken and written language, but also on articulatory disorders in the case of voice, and even in situations of speech impairment, whether resulting from neurological damage or difficulty in performing the necessary movements for its production.

The speech therapist is the professional whose function is the treatment and scientific study of communication and consequently identify and try to treat the disorders and perturbations associated with it.

We verified that there is a gap in the level of technological solutions that help the speech therapist in his activity, so we defined that the main objective of the present dissertation would be to try to understand and analyze technologies that could provide the development of a potential solution to help these professionals.

For this purpose, we contacted speech therapists who readily offered to supervise this work from the point of view of the necessary procedures for the implementation of the solution to be developed. These meetings and suggestions provided the planning for EasyTEFA and it was possible to adapt the requirements to its development. After analyzing different potential speech recognition technologies to be used for prototype development we started the EasyTEFA development planning process.

With the defined requirements and considering the existence of some limitations in the analyzed technologies, the prototype that was tested and evaluated in the usability and functionality aspects was built, designed and implemented and the results obtained were analyzed and explored.

It is concluded that this solution has added value in the treatment of speech problems and can contribute to a more effective analysis and even treatments, especially with children. The work ends with the presentation of some conclusions.

**Keywords:** Speech therapy, audio APIs, EasyTEFA, Speech Recognition, Health



# Agradecimentos

Não poderia deixar de agradecer a todos aqueles que me ajudaram na realização desta dissertação, especialmente ao meu orientador, o Prof. Doutor António Vieira Castro, pela sua sempre pronta disponibilidade, dedicação e empenho.

Agradeço ainda aos meus supervisores, da Escola Superior de Saúde do Politécnico do Porto, a Prof. Doutora Maria João Pinho e o Prof, Doutor André Araújo.

Aos meus colegas de curso agradeço todo o apoio, especialmente ao Tiago Ferreira que sempre me acompanhou e apoiou ao longo de todo o Mestrado.

Aos meus Pais e avós por toda a dedicação e apoio ao longo deste trabalho.

Um agradecimento muito especial à minha namorada pelas horas que lhe roubei.

Ao ISEP pela qualidade de ensino que me proporcionou.

À ESS (Escola Superior de Saúde) e muito particularmente aos especialistas e alunos da terapia da fala pelo apoio e interação constante.

Ao LAMU (Laboratório Multimédia) do DEI (Departamento de Engenharia Informática) do ISEP por ter autorizado a utilização do estúdio de multimédia.



# Índice

|   |           |
|---|-----------|
| <b>Capítulo 1 - Introdução</b> .....                              | <b>1</b>  |
| 1.1 Contextualização .....  | 1         |
| 1.2 O problema .....  | 2         |
| 1.3 Objetivos e contributos esperados .....                       | 2         |
| 1.4 Especificação da hipótese .....                               | 3         |
| 1.5 Metodologia de avaliação .....                                | 4         |
| 1.6 Motivação .....   | 4         |
| 1.7 Organização do presente documento .....                       | 4         |
| <b>Capítulo 2 - Enquadramento teórico</b> .....                   | <b>7</b>  |
| 2.1 A saúde .....   | 7         |
| 2.2 A tecnologia e a saúde .....                                  | 8         |
| 2.3 O caso da terapia da fala .....                               | 9         |
| 2.3.1 Métodos tradicionais de terapia da fala.....                | 10        |
| 2.3.2 Métodos tecnológicos de terapia da fala.....                | 11        |
| 2.4 Uso de tecnologia de reconhecimento de voz emergentes .....   | 13        |
| 2.4.1 Assistentes virtuais por voz.....                           | 14        |
| 2.4.2 Serviços Speech-To-Text .....                               | 18        |
| 2.5 Análise de valor.....   | 19        |
| 2.5.1 5 elementos do modelo de Peter Koen .....                   | 20        |
| 2.5.2 Value for customer Perceived value .....                    | 21        |
| 2.5.3 Proposta de Valor .....                                     | 21        |
| 2.5.4 Modelo Canvas .....   | 21        |
| <b>Capítulo 3 - Análise técnica e de requisitos</b> .....         | <b>23</b> |
| 3.1 Metodologia .....   | 23        |
| 3.2 Análise de soluções existentes de reconhecimento de voz ..... | 24        |
| 3.2.1 Microsoft <i>Speech</i> API.....                            | 24        |
| 3.2.2 Google Cloud Speech-To-Text .....                           | 25        |
| 3.2.3 Speechmatics .....  | 26        |
| 3.2.4 IBM Watson STT.....   | 27        |
| 3.2.5 Comparações.....  | 27        |
| 3.3 Análise conjunta com terapeutas.....                          | 28        |
| 3.3.1 Primeira proposta de solução .....                          | 29        |
| 3.3.2 Proposta de solução final .....                             | 30        |
| 3.4 Requisitos funcionais .....                                   | 31        |
| 3.5 Requisitos não funcionais .....                               | 32        |
| 3.5.1 Usabilidade .....   | 32        |
| 3.5.2 Hardware e Software .....                                   | 32        |

|  |   |           |
|--|---|-----------|
| 3.6  | Reflexão.....                           | 32        |
| <b>Capítulo 4 - Design e desenvolvimento do protótipo.....</b> |   | <b>33</b> |
| 4.1  | Design.....                             | 33        |
| 4.2  | Tecnologias e Linguagens .....          | 36        |
| 4.3  | Desenvolvimento .....                   | 36        |
| 4.3.1  | Reconhecimento de voz.....              | 37        |
| 4.3.2  | Lógica do exercício.....                | 38        |
| 4.3.3  | Protótipo final .....                   | 40        |
| 4.3.4  | Testes preliminares .....               | 47        |
| <b>Capítulo 5 - Avaliação do EasyTEFA .....</b>                |   | <b>49</b> |
| 5.1  | Contextualização .....                  | 49        |
| 5.2  | Avaliação de usabilidade .....          | 53        |
| 5.3  | Avaliação funcional .....               | 57        |
| 5.4  | Apreciação final após a avaliação ..... | 58        |
| <b>Capítulo 6 - Conclusões e trabalho futuro.....</b>          |   | <b>61</b> |
| 6.1  | Conclusões.....                         | 62        |
| 6.2  | Trabalho futuro .....                   | 62        |

# Lista de Figuras

|  |    |
|--|----|
| Figura 1 - A saúde e a sua relação intensa com a tecnologia .....  | 8  |
| Figura 2 - Tecnologias na Saúde .....  | 9  |
| Figura 3 – Processo de uso de objetos e cartas .....   | 11 |
| Figura 4 – Exemplo de exercício da Happies .....   | 12 |
| Figura 5 – Ecrã inicial da aplicação Falar a Brincar .....   | 12 |
| Figura 6 – Amazon Echo .....   | 14 |
| Figura 7 – Percentagem de pessoas que utilizam menos o <i>smartphone</i> desde que tiveram um assistente digital por voz dedicado..... | 15 |
| Figura 8 – Assistente virtual da Microsoft: Cortana .....  | 15 |
| Figura 9 – Exemplo de lembrete baseado localização com a Cortana.....  | 16 |
| Figura 10 – Logo da Bixby.....   | 16 |
| Figura 11 - exemplos do que é possível fazer com a Bixby .....   | 17 |
| Figura 12 – Logo da Siri .....   | 17 |
| Figura 13 – Exemplo de uso da Siri .....   | 18 |
| Figura 14 – O Dictate no MS Word .....   | 19 |
| Figura 15 – Modelo Canvas .....  | 22 |
| Figura 16 – Logo Microsoft Speech SDK.....   | 24 |
| Figura 17 – Logo Google Cloud Speech-To-Text .....   | 25 |
| Figura 18 – Logo Speechmatics .....  | 26 |
| Figura 19 – Logo IBM Watson.....   | 27 |
| Figura 20 – Classes lógicas do protótipo .....   | 34 |
| Figura 21 – Classes gráficas do protótipo.....   | 35 |
| Figura 22 – Método que gera a gramática .....  | 38 |
| Figura 23 – Método que responde ao evento de voz reconhecida .....   | 39 |
| Figura 24 – Ecrã de entrada do EasyTEFA .....  | 40 |
| Figura 25 – Ecrã de Informação e Regras .....  | 41 |
| Figura 26 – Ecrã da funcionalidade adicionar carta.....  | 41 |
| Figura 27 – Ecrã da funcionalidade adicionar carta preenchido .....  | 42 |
| Figura 28 – Ecrã da funcionalidade adicionar lista de jogo .....   | 43 |
| Figura 29 – Ecrã da funcionalidade adicionar lista de jogo preenchido.....   | 43 |
| Figura 30 – Ecrã de escolha de lista de jogo.....  | 44 |
| Figura 31 – Ecrã do jogo .....   | 44 |
| Figura 32 – Ecrã de jogo quando o utilizador acerta .....  | 45 |
| Figura 33 – Ecrã de jogo quando o utilizador erra.....   | 45 |
| Figura 34 – Ecrã de erros do jogo sem erros .....  | 46 |
| Figura 35 – Ecrã de erros do jogo com erros .....  | 46 |
| Figura 36 – Ecrã de créditos .....   | 47 |
| Figura 37 – Mecanismo adotado para a sessão de testes e avaliação do protótipo .....   | 50 |
| Figura 38 – Procedimento do teste de usabilidade .....   | 50 |
| Figura 39 – Processo de respostas após os testes de usabilidade e funcional.....   | 51 |



# Lista de Gráficos

|  |    |
|--|----|
| Gráfico 1 – Distribuição por género dos inquiridos .....                               | 51 |
| Gráfico 2 - Idade dos inquiridos.....  | 52 |
| Gráfico 3 – Experiência dos inquiridos com computadores.....                           | 52 |
| Gráfico 4 – Identificação de problemas de fala nos inquiridos .....                    | 53 |
| Gráfico 5 – Facilidade de navegação nos menus da aplicação .....                       | 54 |
| Gráfico 6 – Simplicidade do processo de adicionar uma carta.....                       | 54 |
| Gráfico 7 - Utilidade da funcionalidade adicionar carta .....                          | 55 |
| Gráfico 8 – Simplicidade do processo de adicionar uma lista de jogo.....               | 55 |
| Gráfico 9 – Utilidade da funcionalidade adicionar lista de jogo .....                  | 56 |
| Gráfico 10 – Simplicidade do processo de começar um jogo / terapia .....               | 56 |
| Gráfico 11 – Problemas com a deteção de voz .....                                      | 57 |
| Gráfico 12 – Dos problemas com a deteção, os que se encontravam num ambiente ruidoso.. | 57 |
| Gráfico 13 – Importância do painel final com os erros .....                            | 58 |



# Acrónimos e Símbolos

## A

- AI** *Artificial Intelligence*
- API** *Application Programming Interface*
- AVC** Acidente Vascular Cerebral

## E

- ESS** Escola Superior de Saúde
- EUNETHTA** European network for Health Technology Assessment

## I

- ISEP** Instituto Superior de Educação do Porto

## M

- ML** *Machine Learning*

## P

- PDF** Portable Document File

## S

- SDK** Software Development Kit
- STT** Speech-to-text

## T

- TI** Tecnologias de Informação.
- TTS** Text-to-speech

## W

- WWW** World Wide Web.



# Capítulo 1 - Introdução

*No primeiro capítulo, Introdução, é feita uma contextualização geral do tema que se relaciona com a terapia da fala. Identifica-se o problema da lacuna de tecnologias de apoio a esta atividade e apresentam-se os objetivos e contributos esperados com a realização do presente estudo. Apresenta-se a Especificação da hipótese e abordam-se as metodologias de avaliação. O capítulo apresenta ainda a motivação e termina com a organização do presente documento.*

## 1.1 Contextualização

A terapia da fala é uma atividade que trabalha os distúrbios orais (fala, canto, etc) provenientes de diversas origens (como a evidenciada em crianças ou resultante de traumas como quedas ou acidentes). A terapia da fala é uma profissão relativamente recente encontrando-se a sua prática profissional em desenvolvimento. Esta profissão reúne conhecimentos de diferentes áreas científicas, estando incluídas as ciências médicas e biológicas, as ciências do comportamento, da linguagem e comunicação, incluindo-se também a linguística, a fonética e a fonologia, a tecnologia aplicada à fala, assim como, a área especializada da patologia e terapêutica da fala e da linguagem.

A terapia da fala concilia o conhecimento científico com as competências clínicas de forma a obter a melhor gestão possível da situação clínica e pedagógica dos utentes (Associação Portuguesa de Terapeutas da Fala, 2018). A principal atividade do Terapeuta da Fala é a prevenção, avaliação, intervenção e estudo científico das perturbações da comunicação humana, englobando não só todas as funções associadas à compreensão e expressão da

linguagem oral e escrita, mas também outras formas de comunicação não verbal. Verifica-se que existe uma grave lacuna de soluções de apoio à terapia da fala que possam apoiar o terapeuta e o próprio paciente. Consideramos que é possível desenvolver uma solução para apoiar os profissionais que atuam nessa área não só a prevenir os distúrbios da voz como melhorá-la, atuando no aperfeiçoamento e promoção da saúde vocal sobretudo ao nível da fala.

## **1.2 O problema**

Embora a necessidade de terapia da fala seja recorrente, não só em crianças, mas também em adultos que por acidentes diversos perdem competências orais, esta área da saúde ainda não foi suficientemente bem explorada no que concerne ao uso de tecnologias.

O terapeuta trabalha muito à base dos seus conhecimentos e baseado em procedimentos clínicos adequados ao tratamento.

Apenas recentemente se verificou um maior desenvolvimento de tecnologias relacionadas com a voz e mais concretamente com o reconhecimento da voz.

No entanto o problema subsiste dado que ainda não é visível a sua aplicação focalizada na terapia da fala.

## **1.3 Objetivos e contributos esperados**

Dada a recente evolução tecnológica e a evolução recente de APIs relacionadas com o reconhecimento de voz pretendemos trabalhar diretamente com profissionais de terapia da fala iniciando o estudo com profissionais da Escola Superior de Saúde do Porto de modo a analisar, desenhar, desenvolver, implementar e testar o potencial de uma solução informática aplicada à área da terapia da fala que será denominada de “EasyTEFA”.

Estas APIs de reconhecimento de voz avançadas permitem a conversão de áudio em texto aplicando modelos de redes neurais avançados, Algumas API reconhecem mais de 120 idiomas e variantes o que poderá significar uma adaptação ao mercado global da solução a criar.

A aplicação a desenvolver terá como objetivos ajudar o terapeuta da fala na sua atividade e apoiar o paciente no tratamento com a proposta de exercícios de comparação de imagens. Quando o paciente disser um dos nomes associados às imagens, o “EasyTEFA” reconhecerá e dará uma resposta visual ao utilizador, de modo a este poder ter uma melhor noção do próprio erro e num cenário ideal, levá-lo a interessar-se a aprender mais sobre como superar o problema.

Idealmente, a solução trabalhará com crianças, mas também tem como objetivo ser adaptável para poder ser usada por adultos.

## **1.4 Especificação da hipótese**

A hipótese a testar determinará se a solução estudada e desenvolvida tem o desempenho devido.

Neste caso, a solução será bem-sucedida se os terapeutas da fala conseguirem obter resultados favoráveis provenientes do uso da solução, face aos métodos tradicionais. Dentro destes resultados espera-se que o paciente consiga ter uma melhor noção de que está a errar e que fique motivado a melhorar.

Caso não se verifique nenhum tipo de vantagem, a solução obtida não será bem-sucedida.

De qualquer modo consideraremos que se estiver funcional e usável e se revelar potencial para proporcionar interação entre o paciente e o software será possível testar a hipótese de que com o protótipo a desenvolver pode ser possível contribuir para abrir portas para o uso de reconhecimento de voz aplicado ao tratamento de casos de terapia da fala.

Pretendemos obter o *feedback* de utilizadores e dos terapeutas no sentido de aferir se é possível preconizar em que aspetos a solução pode ser benéfica, ou não. Será também através da coleta desta informação, que realizaremos ajustes e melhoraremos a solução para que ela possa ir de encontro às necessidades dos utilizadores finais.

## **1.5 Metodologia de avaliação**

Atendendo a que o software a desenvolver pretende identificar problemas relacionados com a terapia da fala consideramos que os testes terão de ser focalizados numa bateria de exercícios providenciada pelos especialistas de terapia da fala de modo a medir a eficácia do reconhecimento da fala pelo sistema. Paralelamente poderão ser considerados os métodos A B sendo que no método A um grupo de terapeutas utilizará o software para trabalhar com cerca de 10 exercícios identificados e os restantes terapeutas (ou alunos do mestrado de terapia da fala) farão os mesmos testes ao paciente recorrendo aos modelos de observação em papel tradicional.

## **1.6 Motivação**

Nos últimos 20 anos, a tecnologia evoluiu de uma maneira imprevisível e embora nos dias de hoje já seja bastante avançada e de evolução mais lenta, ainda há muitas áreas à espera de serem exploradas. Mesmo a nível da saúde, embora existam tecnologias de ponta para algumas áreas específicas, também há áreas em que a tecnologia ainda não é muito utilizada.

Na terapia da fala os métodos mais utilizados para o tratamento dos pacientes ainda são analógicos. Isto não representa qualquer problema, pois os métodos são utilizados e dão resultados. No entanto já não há muito por onde evoluir e nesse sentido, a tecnologia é capaz de ajudar.

Como tal o presente estudo foi motivado pela ideia de ajuda ao próximo, usando tecnologia na área da terapia da fala.

## **1.7 Organização do presente documento**

Este documento está dividido em seis capítulos.

No primeiro capítulo, Introdução, é feita uma contextualização geral do tema que se relaciona com a terapia da fala. Identifica-se o problema da lacuna de tecnologias de apoio a esta atividade e apresentam-se os objetivos e contributos esperados com a realização do presente

estudo. Apresenta-se a Especificação da hipótese e abordam-se as metodologias de avaliação. O capítulo apresenta ainda a motivação e termina com a organização do presente documento.

O segundo capítulo, Enquadramento teórico, é onde é aprofundado, do geral para o particular, o tema deste estudo. É elaborado um estado da arte que relaciona a saúde com a tecnologia e depois é especificado o caso da terapia da fala. Neste estado da arte são ainda alvo de estudo os métodos atualmente utilizados na terapia da fala e as implementações já existentes das tecnologias que se pretendem utilizar no desenvolvimento da solução. Por fim, é elaborado ainda uma análise de valor da solução proposta.

No terceiro capítulo, Análise técnica e de requisitos, é referida a metodologia utilizada e de seguida são elaboradas diversas análises. Primeiro, é elaborada uma análise de soluções atuais de reconhecimento de voz que possam ser utilizadas no desenvolvimento da solução, depois, é feita uma análise da proposta de solução com terapeutas que resulta numa nova proposta de solução. Finalmente, com a proposta de solução final, são levantados os requisitos funcionais e não funcionais da mesma, focando-se os não funcionais na usabilidade e *hardware* e *software*.

O capítulo quatro, Design e desenvolvimento do protótipo, descreve as escolhas técnicas de design, justifica o uso das tecnologias e linguagens e apresenta uma breve explicação do desenvolvimento dos mecanismos fulcrais do protótipo. No fim é mostrado e explicado o protótipo final.

O quinto capítulo, Avaliação do EasyTEFA, descreve, analisa e exhibe os resultados a avaliação de usabilidade e funcional do protótipo. Com toda esta informação, é ainda apresentado neste capítulo uma apreciação final.

Finalmente, no sexto capítulo, Conclusões e trabalho futuro, são tiradas conclusões dos resultados do presente estudo e do protótipo resultante do mesmo. Usando essas conclusões é ainda realizada uma pequena reflexão sobre o que poderá ser o trabalho futuro deste estudo e protótipo.



## Capítulo 2 - Enquadramento teórico

*É neste capítulo que é aprofundado, do geral para o particular, o tema deste estudo. É elaborado um estado da arte que relaciona a saúde com a tecnologia e depois é especificado o caso da terapia da fala. Neste estado da arte são ainda alvo de estudo os métodos atualmente utilizados na terapia da fala e as implementações já existentes das tecnologias que se pretendem utilizar no desenvolvimento da solução. Por fim, é elaborado ainda uma análise de valor da solução proposta.*

### 2.1 A saúde

Segundo Carta de Ottawa, publicada em 1986, a “Saúde é um recurso para a vida, não o objetivo de viver” (Carta de Ottawa, 1986), que, embora tenha sido alvo de diversas definições ao longo do tempo se assume como um “estado completo de bem-estar físico, mental e social, e não apenas na ausência de doença”, constituindo “um dos direitos fundamentais de todo o ser humano sem distinção de raça, religião, opiniões políticas e condições económicas e sociais” (OMS, 1946). Mais tarde, em 1984, a OMS alargou o conceito para “um estado em que o indivíduo ou um grupo são capazes de concretizar aspirações, satisfazer necessidades, e lidar com o ambiente que o rodeia. (...) É um conceito positivo, que dá ênfase aos recursos pessoais e sociais de cada um, assim como às suas capacidades físicas” (Carta de Ottawa, 1986)

## 2.2 A tecnologia e a saúde

De acordo com a European Network for Health Technology Assessment (EUNETHTA), a designação de tecnologia da saúde, traduzida para português pela Comissão Europeia (2012), define-se como “a aplicação dos conhecimentos científicos aos cuidados de saúde e à prevenção da doença”, sendo exemplos desta “os métodos de tratamento e de diagnóstico, equipamentos médicos, métodos de prevenção e reabilitação, sistemas organizativos e de apoio, no quadro dos quais são prestados os cuidados de saúde”.



Figura 1 - A saúde e a sua relação intensa com a tecnologia <sup>1</sup>

Tendo por base a definição acima, a tecnologia na saúde abrange diversas áreas e permite que haja um maior desenvolvimento de procedimentos e tratamentos (como se pretende ilustra com a Figura 1).

Este desenvolvimento, por sua vez, permite que se atue mais preventivamente, melhorando a qualidade de vida dos humanos e leva também à inovação. Com novos conhecimentos e novas tecnologias, é possível criar novas ferramentas para facilitar o diagnóstico de uma determinada condição (ex: termómetro para medir a febre) ou para ajudar um indivíduo a ter melhor condição física (ex: prótese para uma perna amputada que permite o indivíduo voltar a andar).

No final, a aplicação da tecnologia na saúde leva a que a qualidade de vida dos indivíduos seja melhor e pode ser usada para desenvolvimento, prevenção, inovação, como ferramenta ou até para melhorar a qualidade de vida (Figura 2).

---

<sup>1</sup> Imagem disponível em: <https://gestaoopme.com.br/wp-content/uploads/2016/07/tecnologia-em-saude-750x400.jpg>

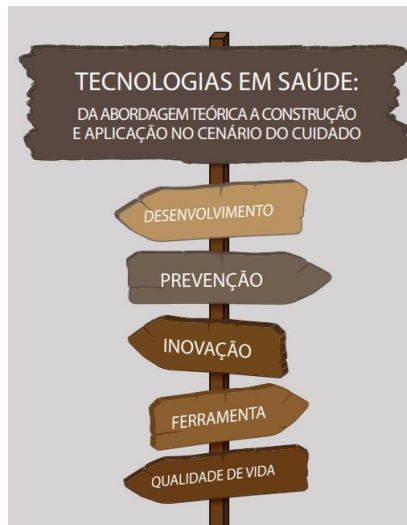


Figura 2 - Tecnologias na Saúde

### 2.3 O caso da terapia da fala

Atualmente, a terapia da fala trata a área da linguagem, tratando situações de afasia (perturbação da linguagem resultante de lesão neurológica), atrasos no desenvolvimento da linguagem, perturbações específicas da linguagem falada e escrita, e no campo da fluência/voz, intervindo nas perturbações articulatórias (substituições, omissões ou outros erros no discurso das crianças), na disartria (perturbação da fala resultante de lesão neurológica), na apraxia (dificuldade em executar movimentos neuromusculares necessários à produção da fala), na difluência (gaguez) e na disfonia (perturbação da qualidade vocal caracterizada por rouquidão, esforço vocal, maus hábitos vocais, má utilização do aparelho vocal). (CUF, s.d.)

A comunicação exerce-se de forma verbal e não verbal, englobando todas as funções associadas à compreensão e à expressão da linguagem oral e escrita. Para que seja eficaz, a terapia da fala tem de abordar todas as componentes adjacentes à comunicação.

O nosso dia a dia seria praticamente impossível se não nos comunicássemos com outras pessoas. Para comunicar-nos, falamos, gesticulamos, escrevemos e eventualmente ainda outras formas. Na prática, comunicar é a capacidade de descrever, partilhar, participar e interagir com outras pessoas.

Ao longo deste estudo, iremos debruçar-nos essencialmente sobre a comunicação verbal, que é a forma mais utilizada de comunicação e relacionamento humano.

Através da fala podemos transmitir ideias e expressar pensamentos, seja qual for o nível de complexidade e detalhe. A comunicação verbal abrange sobretudo a oralidade que é proveniente da capacidade de falar.

A fala ela está inserida no nosso dia a dia, em diversas situações sociais. Para que seja eficaz, a comunicação verbal deve ocorrer na forma passiva e ativa, mas é fundamental que Ambos os interlocutores tenham em pleno as suas faculdades verbais e uma fala fluente.

Se isto não acontece pode ser devido a problemas da fala e é necessário recorrer à terapia que pode originar a sua correção ou minimizar os efeitos desse problema. (Matérias Escolares: Ensino Fundamental e Médio, 2019)

A intervenção com terapia da fala pode ocorrer em várias alturas da vida, quer seja nos primeiros anos de vida para ajudar a combater dificuldades e falhas comunicativas, quer seja mais tarde em adulto para quem tiver problemas derivados de doença, como por exemplo um AVC (Acidente Vascular Cerebral).

### **2.3.1 Métodos tradicionais de terapia da fala**

Na terapia da fala existem vários métodos e soluções para lidar com os problemas dos pacientes. Estes problemas, como referido anteriormente, podem ser muito variados e subjetivos, dependendo de paciente para paciente. Esta variedade de casos para tratar torna o trabalho de um profissional terapeuta da fala bastante desafiante, especialmente na escolha de que ferramentas vai utilizar para atacar o problema apresentado pelo paciente. São apresentados abaixo exemplos de métodos e ferramentas utilizadas pelos terapeutas nas suas sessões de terapia da fala.

#### **2.3.1.1 *Uso de objetos e cartas***

O uso de objetos e cartas que representem animais, objetos e letras, como se tenta exemplificar com a Figura 3, é muito comum, principalmente em tratamentos com crianças. Esta é uma ótima maneira de diagnosticar quais são os sons e letras que a criança tem mais dificuldade em dizer e que acabam por lhe dificultar a comunicação com os outros. (Pequenada, s.d.)



Figura 3 – Processo de uso de objetos e cartas<sup>2</sup>

Este tipo de exercício é personalizável a cada criança, pois o terapeuta tem vários objetos e cartas para mostrar e assim diagnosticar os distúrbios do paciente.

#### **2.3.1.2 Uso de espelho**

Através do uso de um espelho o terapeuta consegue mostrar visualmente ao paciente como é que este deve posicionar os lábios e a língua de modo a produzir os sons desejados. (Pequenada, s.d.) Desta forma o paciente tem uma referência visual e facilita o processo de cópia, podendo acelerar a aprendizagem.

### **2.3.2 Métodos tecnológicos de terapia da fala**

Na terapia da fala, continuam a utilizar-se muitos métodos não tecnológicos com os pacientes. Não quer dizer que não existam soluções tecnológicas que possam ajudar na terapia da fala, até porque existem algumas, mas a utilização destas soluções não faz uso do máximo potencial tecnológico atual. São apresentados abaixo exemplos de ferramentas tecnológicas que podem ser utilizadas pelos terapeutas nas suas sessões de terapia da fala.

#### **2.3.2.1 Happies**

A Happies é uma aplicação móvel, disponível quer para *smartphones* Android, quer para Apple. Esta é “uma aplicação de apoio ao desenvolvimento da motricidade oro facial, com mais de 30 exercícios suporte ao desenvolvimento de bochechas, lábios e língua.” (Azevedo, 2018)

---

<sup>2</sup> Imagem disponível em: <https://www.saudecuf.pt/sites/saudecuf/files/2019-05/filho-terapia-fala-cuf.png>



Figura 4 – Exemplo de exercício da Happies<sup>3</sup>

Os exercícios desta aplicação focam-se na repetição de um movimento de modo a ensinar o como fazê-lo (Figura 4).

A utilização desta aplicação permite aos pais e terapeutas avaliar o desempenho da criança, tentando que a terapia esteja presente durante mais tempo, além da sessão em si. (Boas Notícias, 2018)

### 2.3.2.2 Falar a Brincar

Falar a Brincar (Figura 5) é mais uma aplicação, presente para Android, que procura ajudar na terapia da fala. O foco desta aplicação é melhorar a consciência fonológica e procura usar exercícios já utilizados nas sessões de terapia da fala para as crianças começarem a conhecer os exercícios. (JJMR Productions, 2017)



Figura 5 – Ecrã inicial da aplicação Falar a Brincar<sup>4</sup>

---

<sup>3</sup> Imagem disponível em: <https://apps.apple.com/us/app/happies/id1224164066?l=pt&ls=1>

### **2.3.2.3 Palavras aos Bocadinhos**

Esta é mais uma aplicação que tenta ser uma ferramenta na prática de terapia da fala. Está disponível para dispositivos Apple e não é completamente gratuita.

De um ponto de vista de terapia da fala, esta aplicação móvel permite trabalhar a consciência silábica em diferentes níveis: Segmentação, Identificação, Síntese, Ajuda à Evocação, Adição, Omissão, Substituição e Inversão. (EMS Software, 2015)

## **2.4 Uso de tecnologia de reconhecimento de voz emergentes**

No mundo atual da tecnologia existem diversas soluções que fazem uso da fala para executar ações ou comandos dentro da aplicação ou para simplesmente transformar em texto o que utilizador disse em voz alta (*text to speech* ou TTS).

As empresas como a Google, Microsoft, Apple, Samsung e outros grandes das tecnologias, têm vindo a investir neste tipo de tecnologia para que os seus assistentes pessoais (Google Assistant, Cortana, Siri, Bixby, etc) consigam captar melhor o que cada utilizador diferente diga. Isto obriga a uma certa tolerância a erros de dicção e à captação de inúmeros sotaques diferentes para uma captação mais eficaz. Conseguindo obter uma boa tradução do discurso em texto, estes assistentes pessoais conseguem executar bastantes comandos e facilitar a vida dos seus utilizadores. Algumas das muitas funções disponibilizadas pelos assistentes são: efetuar chamadas de voz para um contacto específico, enviar uma mensagem de texto, criar um evento no calendário, criar um alarme para uma hora específica, começar a tocar uma música ou até mesmo fazer uma pesquisa na internet.

Todas estas soluções têm um impacto na qualidade de vida de um individuo que faça uso das mesmas. São apresentados alguns exemplos no ponto 2.4.1 Assistentes virtuais por voz, e 2.4.2 Serviços *Speech-To-Text*.

---

<sup>4</sup>Imagem disponível em:  
[https://play.google.com/store/apps/details?id=com.productionsjjmrgmail.conscienciafonologica&hl=pt\\_PT](https://play.google.com/store/apps/details?id=com.productionsjjmrgmail.conscienciafonologica&hl=pt_PT)

### 2.4.1 Assistentes virtuais por voz

Os assistentes virtuais por voz já existiam nos *smartphones* e até mesmo computadores há algum tempo, mas recentemente saíram para as casas dos consumidores que decidirem comprar um assistente digital por voz dedicado. Como exemplos temos o dispositivo Echo com Alexa da Amazon (Figura 6).



Figura 6 – Amazon Echo <sup>5</sup>

Este tipo de dispositivo é um *smart speaker* com uma assistente suportada por inteligência artificial que permite ao utilizador executar diversos serviços através de comandos de voz – por música a tocar, executar tarefas como ligar ou desligar as luzes de diferentes zonas da casa ou outras ainda relacionadas, não só, mas também com sistemas de casa inteligente e saber notícias ou o tempo, entre outros.

Segundo um estudo criado pela Accenture, 66% dos utilizadores de um assistente digital por voz dedicado usam menos o *smartphone*. Dos utilizadores todos, 64% utiliza menos os seus *smartphones* para entretenimento, 58% utiliza menos para compras online e 56% para pesquisas de informação geral. (Sovie, Roberts, Murdoch, McMahon, & Schoelwer, 2018)

---

<sup>5</sup> Imagem disponível em <http://stackoverflow.co/wp-content/uploads/2019/07/voice-control-device-amazon-echo-dot-black-loading-zoom-voice-control-devices.jpg>

66% of consumers who own a stand-alone DVA use their smartphone for fewer activities

**Q: To what extent do you agree with the following statement?**

**Since I got my digital voice assistant device, I use my smartphone for fewer activities**

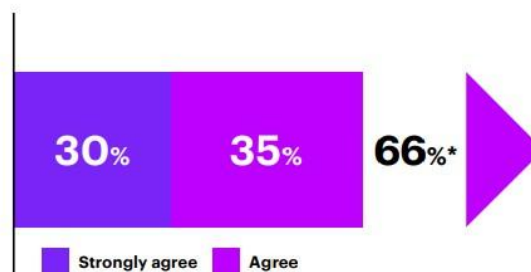


Figura 7 – Percentagem de pessoas que utilizam menos o *smartphone* desde que tiveram um assistente digital por voz dedicado

#### 2.4.1.1 O caso da Cortana (Microsoft)

De acordo com o site da Microsoft<sup>6</sup> a Cortana “é o assistente digital verdadeiramente pessoal”. Ela permite executar tarefas simples como criar lembretes ou iniciar qualquer aplicação presente no Windows.



Figura 8 – Assistente virtual da Microsoft: Cortana <sup>7</sup>

Segundo a Microsoft, os lembretes são bastante versáteis e podem, por exemplo, ser baseados numa localização (Figura 9).

<sup>6</sup> <https://www.microsoft.com/pt-br/windows/cortana>

<sup>7</sup> <https://www.hackread.com/wp-content/uploads/2018/03/cortana-lets-hacker-infect-windows-pc-when-locked-1.jpg>

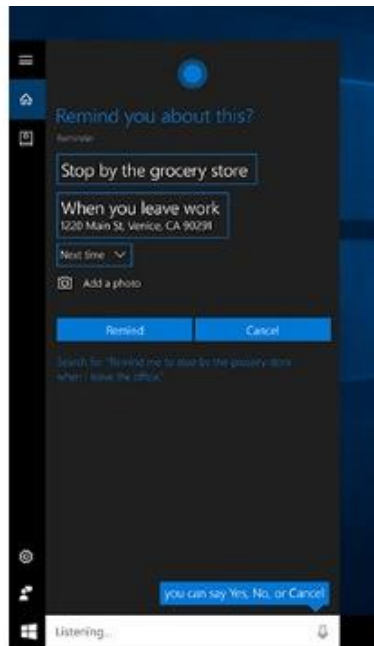


Figura 9 – Exemplo de lembrete baseado localização com a Cortana<sup>8</sup>

Basta dizer "Ei, Cortana, me lembra de..." para começar a registrar um lembrete.

#### 2.4.1.2 O caso da Bixby (Samsung)

A Bixby é a assistente pessoal da Samsung e equipa os seus dispositivos de topo. Fornece assim a capacidade de execução de tarefas quotidianas como ler notificações, enviar mensagens de texto ou até mesmo pesquisar informação.



Figura 10 – Logo da Bixby<sup>9</sup>

A Bixby dá aos seus utilizadores liberdade para se concentrarem no que é mais importante podendo torná-los ainda mais produtivos.

Na figura seguinte apresentam-se alguns exemplos do que é possível fazer com a Bixby.

---

<sup>8</sup> Imagem disponível em: <https://www.microsoft.com/pt-br/windows/cortana>

<sup>9</sup> Imagem disponível em: [https://www.phonearena.com/news/Samsungs-Bixby-will-soon-work-with-Google-Play-YouTube-and-Google-Maps\\_id112504](https://www.phonearena.com/news/Samsungs-Bixby-will-soon-work-with-Google-Play-YouTube-and-Google-Maps_id112504)

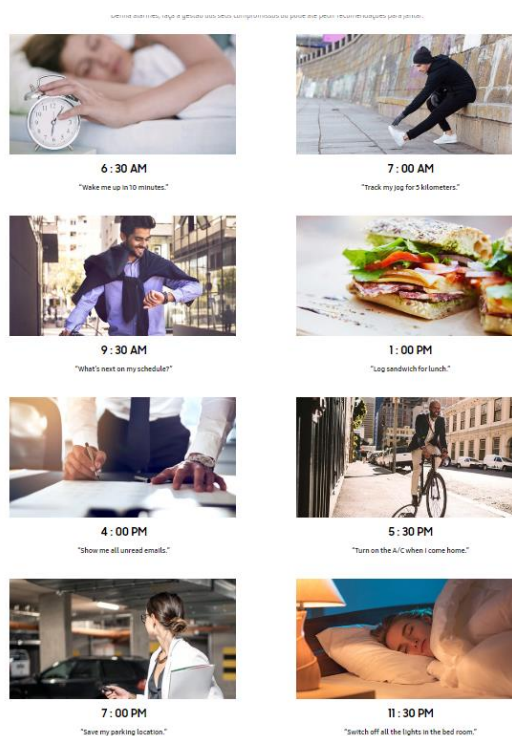


Figura 11 - exemplos do que é possível fazer com a Bixby<sup>10</sup>

Como se pode ver, podemos facilmente, através de comandos de voz, definir por exemplo o despertador para as “6:30 AM” dizendo apenas “Wake me up at 6:30 AM” ou solicitar a visualização dos emails dizendo “Show me all unread emails.”

### 2.4.1.3 O caso da Siri (Apple)

Tal como os outros grandes da tecnologia, a Apple possui uma assistente pessoal, a Siri. Esta é muito potente e permite que se use o *smartphone* sem mãos para quase todas as funções do dia a dia.



Figura 12 – Logo da Siri<sup>11</sup>

---

<sup>10</sup> Exemplos disponíveis em <https://www.samsung.com/pt/apps/bixby/> consultados em março 2019

Para evocar a Siri o utilizador pode dizer uma frase em voz alta que a ativa e de seguida o comando de voz para fazer a ação que pretende. Na figura seguinte está exemplificado um desses exemplos.



Figura 13 – Exemplo de uso da Siri<sup>12</sup>

Através da frase “Hey Siri, cal Mom on speaker”, a Siri põe a chamar para a mãe em alta-voz.

## 2.4.2 Serviços Speech-To-Text

Também conhecidos como serviços de ditado, as soluções deste tipo facilitam o transformar de discurso de uma determinada língua em texto. Esta tecnologia não só pode tornar a vida de uma pessoa que precise de escrever muito ao computador mais fácil, como também pode tornar possível a escrita de todo por parte de um individuo que por alguma razão não consigo utilizar um teclado corretamente.

---

<sup>11</sup> Imagem disponível em: <https://en.wikipedia.org/wiki/Siri>

<sup>12</sup> Imagem disponível em: <https://www.apple.com/siri/>

### 2.4.2.1 Microsoft Office

A Microsoft já possui uma solução destas integradas no próprio Word (Figura 14), Outlook e PowerPoint. Isto permite a qualquer utilizador de uma destas aplicações, e que fale uma das línguas suportadas, que use a sua voz para escrever um documento, email ou diapositivo.

No site oficial<sup>13</sup> tem informação sobre perguntas frequentes, e as capacidades da solução.

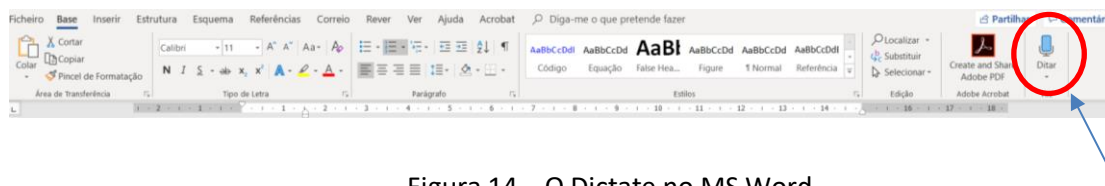


Figura 14 – O Dictate no MS Word

### 2.4.2.2 Google Cloud Speech-to-Text

A Google disponibiliza uma API que permite a desenvolvedores que utilizem os serviços de reconhecimento de fala avançado próprio. Tal como o nome sugere, este serviço é baseado na nuvem e não é gratuito, mas permite aos desenvolvedores utilizar ferramentas muito interessantes, como por exemplo, o reconhecimento automático de idioma. No site oficial<sup>14</sup> é possível experimentar o serviço gratuitamente e obter mais informação sobre o como funciona esta API.

### 2.4.2.3 Microsoft Speech API

A Microsoft desenvolveu uma API gratuita para os desenvolvedores de Visual Basic, uma das linguagens de programação da Microsoft, poderem usar *engines* de TTS ou de reconhecimento de fala. Isto permite que as aplicações ajam de forma independente da internet quer para o TTS quer para o reconhecimento da fala.

## 2.5 Análise de valor

A análise de valor de uma nova solução é sempre importante dado que obriga a uma reflexão prévia cuidada que pode demonstrar se o que se pretende fazer tem ou não potencial.

---

<sup>13</sup> <https://dictate.ms/>

<sup>14</sup> <https://cloud.google.com/speech-to-text/>

### 2.5.1 5 elementos do modelo de Peter Koen

Os 5 elementos chave do modelo “the new concept development model” (NCD) são:

- Identificação das oportunidades: No mercado das soluções para ajuda na terapia da fala as soluções são ainda escassas, sendo assim uma oportunidade para este projeto. Beneficia-se ainda do facto de se tratar de uma área relacionada com a saúde, o que significa que existe um grande mercado à espera de novas soluções;
- Análise de oportunidades: Embora não se possa quantificar o número de crianças ou outros indivíduos adultos com problemas de fala;
- Geração de ideia: recentemente, após um contacto com especialistas da terapia da fala da Escola Superior de Saúde do Instituto Politécnico do Porto, foi possível perceber que existia uma grave lacuna no que respeita a soluções tecnológicas para apoio ao tratamento dos problemas de fala. Estes terapeutas, após uma reunião com a equipa que irá desenvolver o presente estudo, perceberam que é possível desenvolver uma solução protótipo no sentido de ser testada com pessoas com problemas de fala. A ideia de desenvolver uma solução desta natureza pretende ajudar os terapeutas na sua tarefa e só é possível pelo facto de recentemente terem aparecido novas APIs de áudio;
- Seleção da ideia: após a análise das várias abordagens possíveis, e após algumas reuniões com especialistas da terapia da fala focalizamos uma solução que permita solicitar ao paciente a leitura de textos previamente escritos pelos terapeutas devendo a solução a desenvolver identificar as palavras faladas erradamente para que o terapeuta possa insistir nesses termos. Está ainda pensada uma solução alternativa para crianças que não sabem ler, sendo que nestes casos o terapeuta irá incentivar a criança a repetir termos ou sons que normalmente não conjuga corretamente. Neste caso poderá ser proposta uma espécie de jogo em que a criança é recompensada com pontos cada vez que consegue dizer o termo corretamente;
- Definição do conceito: a solução a ser desenvolvida consiste na utilização de microfones que irão capturar a fala do paciente e identificar problemas de oralidade. Dada a enorme quantidade de diferentes casos possíveis de problemas, pretende-se que a solução seja adaptativa, obrigando à interação do terapeuta consoante o problema.

### **2.5.2 Value for customer Perceived value**

A criação de uma aplicação para auxiliar um terapeuta da fala aumenta o repertório de ferramentas com que o mesmo pode diagnosticar ou tratar um paciente. Isto tem um grande valor não só para o terapeuta como para o paciente, que é diagnosticado e/ou tratado mais eficazmente. No entanto, dependendo do tempo limite da sessão de terapia, o terapeuta pode se ver numa situação em que precise de escolher entre ferramentas e o facto de ter mais uma pode dificultar essa mesma escolha.

### **2.5.3 Proposta de Valor**

A solução descrita neste estudo apresenta diversas propostas de valor. A primeira proposta de valor desta solução é a polivalência e abrangência em que pode ser utilizada, sendo versátil.

Facilita também a captação da atenção de um paciente com 2 a 10 anos apresentando assim outra proposta de valor, ser cativante para um determinado público.

Por fim, insere fatores de ação reação para o paciente, evidenciando de forma mais clara o problema.

### **2.5.4 Modelo Canvas**

Na Figura 15 apresentada abaixo, está descrito o modelo de negócio.



Figura 15 – Modelo Canvas

## Capítulo 3 - Análise técnica e de requisitos

*Neste capítulo é analisado e descrito todo o processo de desenvolvimento que levou à criação do protótipo EasyTEFA, desde as decisões tomadas e os seus porquês, ao seu relacionamento e aplicabilidade no âmbito da terapia da fala e ajuda ao terapeuta.*

*São analisadas diversas soluções de processamento e reconhecimento de voz de modo a ser selecionada a mais adequada ao contexto do projeto e ao seu desenvolvimento.*

*É também efetuada a documentação da visão tida inicialmente e de como esta foi adaptada em concordância com as necessidades específicas dos terapeutas da fala, de modo a que o projeto seja o mais útil possível num contexto de ferramenta auxiliar de uma sessão de terapia da fala. É assim feito e documentado este ajuste medido, quer das expectativas, quer dos requisitos do presente projeto.*

### 3.1 Metodologia

O presente estudo visa encontrar lacunas nas atuais práticas da terapia da fala, de modo a tentar ajudar o terapeuta, munindo-o de mais uma ferramenta que o consiga auxiliar no desempenho de suas funções.

Conforme mencionado anteriormente, a tecnologia está em constante evolução e existem várias soluções de reconhecimento de voz. Este projeto tenciona explorar esta tecnologia na

área da saúde, mais especificamente, na terapia da fala. Como tal é feita uma análise neste sentido.

Com a intenção de tornar o projeto o mais útil possível para os terapeutas, são analisadas com estes ideias e levantados determinados requisitos e mais valias, ajustando-se assim o projeto à realidade.

Finalmente, com o objetivo de desenvolver um protótipo funcional, foi feito um planeamento da fase de desenvolvimento que foi posto em prática.

## **3.2 Análise de soluções existentes de reconhecimento de voz**

O mundo da tecnologia está em constante evolução e são cada vez mais as soluções existentes no mercado para reconhecimento de voz. De modo a escolher a melhor alternativa possível para o contexto do projeto, é efetuada uma análise de várias destas soluções, comparativamente.

### **3.2.1 Microsoft *Speech* API**

A Microsoft Speech API, também conhecida por SAPI, é uma API desenvolvida pela Microsoft, que torna possível o uso de *Speech Recognition* e *Speech Synthesis* em aplicações desenvolvidas para o ambiente Windows.

A SAPI 5 é a versão mais atual e que integra nativamente nos sistemas operativos mais recentes da Microsoft, sendo utilizada pelas aplicações Microsoft Office, por exemplo.

De modo a implementar as funcionalidades disponibilizadas pela SAPI em aplicações não Microsoft, os programadores podem utilizar o Microsoft Speech Software Development Kit.



Figura 16 – Logo Microsoft Speech SDK<sup>15</sup>

---

<sup>15</sup> Imagem disponível em: <https://getintopc.com/software/development/speech-sdk-free-download/>

### 3.2.1.1 Vantagens

As vantagens do uso desta tecnologia são:

- É uma tecnologia gratuita;
- Suporta vários idiomas, entre os quais, português de Portugal;
- A integração com aplicações Windows é muito facilitada;
- Para além de reconhecimento de voz, suporta *Text-To-Speech*.

### 3.2.1.2 Desvantagens

A grande desvantagem do uso da SAPI, é a limitação de distribuição da aplicação apenas a ambientes Windows.

## 3.2.2 Google Cloud Speech-To-Text

A *Google Cloud Speech-To-Text* é uma API desenvolvida pela Google e que fornece o serviço de reconhecimento de voz. Esta API faz uso de *machine learning* para melhorar o seu desempenho ao longo do tempo.

Esta API pode ser usada em várias linguagens de programação, aumentando assim a sua abrangência e versatilidade.



Figura 17 – Logo Google Cloud Speech-To-Text<sup>16</sup>

### 3.2.2.1 Vantagens

As vantagens desta API são:

- Reconhece 120 idiomas e variantes, entres os quais, português de Portugal;

---

<sup>16</sup> Imagem disponível em: <https://cloud.google.com/speech-to-text/>

- Identifica automaticamente o idioma falado;
- Utiliza *machine learning* para melhorar o seu desempenho ao longo do tempo.

### 3.2.2.2 Desvantagens

Infelizmente, esta API não é de utilização gratuita, sendo essa uma das grandes desvantagens para o seu uso no presente projeto.

### 3.2.3 Speechmatics

A Speechmatics API é uma API especializada em transformar dados de voz em texto para análise e acessibilidade. A Speechmatics API faz também o uso de *machine learning* para melhorar os seus resultados ao longo do tempo.



Figura 18 – Logo Speechmatics<sup>17</sup>

#### 3.2.3.1 Vantagens

A Speechmatics API tem como vantagens:

- É fácil de utilizar;
- Bastante precisa no reconhecimento;
- Reconhece vários idiomas, entre os quais, português;
- Uso de *machine learning* para melhorar o desempenho ao longo do tempo.

#### 3.2.3.2 Desvantagens

A grande desvantagem desta API é o facto de não ser gratuita, cobrando por cada chamada do serviço.

---

<sup>17</sup> Imagem disponível em: <https://www.redbeemedia.com/news/red-bee-media-to-showcase-real-time-live-automatic-captioning-of-the-abcs-24-hour-online-news-stream-at-nab-2019-in-partnership-with-speechmatics/speechmatics-logo/>

### 3.2.4 IBM Watson STT

Outra API de reconhecimento de voz, desenvolvida pela IBM, é a IBM Watson STT. Esta API faz parte de todo um conjunto de outros produtos da IBM, a IBM Watson, que fazem uso de inteligência artificial para processamento e análise de dados.



Figura 19 – Logo IBM Watson<sup>18</sup>

#### 3.2.4.1 Vantagens

As vantagens da IBM Watson STT resumem-se a:

- Facilidade de uso;
- Uso de inteligência artificial para obtenção de resultados mais precisos;
- Reconhece múltiplos idiomas.

#### 3.2.4.2 Desvantagens

Infelizmente a IBM Watson STT tem desvantagens:

- É gratuita apenas até 500 minutos por mês;
- Das variantes do português, apenas reconhece o português do Brasil.

### 3.2.5 Comparações

De modo a facilitar a comparação das APIs enumeradas acima é apresentada a tabela 1.

Os parâmetros a comparar foram os seguintes:

---

<sup>18</sup> Imagem disponível em: <https://futuretv.dk/go/en/program-2018/ibm-watson-logo/>

- Custo;
- Reconhecimento da Língua Portuguesa;
- Usa AI e/ou Machine Learning;
- Pode ser usada com C#.

| API                         | Custo                               | Reconhece português                                | Usa AI e/ou M. L | Pode ser usada com C# |
|-----------------------------|-------------------------------------|--|------------------|-----------------------|
| Microsoft Speech API        | Gratuito                            | Sim, português de Portugal                         | Não              | Sim                   |
| Google Cloud Speech-To-Text | Pago                                | Sim, português de Portugal                         | Sim              | Sim                   |
| Speechmatics                | Pago por cada utilização de serviço | Sim, mas não especifica se é de Portugal ou Brasil | Sim              | Sim                   |
| IBM Watson STT              | Pago acima de 500 minutos por mês   | Sim, mas só português do Brasil                    | Sim              | Não                   |

Tabela 1 – Comparação de APIs de reconhecimento de voz

Como se pode verificar, apenas a Microsoft Speech API era gratuita e cumpria os requisitos mínimos que procurávamos.

### 3.3 Análise conjunta com terapeutas

No âmbito do projeto abrangido por esta dissertação, foi possível manter contato com profissionais de saúde na área da terapia da fala, da Escola Superior de Saúde do Porto. Com esta relação privilegiada, foi possível encaminhar o projeto, dentro das linhas iniciais, para uma maior utilidade num contexto de uma sessão de terapia da fala.

Depois de reunir com a Professora Doutra Maria João Cunha e o Professor Doutor André Araújo e de discutidas várias ideias, percebeu-se, de uma forma geral, de que forma a tecnologia de reconhecimento de voz poderia ajudar numa sessão de terapia da fala. Neste sentido e como referido no Capítulo 2, foram analisadas possíveis soluções já existentes com o

mesmo objetivo. Embora já existam soluções tecnológicas na área da terapia da fala, nenhuma das analisadas no Estado da Arte faz uso de tecnologias de reconhecimento de voz. Partindo deste ponto, foram exploradas duas soluções, a primeira sendo modificada para o que veio a ser a segunda e final solução, por base nos requisitos e limitações obtidos da primeira.

### **3.3.1 Primeira proposta de solução**

Inicialmente, a solução teorizada passou por criar uma aplicação que mostrasse uma frase para o paciente ler. Esta frase poderia ser escrita pelo terapeuta de modo a incidir e trabalhar para melhorar os problemas do paciente. Fazendo uso das tecnologias de reconhecimento de voz, a aplicação iria captar a leitura do paciente e mostrar no ecrã o que foi lido, comparando palavra a palavra e mostrando, através de contraste, as palavras ditas incorretamente (utilização da cor vermelha nos erros, por exemplo).

Embora pareça ser uma solução interessante, tem algumas limitações, listadas em baixo:

- É necessário que o paciente saiba ler, limitando a abrangência do público alvo;
- A tecnologia *Speech-To-Text* tem limitações no reconhecimento de voz, e a língua portuguesa é uma das mais complexas, em parte por ter muitas variâncias fonéticas. Isto dificulta o processo de captação do que foi dito. Isto pode tornar as captações de voz menos fiéis e conseqüentemente diminuir a eficácia da solução;
- Outra variável que impactua negativamente os resultados é a existência de barulho ambiente, que provoca falsos reconhecimentos.

No entanto, esta proposta apresenta ainda alguns pontos positivos que devem ser considerados, listados abaixo:

- Pode ser útil em vários problemas da fala, não só relacionados com perturbações articatórias (substituições, omissões ou outros erros no discurso das crianças), mas também na disartria (perturbação da fala resultante de uma lesão neurológica) e na apraxia (dificuldade em executar movimentos neuromusculares necessários à produção da fala);
- É personalizável pelo terapeuta de modo a atacar o problema do paciente da maneira que o terapeuta considerar mais eficaz;

- Fornece uma reação com representação visual ao paciente sobre o erro cometido, com o objetivo de facilitar a compreensão.

Todos estes pontos, positivos e negativos, foram considerados e numa nova reunião com a Professora Doutora Maria João Cunha, ajustou-se a solução.

### **3.3.2 Proposta de solução final**

Após serem analisados todos os pontos positivos e negativos da solução anterior e em conjunto com a Professora Doutora Maria João Cunha, foi definida uma nova solução. Esta é mais específica e particular, focada nas perturbações articulatórias (substituições e omissões), mas com uma abrangência de público alvo maior, pois não depende também da leitura, mas sim apenas da fala.

Esta solução consiste então no desenvolvimento de uma aplicação que faz uso de tecnologias de reconhecimento de voz para identificar palavras ditas pelo paciente, que estão associadas a imagens. O uso de imagens em alternativa ao uso de texto permite que uma criança que não saiba ler possa utilizar a aplicação. São então apresentadas duas imagens que representam duas palavras à escolha do terapeuta, em que uma delas é considerada correta. Esta palavra correta é a que o paciente tem dificuldade em dizer. É dada uma resposta visual ao paciente sobre o que a aplicação interpretou para que ele, com o apoio da terapeuta, consiga perceber e tentar corrigir/melhorar a sua fala. Os pares certo/errado são inseridos numa lista para que o paciente tenha vários exercícios de treino.

Esta proposta, tendo sido atingida em conjunto com a Professora Doutora Maria João Cunha, tem considerações de um profissional da área. A solução em si continua a ter pontos positivos e negativos, mas está mais ajustada à realidade de uma sessão de terapia da fala. Alguns dos pontos positivos são transversais à solução anterior.

São abaixo listados todos os pontos positivos da solução final:

- Pode ser útil em vários problemas da fala, maioritariamente relacionados com perturbações articulatórias (substituições, omissões ou outros erros no discurso das crianças);
- É personalizável pelo terapeuta de modo a atacar o problema do paciente da maneira que o terapeuta considerar mais eficaz;

- Fornece uma reação com representação visual ao paciente sobre o erro cometido, com o objetivo de facilitar a compreensão;
- A personalização permite abranger também um público adulto;
- As tecnologias *Speech-To-Text* têm mais facilidade em captar apenas uma palavra ou em vez de várias seguidas ao construir uma frase inteira.

Nesta solução não deixaram de existir pontos negativos, mas fez-se os possíveis para que fossem limitados aos das tecnologias e mesmo estes que fossem trabalhados de modo a serem atenuados. Segue abaixo a listagem dos pontos negativos:

- A tecnologia *Speech-To-Text* tem limitações no reconhecimento de voz, e a língua portuguesa é uma das mais complexas, em parte por ter muitas variâncias fonéticas. Isto dificulta o processo de captação do que foi dito. Isto pode tornar as captações de voz menos fiéis e conseqüentemente diminuir a eficácia da solução;
- Outra variável que impactua negativamente os resultados é a existência de barulho ambiente, que provoca falsos reconhecimentos.

### 3.4 Requisitos funcionais

Os requisitos funcionais são fundamentais no processo de desenvolvimento de um projeto. É através deles que é possível determinar certas necessidades de uma aplicação para que esta tenha um bom funcionamento. Os requisitos funcionais foram os seguintes:

- Existir um menu principal onde o utilizador pode navegar pela aplicação;
- Ser possível adicionar Cartas;
  - Escolher a palavra que representa a imagem
- Ser possível criar Listas de Jogo, também com Cartas novas, definindo as cartas certas e erradas;
- No ecrã dos exercícios, ter duas imagens;
- A aplicação captar a palavra dita pelo utilizador e associar com a imagem correta;
- No final de um exercício têm que ser mostrados os resultados dos erros.

## **3.5 Requisitos não funcionais**

Para que a aplicação possa funcionar devidamente, é ainda necessário cumprir certos requisitos não funcionais.

### **3.5.1 Usabilidade**

A nível de usabilidade, esta solução tem de ser intuitiva e fácil de usar, de modo a diminuir o tempo gasto a tentar compreender a mesma. Como tal, dentro da aplicação, é incluída uma secção dedicada à explicação de como esta funciona.

### **3.5.2 Hardware e Software**

Para que esta solução possa ser utilizada, o utilizador tem que utilizar um computador com sistema operativo Windows 8 ou mais recente. Tem também que ter os serviços de reconhecimento de voz do próprio Windows instalados e ativos.

Do ponto de vista de hardware, é essencial o uso de um microfone para que a aplicação funcione. E ainda neste tópico, não se aconselha o uso de microfones incorporados nos computadores portáteis, nem nas webcams. A qualidade do microfone influencia o reconhecimento de voz e conseqüentemente os resultados dos exercícios a que o utilizador está a tentar responder.

## **3.6 Reflexão**

Após a análise de tecnologias com potencial para suportar o desenvolvimento previsto e numa fase em que já tinham sido testadas algumas APIs e soluções relacionadas com o uso do reconhecimento de voz, considerou-se estarem reunidas as condições mínimas para poder avançar para o estudo do “Design e desenvolvimento do protótipo” após algumas reuniões com os especialistas em terapia da fala da ESS.

# Capítulo 4 - *Design* e desenvolvimento do protótipo

*O atual capítulo descreve as escolhas técnicas de design, justifica o uso das tecnologias e linguagens e apresenta uma breve explicação do desenvolvimento dos mecanismos fulcrais do protótipo. No fim é mostrado e explicado o protótipo final.*

## 4.1 Design

De um ponto de vista de *design* da aplicação, esta terá uma arquitetura consideravelmente simples. Terá que ter uma camada de lógica que irá interagir com a API de reconhecimento de voz e identificar a palavra que foi dita, para a poder comparar com as opções disponíveis em forma de nomes das cartas. Terá que ter ainda uma camada gráfica que será a visualizada pelo utilizador e com a qual este interage. É nesta camada que será mostrado ao utilizador os menus para executar ações dentro da aplicação. O utilizador poderá então aceder à criação de cartas, criação de listas de cartas para jogar, às regras e informações, aos créditos e aos exercícios de voz. Ao entrar para os exercícios, o utilizador poderá ainda escolher qual a lista, se existir mais que uma, que pretende utilizar para fazer os exercícios.

Da camada lógica, existiram quatro classes, como representado na Figura 20.

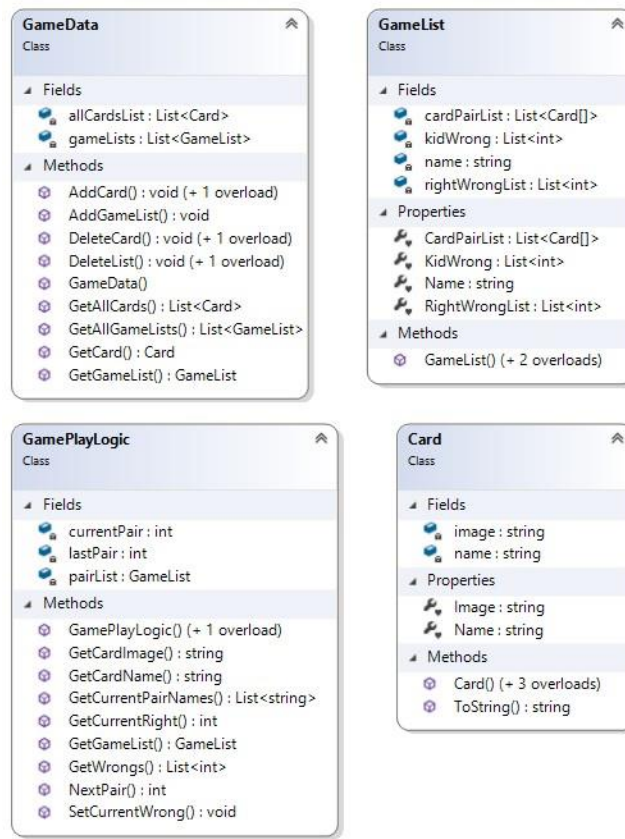


Figura 20 – Classes lógicas do protótipo

A classe “*GameData*” será a responsável pelo armazenamento e tratamento dos dados. Vai ser esta classe a que regista todas as cartas e listas de exercícios. É através da mesma que se poderá obter os dados necessários para a lógica dos exercícios.

A classe “*GamePlayLogic*” será a classe que comanda os exercícios de voz. É esta que recebe da classe “*GameData*” a informação necessária para começar e correr uma lista de exercícios. É com esta classe que a interface gráfica irá interagir para obter os dados e os apresentar no ecrã. Quando se escolher começar um exercício, será dada a opção de escolher uma lista

A classe “*GameList*” representa a lista dos exercícios. Contém as cartas que vão ser utilizadas, a informação sobre as que são consideradas certas ou erradas e ainda é aqui que fica registado as cartas que o paciente disse e não corresponderam às que o terapeuta pretendia.

Por fim, a classe “*Card*”, da camada lógica, contém a informação relevante para uma carta a ser mostrada num exercício. Esta informação é um nome, que é a palavra que tem que ser dita, e a localização do ficheiro do tipo imagem a ser carregado.

Passando então à análise da camada gráfica, podemos começar por observar as suas classes, associadas às vistas do utilizador, na Figura 21 abaixo.

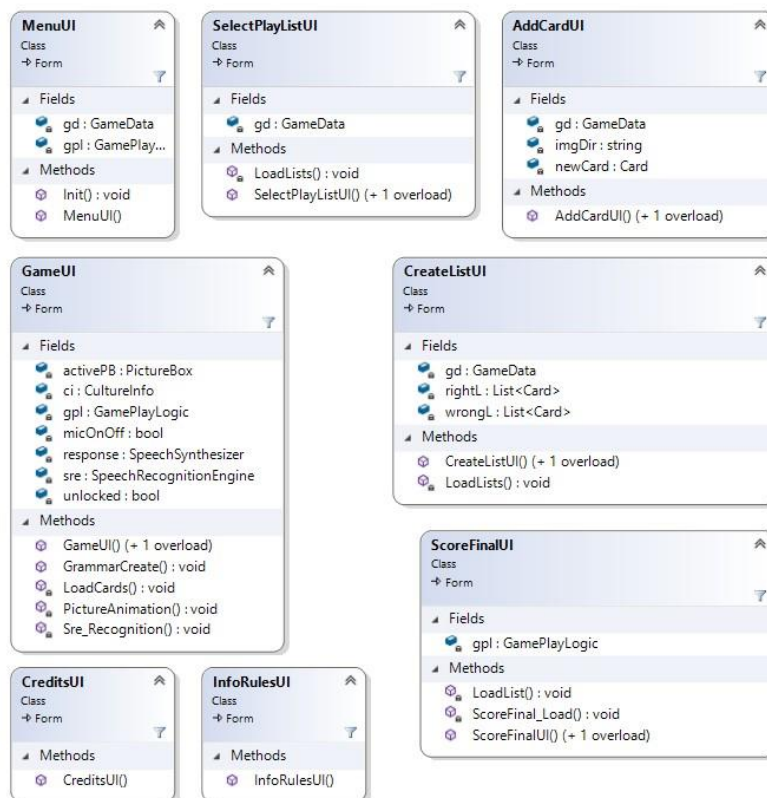


Figura 21 – Classes gráficas do protótipo

Estas classes são as responsáveis pela interação com o utilizador. Cada uma representa uma vista e vai lidar com as classes lógicas para mostrar a informação adequada a cada situação.

A classe “MenuUI” é a que representa o menu inicial da aplicação e é através dela que se navegará para as outras vistas e se executam as diversas funcionalidades fornecidas pela aplicação.

Começando por explicar o mais básico depois do menu, são as vistas pertencentes às classes “CreditsUI” e “InfoRulesUI”. Estas mostrarão informação estática, respetivamente, sobre os créditos da aplicação e sobre informações importantes e regras de utilização.

Para a funcionalidade de adicionar cartas, será implementada uma vista através da classe “AddCardUI”. Esta vista recolhe dados inseridos pelo utilizador e adiciona a carta à classe lógica “GameData” para poder ser usada mais tarde nas outras funcionalidades da aplicação.

À semelhança da classe anterior, a classe *“CreateListUI”* também irá recolher a informação necessária, fornecida pelo utilizador, para poder criar uma *GameList* que poderá mais tarde ser usada para iniciar um exercício.

As três classes restantes, *“SelectPlayListUI”*, *“GameUI”* e *“ScoreFinalUI”* são relativas à execução dos exercícios. A primeira vista a ser mostrada será a respetiva à classe *“SelectPlayListUI”*, onde o utilizador poderá escolher uma das listas de exercício gravadas para começar o exercício. De seguida, será aberta uma vista da classe *“GameUI”* que será responsável por mostrar os dados do exercício e comunicar com a camada lógica os resultados obtidos pelo utilizador. Quando a lista terminar, será mostrada a vista representada pela classe *“ScoreFinalUI”* que será responsável por mostrar, listando, os erros do utilizador.

## 4.2 Tecnologias e Linguagens

No presente estudo, foi feita uma análise de várias tecnologias de reconhecimento de voz com intuito de facilitar a decisão sobre qual utilizar no desenvolvimento do protótipo. As informações recolhidas foram cruzadas com os requisitos levantados e com a análise conjunta com profissionais de terapia da fala. Estas informações, explicitas no Capítulo 3, permitiram que se chegasse à conclusão de que para este projeto seria utilizada a tecnologia Microsoft *Speech* API. Esta decisão incidiu no fácil desenvolvimento e integração com sistemas Windows. A linguagem de programação mais acessível para utilizar com esta tecnologia é a linguagem utilizada pela Microsoft, Visual C#. Como alternativa poder-se-ia também ter utilizado C/C++, JavaScript e Visual Basic.

## 4.3 Desenvolvimento

A atual secção descreve os pontos fulcrais do desenvolvimento do protótipo da solução. Este desenvolvimento é o resultado de uma extensa análise sobre uma solução com o objetivo de se tornar útil num cenário real de uma sessão de terapia da fala. Nas subsecções que se seguem é explicada a implementação dos mecanismos chave desta solução e apresentados os resultados finais.

### 4.3.1 Reconhecimento de voz

A solução analisada no presente estudo faz uso de tecnologias de reconhecimento de voz. Como tal é importante perceber como funciona o sistema de reconhecimento de voz utilizado.

A tecnologia escolhida, tal como referido anteriormente, foi a Microsoft *Speech* API. A sua integração com C# é bastante intuitiva e fácil. Com esta tecnologia, para criar um motor de reconhecimento de voz, é preciso primeiro definir o idioma que se pretende que o mesmo reconheça. Esta definição é feita através da instanciação de um objeto estático *CultureInfo*, que no caso deste protótipo passa como argumento do construtor “pt-PT” para português de Portugal. Depois disto é então criado o motor de reconhecimento de voz com a informação necessária, como pode ser visto na linha 134 da Figura 22. Falta agora definir a gramática para o motor de reconhecimento de voz utilizar. Esta gramática representa as palavras que o motor de reconhecimento de voz vai identificar. Como tal, é necessário ir buscar esta informação à camada lógica, tal como pode ser observado na linha 143 da Figura 22 abaixo, com a expressão “`gp1.GetCurrentPairNames().ToArray()`”.

O objeto “gp1” é uma instância da classe “*GamePlayLogic*” e nesta expressão estamos a aceder ao nome de cada carta no par de cartas para o exercício em execução e transformá-lo num objeto do tipo *Array*. Este objeto é usado para construir um objeto do tipo *GrammarBuilder* (linhas 145 e 146 da Figura 22) que é então usado para a construção do objeto do tipo *Grammar* (linha 150 da Figura 22 abaixo). Finalmente, na linha 155, essa gramática, o objeto do tipo *Grammar*, é carregada para o motor de reconhecimento de voz para poder ser utilizada.

Este código é o responsável pelo reconhecimento de voz, mas agora é preciso dizer à aplicação o que fazer sempre que detetar alguma das palavras correspondentes. Para isso, é acrescentado o método “*Sre\_Recognition*” que é responsável por lidar com o evento de fala reconhecida. Este método é explicado no subcapítulo seguinte pois faz parte integrante da lógica por detrás dos exercícios que o utilizador terá para executar.

```

130 public void GrammarCreate()
131 {
132     try
133     {
134         sre = new SpeechRecognitionEngine(ci);
135     }
136     catch (Exception ex)
137     {
138         MessageBox.Show("ERRO ao integrar lingua escolhida:" + ex.Message);
139     }
140
141     var grammar = new Choices();
142
143     grammar.Add(gpl.GetCurrentPairNames().ToArray());
144
145     var gb = new GrammarBuilder();
146     gb.Append(grammar);
147
148     try
149     {
150         var g = new Grammar(gb);
151
152         try
153         {
154             sre.RequestRecognizerUpdate();
155             sre.LoadGrammarAsync(g);
156
157             sre.SpeechRecognized += Sre_Recognition;
158
159             sre.SetInputToDefaultAudioDevice();
160             response.SetOutputToDefaultAudioDevice();
161             sre.RecognizeAsync(RecognizeMode.Multiple);
162         }
163         catch (Exception ex)
164         {
165             MessageBox.Show("ERRO ao criar reconhecedor: " + ex.Message);
166         }
167     }
168     catch (Exception ex)
169     {
170         MessageBox.Show("ERRO ao criar a gramática: " + ex.Message);
171     }
172 }

```

Figura 22 – Método que gera a gramática

### 4.3.2 Lógica do exercício

O exercício que o utilizador terá que realizar apresenta duas imagens e com o acompanhamento de um profissional terapeuta da fala, o utilizador terá que tentar dizer a palavra que o terapeuta considera certa. Como foi explicado no subcapítulo anterior, o motor de reconhecimento de voz estará apenas à escuta das palavras representadas pelas imagens. Sempre que uma for deteta é lançado o evento que é tratado pelo método representado na Figura 23 abaixo.

A palavra que é detetada pelo motor de reconhecimento de voz é obtida através do código “`string phrase = e.Result.Text;`”. Com isto pode ser usada nas validações seguintes.

A primeira condição (linhas 180 e 194 da Figura 23 abaixo) garante que a palavra captada é igual a uma das cartas do par atual do exercício e que este não se encontra bloqueado pela variável “*unlocked*”. Este mecanismo de bloqueio foi implementado para evitar que depois de ter sido captada uma palavra e tomada a ação de resposta visual, não pudesse ser dita a outra pelo utilizador. Deste modo a captação existe apenas uma vez por cada par de cartas.

Se o código passar a primeira condição passa à segunda (linhas 182 e 196 da Figura 23 abaixo). É neste bloco de código que é feita a validação se a opção dita é a que se pretende que seja a correta ou não. Mediante o resultado é dada resposta visual ao utilizador através do método “*PictureAnimation()*”, no qual são passados os argumentos para alterar a imagem. Caso seja a opção considerada errada pelo profissional terapeuta da fala, é registado pelo código das linhas 189 e 203 da Figura abaixo. Por fim é bloqueada a interface de reconhecimento de voz pela atribuição do valor “*false*” à variável “*unlocked*”.

```
175 void Sre_Recognition(object sender, SpeechRecognizedEventArgs e)
176 {
177     string phrase = e.Result.Text;
178
179     if (phrase.Equals(gpl.GetCardName(0)) && unlocked)
180     {
181         if (gpl.GetCurrentRight() == 0)
182         {
183             PictureAnimation(gameImage1, Color.Green);
184         }
185         else
186         {
187             PictureAnimation(gameImage1, Color.Red);
188             gpl.SetCurrentWrong();
189         }
190         unlocked = false;
191     }
192
193     else if (phrase.Equals(gpl.GetCardName(1)) && unlocked)
194     {
195         if (gpl.GetCurrentRight() == 1)
196         {
197             PictureAnimation(gameImage2, Color.Green);
198         }
199         else
200         {
201             PictureAnimation(gameImage2, Color.Red);
202             gpl.SetCurrentWrong();
203         }
204         unlocked = false;
205     }
206 }
207
```

Figura 23 – Método que responde ao evento de voz reconhecida

### 4.3.3 Protótipo final

Após o término da fase de desenvolvimento do protótipo da solução proposta, resultou um protótipo funcional, capaz de demonstrar as capacidades da tecnologia e que cumpre os requisitos estabelecidos.

Para além dos requisitos e da proposta final definida, foram tomadas decisões face à interface que são melhor explicadas ecrã a ecrã.



Figura 24 – Ecrã de entrada do EasyTEFA

O ecrã de entrada da aplicação, representado na Figura 24 acima, tem o título com um tipo de letra escolhido de modo a ser atraente a uma criança, daí ser mais arredondado. O tipo de letra presente nos botões também foi escolhido de forma a ser mais cativante e menos sério, sendo utilizado também para algum texto noutras secções com o objetivo de manter a interface gráfica uniforme por toda a aplicação. Os botões abrem outras secções da aplicação descritas abaixo:

- Ecrã de informação e regras (Figura 25): Neste ecrã é onde se encontra a informação relativa ao uso da aplicação e às regras do exercício. É explicado o processo de adicionar carta, adicionar lista de jogo e como é suposto o exercício funcionar. No texto explicativo é usado outro tipo de letra para facilitar a leitura do mesmo. Como o texto é extenso, foi usado um elemento de interface gráfica com barra de deslocamento para não ser necessária uma janela muito grande. Ao mesmo tempo, mudando o texto, não é necessário ajustar mais nada na interface.

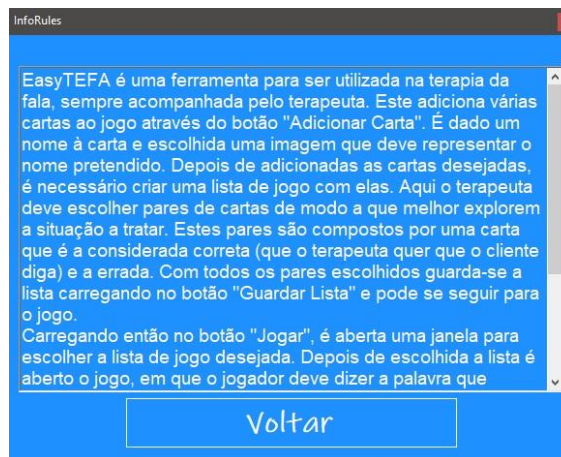


Figura 25 – Ecrã de Informação e Regras

- Adicionar carta (Figura 26 e 27): O ecrã de adicionar carta é a interface que permite ao terapeuta, como utilizador, adicionar cartas que considere que precisa de praticar com paciente. Na Figura 26 está representado o ecrã quando é aberta pela primeira vez a funcionalidade adicionar carta. O botão para escolher a imagem da carta segue o padrão de colocar “...” para o utilizador saber que vai escolher o ficheiro através de uma janela do explorador de ficheiros.

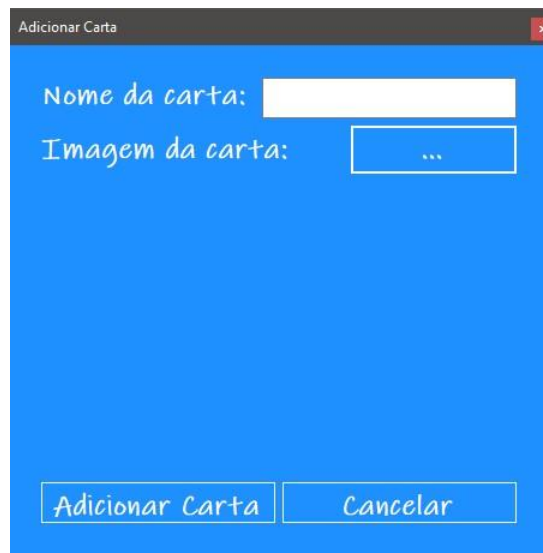


Figura 26 – Ecrã da funcionalidade adicionar carta

Na Figura 27 é demonstrado como ficará o ecrã de adicionar carta preenchido. Como se pode constatar o espaço em branco na figura anterior serve se espaço para uma previsualização da imagem que o utilizador escolheu.



Figura 27 – Ecrã da funcionalidade adicionar carta preenchido

- Adicionar lista de jogo (Figura 28 e 29): É através deste ecrã que o terapeuta, como utilizador, pode adicionar uma lista de jogo, com vários pares de cartas para o paciente ter mais que um exercício para executar. Na figura seguinte vemos como é o ecrã imediatamente após ser aberto. O campo “Nome da lista” fica em cima e à esquerda para que a experiência do utilizador seja uniforme face a tarefas similares que são executadas na aplicação.

Na lista com título “Todas as cartas”, aparecem todas as cartas carregadas em sistema. Selecionando uma delas e carregando no botão “-> Corretas”, é acrescentada aquela carta para a lista das cartas corretas. Se se carregar no botão “-> Erradas” acontece o mesmo para a lista “Cartas Erradas”.

Em cada uma dessas listas, se o utilizador selecionar uma das cartas adicionadas, pode eliminá-la da lista ao carregar de seguida no botão do “X”. O “X” é associado com eliminar, sendo essa a razão para o escolher como face do botão.

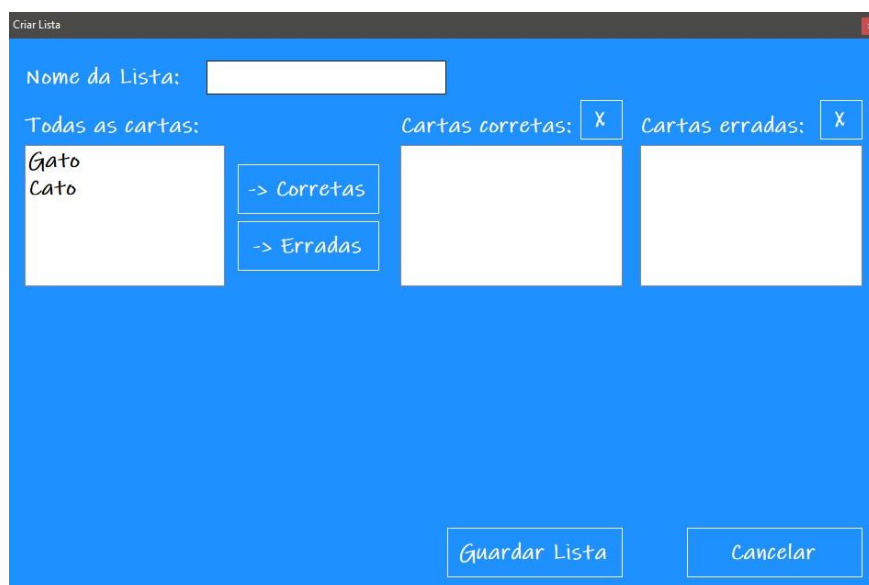


Figura 28 – Ecrã da funcionalidade adicionar lista de jogo

Na Figura 29 abaixo podemos ver o ecrã da funcionalidade adicionar lista de jogo preenchido como exemplo. Se o utilizador quiser guardar a lista basta carregar no botão “Guardar lista”.



Figura 29 – Ecrã da funcionalidade adicionar lista de jogo preenchido

- Jogar (Figuras 30, 31, 32, 33, 34 e 35): Ao carregar no botão “Jogar”, aparece um ecrã com uma lista que permite ao terapeuta escolher a lista de jogo com que o paciente

vai jogar. Como se trata de uma janela pequena, optou-se por dispor os botões organizados verticalmente.

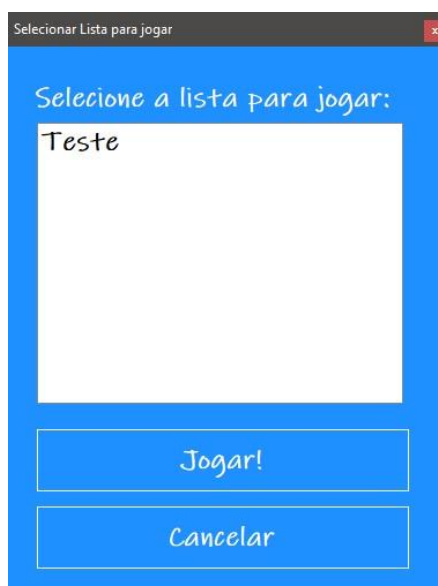


Figura 30 – Ecrã de escolha de lista de jogo

Após se escolher a lista de jogo e clicar em “Jogar” é aberto o ecrã de jogo, representado na figura abaixo. Tendo em conta a proposta final de solução, esta foi a interface escolhida: uma imagem ao lado de outra, com espaço em baixo para os botões de parar o microfone e de exercício seguinte. Optou-se por usar símbolos para os botões por indicarem funções simples de explicar deste modo.

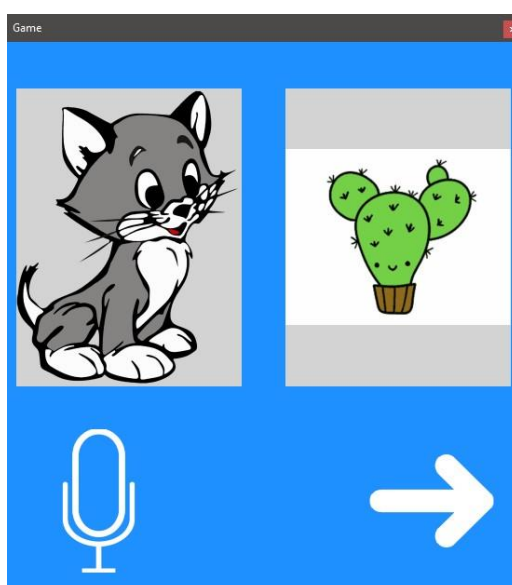


Figura 31 – Ecrã do jogo

Quando o utilizador responde a palavra que o terapeuta pretendia que fosse dita é dado uma resposta visual, mudando o fundo da imagem para verde, como apresentado na figura seguinte.

Foi escolhida a cor verde para assinalar a resposta correta por ser uma cor associada em muitos outros sistemas a uma resposta certa e assim ser mais intuitiva para os utilizadores.

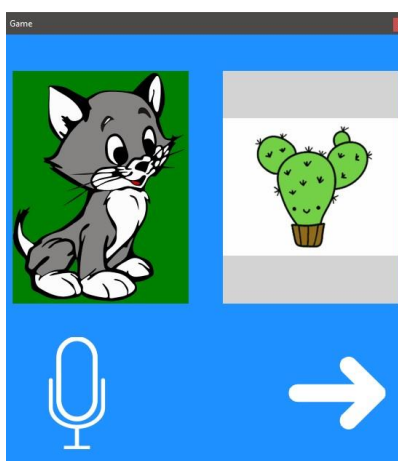


Figura 32 – Ecrã de jogo quando o utilizador acerta

Por outro lado, quando o paciente diz a palavra incorreta, aparece na resposta visual uma alteração da cor de fundo da imagem para vermelho, como se pode ver na Figura 33. A escolha foi semelhante à anterior mas neste caso porque o uso do vermelho é usualmente utilizado para representar a falha ou a resposta errada.

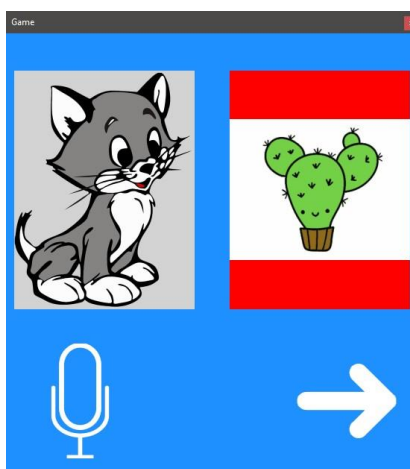


Figura 33 – Ecrã de jogo quando o utilizador erra

Por fim, quando a lista de jogo termina, é apresentado num novo ecrã os resultados, mostrando quais as imagens que deviam ter sido ditas, mas que não foram. Este ecrã pode aparecer a vazio, o que quer dizer que o utilizador não teve erros (Figura 34).

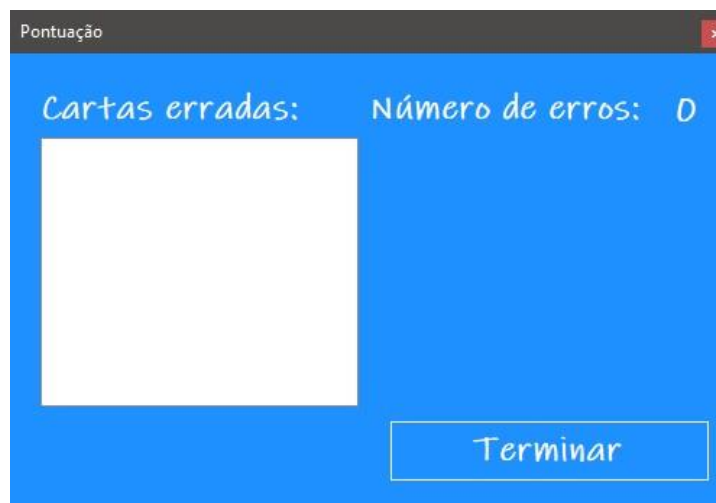


Figura 34 – Ecrã de erros do jogo sem erros

Também pode aparecer com uma lista, exibindo as palavras que o utilizador precisa de treinar e, como demonstrado na Figura 35, selecionando uma dessas palavras, aparece a imagem a ela associada.



Figura 35 – Ecrã de erros do jogo com erros

- Créditos (Figura 36): Por fim, quando se clica no botão “Créditos”, é aberto o ecrã de créditos, com informação sobre o autor e contexto do desenvolvimento da aplicação. Tal como no ecrã das informações e regras, aqui o tipo de letra é diferente para ser mais legível. Foi ainda incluído o logótipo do ISEP por ser a escola onde o presente estudo se desenvolveu e concluiu.



Figura 36 – Ecrã de créditos

#### 4.3.4 Testes preliminares

Após o término da fase de desenvolvimento do protótipo realizamos uma série de testes preliminares que evidenciaram que o protótipo poderia agora seguir para uma avaliação por outros utilizadores.

A avaliação deveria considerar essencialmente dois grandes factos:

- A usabilidade da solução proposta;
- A funcionalidade da mesma.

No capítulo seguinte, apresenta-se o processo de avaliação do EasyTEFA



# Capítulo 5 - Avaliação do EasyTEFA

*No presente capítulo descreve, analisa e exhibe os resultados a avaliação de usabilidade e funcional do protótipo. Com toda esta informação, é ainda apresentado neste capítulo uma apreciação final.*

## 5.1 Contextualização

A avaliação de uma solução é extremamente importante e reflete o uso de boas práticas de engenharia.

Mais do que saber se funciona é preciso identificar claramente como se comporta perante os utilizadores e validar se cumpre os requisitos funcionais para o propósito com que foi criada.

Nos pontos seguintes apresentam-se a avaliação dos resultados.

Com a finalidade de se realizar a avaliação do EasyTEFA, foram planeadas algumas sessões onde um grupo de 23 utilizadores foi convidado a utilizar e a testar a aplicação desenvolvida para suportar o presente estudo.

A cada um dos 23 inquiridos foi pedido que executassem uma série de tarefas de acordo com o Anexo 1 e que em seguida, mudassem de computador para responder de imediato às questões colocadas.



Figura 37 – Mecanismo adotado para a sessão de testes e avaliação do protótipo

As sequências das atividades focalizadas no teste de usabilidade eram dadas ao utilizador através de um tablet enquanto ao mesmo tempo era disponibilizado o EasyTEFA para as realizar.

Como se apresenta na figura 38 cada teste decorreu com a disponibilização ao utilizador do teste de usabilidade (Anexo 2) e com a aplicação EasyTEFA num computador.



Figura 38 – Procedimento do teste de usabilidade

Após esta etapa, como referido passavam para um outro computador para responder ao questionário (Anexo 1).

De salientar que antes de passar à resposta do questionário (figura 39) os utilizadores foram convidados assumir o papel de paciente e a testar a aplicação nessa perspetiva.

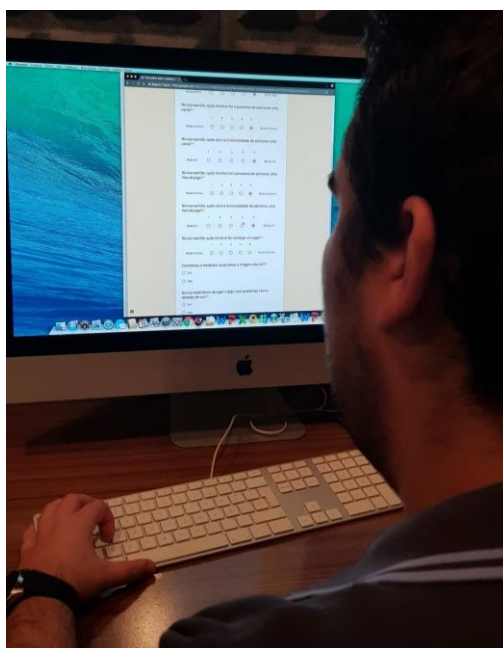


Figura 39 – Processo de respostas após os testes de usabilidade e funcional

Neste grupo alvo participaram 30% de utilizadores do género Feminino e 70% utilizadores do género Masculino tal como se apresenta no Gráfico 1.

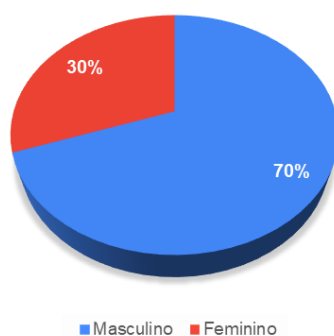


Gráfico 1 – Distribuição por género dos inquiridos

Dado que os inquiridos foram essencialmente estudantes do Mestrado em Engenharia Informática do ISEP tratou-se de um grupo com idades compreendidas entre os 20 e os 28, distribuídas de acordo com o Gráfico 2.

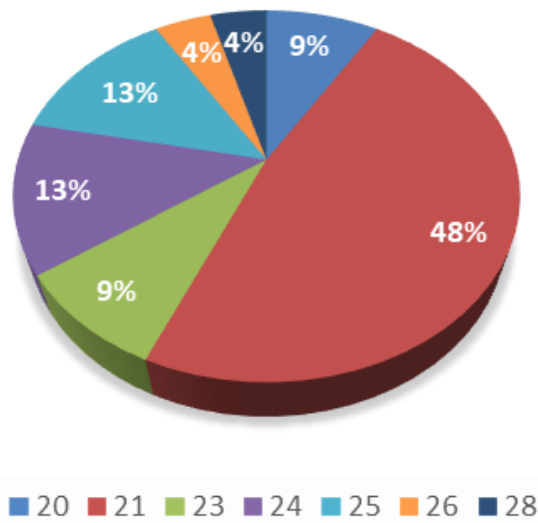


Gráfico 2 - Idade dos inquiridos

Pode-se acrescentar ainda que 100% dos inquiridos possuía o grau de licenciado.

Numa escala de 1 a 5, em que 5 significa muito experiente a utilizar um computador e 1 significa nada experiente a utilizar um computador, 74% dos inquiridos respondeu que é muito experiente e 26% respondeu como sendo experiente, tal como é apresentado no seguinte gráfico.

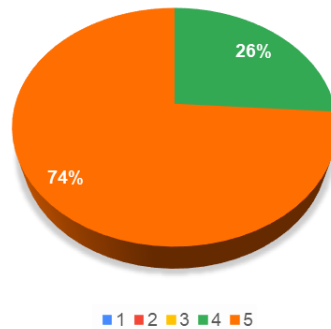


Gráfico 3 – Experiência dos inquiridos com computadores

Devido ao foco da aplicação ser ajudar pessoas com problemas da fala, o grupo de inquiridos foi questionado se tem ou já teve problemas da fala.

Tal como demonstrado no gráfico 4, 9% responderam que sim e os restantes 91% que não.



Gráfico 4 – Identificação de problemas de fala nos inquiridos

## 5.2 Avaliação de usabilidade

Considerando que a aplicação EasyTEFA é uma solução desenvolvida para ser usada quer pelos profissionais de terapia da fala quer pelos seus pacientes para além da usabilidade seria necessário validar os parâmetros funcionais.

No contexto de terapia da fala, o profissional irá utilizar os menus de navegação para adicionar cartas e listas de jogo, que são compostas por cartas. Quando estiver pronto a dar início à resolução dos exercícios, o profissional terá de escolher uma das listas de jogo que criou anteriormente. Quando o paciente terminar os exercícios, é apresentada a lista de erros referentes à lista que se acabou de jogar, com o propósito de, em conjunto com o profissional, trabalhar para melhorar o problema a ser tratado.

Ao grupo de inquiridos foi fornecido uma guião (Anexo 2) com indicações para realizar uma série de procedimentos sobre a aplicação.

Ao cumprir os passos presentes no documento, o grupo teve uma experiência de utilização como se fosse o terapeuta a criar dados para os exercícios e, posteriormente, foi-lhes solicitado que simulassem o papel de “paciente” e que executassem alguns exercícios.

Analisando as respostas à primeira pergunta sobre usabilidade: - “Na sua opinião, quão fácil foi navegar pelos menus da aplicação?”, em que 1 significava muito difícil e 5 muito fácil, verificou-se que 91% dos inquiridos considerou que foi muito fácil e 9% refere que foi fácil (Gráfico 5).

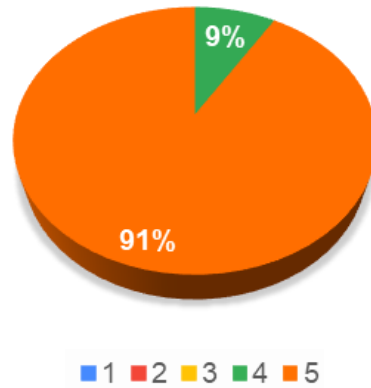


Gráfico 5 – Facilidade de navegação nos menus da aplicação

Pretendíamos avaliar se a interface era intuitiva em alguns processos pelo que a pergunta seguinte foi: - “Na sua opinião, quão intuitivo foi o processo de adicionar uma carta?”.

As respostas, representadas no Gráfico 6, refletem que 83% dos inquiridos considerou que o processo de adicionar uma carta foi muito intuitivo e 17% referem que o processo é intuitivo.

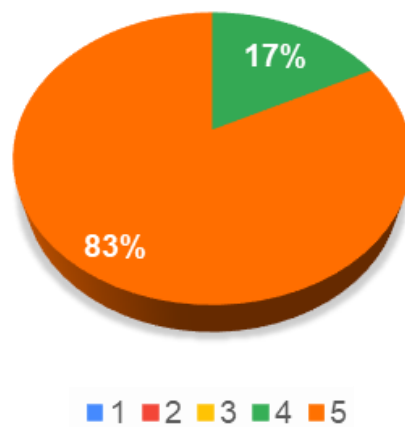


Gráfico 6 – Simplicidade do processo de adicionar uma carta

Relativamente ao mesmo processo ainda se perguntou: - “Na sua opinião, quão útil é a funcionalidade de adicionar uma carta?”. As respostas, representadas no Gráfico 7, refletem que 83% dos inquiridos considerou que a funcionalidade de adicionar uma carta foi muito útil e 17% referem que a funcionalidade útil.

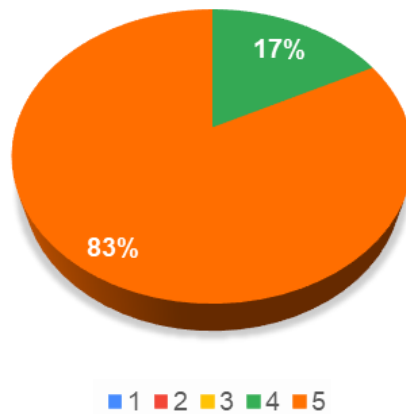


Gráfico 7 - Utilidade da funcionalidade adicionar carta

Dado que seria necessário criar uma lista de cartas para suportar a terapia, questionamos o seguinte: - “Na sua opinião, quão intuitivo foi o processo de adicionar uma lista de jogo?”.

As respostas, representadas no Gráfico 8, refletem que pouco mais de metade dos inquiridos, 57%, considerou o processo muito intuitivo e 35% considerou como intuitivo. No entanto, 4% dos inquiridos respondeu como sendo pouco intuitivo e outros 4% como moderadamente intuitivo.

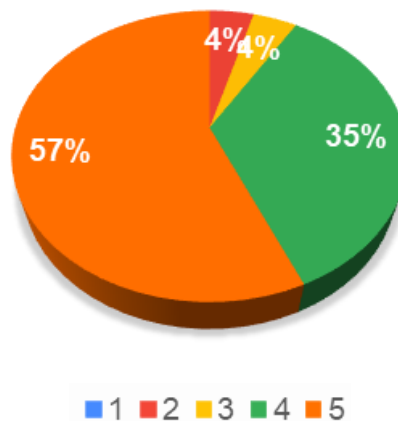


Gráfico 8 – Simplicidade do processo de adicionar uma lista de jogo

Relativo ao mesmo processo, ainda foi questionado: - “Na sua opinião, quão útil foi o processo de adicionar uma lista de jogo?”.

As respostas, representadas no Gráfico 9, refletem que 78% dos inquiridos considerou que a funcionalidade de adicionar uma lista de jogo foi muito útil e 22% consideraram a funcionalidade útil.

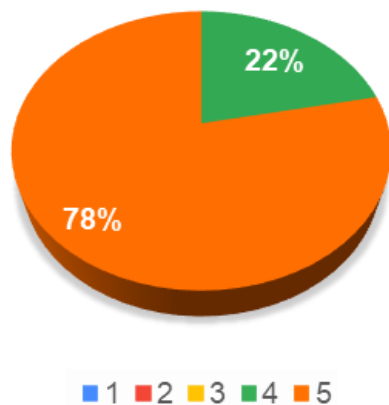


Gráfico 9 – Utilidade da funcionalidade adicionar lista de jogo

Para começar um jogo/terapia, é necessário escolher a lista adequada e dar início ao exercício, como tal foi perguntado: “Na sua opinião, quão intuitivo foi começar um jogo?”

O Gráfico 10, evidencia que a grande maioria dos inquiridos, 96%, considerou o processo muito intuitivo e 4% o considerou como intuitivo.

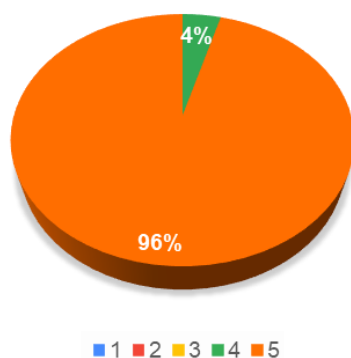


Gráfico 10 – Simplicidade do processo de começar um jogo / terapia

### 5.3 Avaliação funcional

Na continuidade dos testes feitos, o grupo de 23 utilizadores convidados testou também a funcionalidade da aplicação EasyTEFA. Para o efeito, como referido anteriormente assumiram agora, o papel de “paciente”.

O reconhecimento de voz é uma funcionalidade chave desta aplicação e por isso foi questionado: - “Na sua experiência de jogar o jogo, teve problemas com a deteção de voz?”. Dos inquiridos, 74% responderam que não tiveram problemas com a deteção de voz durante o jogo e 26% referiram que sim, como se pode constatar no Gráfico 11.

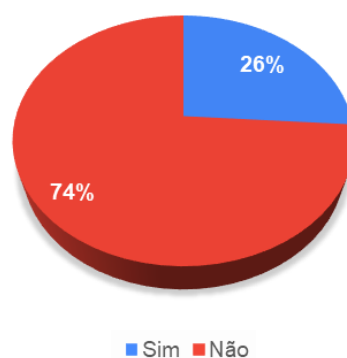


Gráfico 11 – Problemas com a deteção de voz

No caso dos inquiridos que tiveram problemas com a deteção de voz, foi feita também uma pergunta para despiste de problemas. Verificou-se então que 50% dos afetados se encontravam num ambiente com barulho ambiente, os outros 50% encontravam-se num ambiente silencioso. As respostas estão ilustradas pelo Gráfico 12.

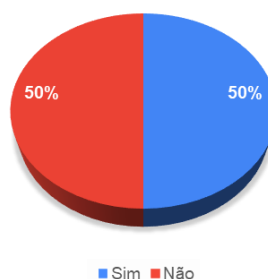


Gráfico 12 – Dos problemas com a deteção, os que se encontravam num ambiente ruidoso

O uso do EasyTEFA tem como objetivo focar na componente educativa e de aprendizagem, sendo assim incluído um ecrã final, após todos os exercícios, que mostra os erros do paciente.

Nesse contexto, foi questionado aos inquiridos: - “Tendo em conta o contexto educativo e de aprendizagem associado ao jogo, o quão importante considera o painel final com os erros?”. As respostas, representadas no Gráfico 13, refletem que 78% dos inquiridos considerou que o painel final com os erros foi muito importante, 18% consideraram a funcionalidade importante e 4% considerou a presença do painel final com os erros moderadamente importante.

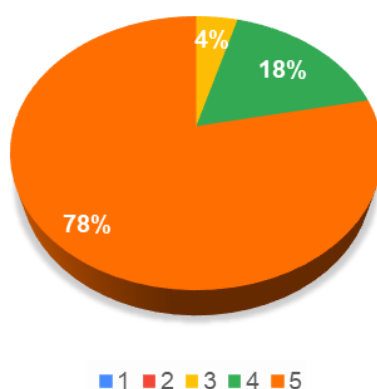


Gráfico 13 – Importância do painel final com os erros

## 5.4 Apreciação final após a avaliação

Considerando as respostas dos testes feitos pelo grupo de inquiridos, podem-se tirar algumas ideias de como trabalhar e melhorar a aplicação no futuro. Embora na maior parte dos casos não tenha aido respostas completamente negativas, há alguns casos que podem ser preocupantes e deverão ser analisados.

No processo de adicionar lista de jogo, houve uma percentagem elevada de utilizadores que não considerou o processo muito intuitivo. Houve mesmo respostas a considerar o processo pouco e moderadamente intuitivo. Se estes dados forem cruzados com as respostas à questão do quão útil é a funcionalidade de adicionar uma lista de jogo, podemos ver que não se relacionam. Disto pode-se concluir que a interface gráfica para este processo pode e deve ser melhorada.

Relativamente aos dados de erros com a deteção de voz, a taxa de acerto é bastante boa para efeitos de um protótipo, especialmente tendo em conta que metade das pessoas que tiveram problemas se encontravam num ambiente ruidoso. Estes resultados reforçam que os

exercícios devem ser executados num ambiente silencioso e de preferência com um microfone bom, de modo a diminuir as camadas de ruído com que o motor de reconhecimento de voz tem de lidar, facilitando assim o reconhecimento.

Podemos ainda considerar o uso de *feedback* visual para mostrar ao “paciente” se a palavra que disse era a correta ou não bastante importante com base nas repostas à questão: - “Considerou o *feedback* visual sobre a imagem dita útil?”, que foram de 100% dos inquiridos referir que o *feedback* visual era útil.

De forma idêntica à usabilidade do processo adicionar lista de jogo, o painel final dos exercícios, que mostra os erros, foi considerado moderadamente importante por 4% dos inquiridos. Estes valores devem ser lidos de um modo crítico e com a finalidade de melhorar a solução.

Finalmente, como 100% inquiridos respondeu que sim à questão: - “Recomendaria esta aplicação a alguém?”, podemos concluir que, embora tenha alguns pontos a melhorar, o protótipo da aplicação EasyTEFA está bem encaminhado e tem potencial para evoluir.



## Capítulo 6 - Conclusões e trabalho futuro

O objetivo desta dissertação era o estudo de uma possível solução na área da saúde, mais concretamente na terapia da fala, que captasse e reconhecesse a voz de um paciente. Inicialmente a proposta de solução passava por escrever uma frase que o paciente teria que ler e a aplicação identificaria os erros ditos, salientando-os a vermelho. Na continuidade do estudo e com o contato mantido com terapeutas da fala da Escola Superior de Saúde, a solução foi ajustada para melhor conviver num contexto de terapia da fala e assim poder ser mais útil como ferramenta para um terapeuta. Esta nova proposta de solução consistiu na mostragem de duas imagens que representem palavras parecidas ou com sons parecidos, que o paciente tem dificuldade em dizer. O terapeuta define estas imagens e define também qual delas são consideradas corretas ou erradas. Esta solução foi então mais abrangente que a anterior pois não é necessário o paciente saber ler.

Por via de vir a desenvolver tal solução, foram analisadas as tecnologias de reconhecimento de voz já existentes no mercado e o estado da arte das mesmas. Depois de um período de análise e ser tomada a decisão sobre qual tecnologia utilizar e como implementar a solução, o protótipo começou a aproximar-se do seu estado final.

Finalmente, o protótipo tornou-se funcional e serviu de prova de conceito para a solução pretendida. Através deste protótipo é possível entender, até certo nível, o seu potencial.

## 6.1 Conclusões

Da análise feita após o desenvolvimento do protótipo da solução conclui-se que este ainda pode evoluir mais e ser orientado para casos de uso diferentes. A solução pode ser adaptada para adultos, embora o foco inicial tenha sido as crianças.

Concluiu-se também que ainda pode ser feito algum trabalho para melhorar a experiência do utilizador final, principalmente a nível de interface gráfica.

## 6.2 Trabalho futuro

O trabalho futuro desta solução pode passar pela utilização de uma API de reconhecimento de voz diferente, paga, para tentar obter melhores resultados e menos falsos positivos. Outra melhoria desta solução passa por não limitar a solução aos sistemas Windows.

De um ponto de vista de funcionalidades, há muitas novas que podem ser implementadas de modo a facilitar o trabalho dos terapeutas da fala. Por exemplo: guardar as cartas numa base de dados para não terem que ser reinseridas todas as vezes que se correr a aplicação, criar uma ficha de paciente integrada que guarde as listas específicas e a performance do mesmo para consultas futuras e eventual exportação para um ficheiro Excel ou PDF.

O exercício neste momento apenas mostra se é bem ou mal falada a palavra pretendida. Uma sugestão de trabalho futuro é trabalho no tornar este exercício, ou uma coleção de vários que se podem juntar a este, num jogo sério.

Perspetiva-se a divulgação deste trabalho sob a forma de artigo científico e a sua apresentação em conferências relacionadas com o tema.

Pretendemos evoluir o protótipo e trabalhar com especialistas da terapia da fala para que o usem e nos proporcionem informações que o possam fazer evoluir ainda mais.

# Referências

- Azevedo, A. (2018). Obtido em julho de 2019, de <https://apps.apple.com/us/app/happies/id1224164066?l=pt&ls=1>
- Boas Noticias. (2018). Obtido em julho de 2019, de <https://boasnoticias.pt/jogo-ajuda-criancas-com-dificuldades-de-fala/>
- Carta de Ottawa. (1986). Obtido em fevereiro de 2019, de <https://www.dgs.pt/documentos-e-publicacoes/carta-de-otawa-pdf1.aspx>
- CUF. (s.d.). *Saúde CUF*. Obtido em fevereiro de 2019, de <https://www.saudecuf.pt/areas-clinicas/terapia-da-fala>
- EMS Software. (2015). Obtido em julho de 2019, de <https://apps.apple.com/pt/app/palavras-aos-bocadinhos/id963095946>
- JJMR Productions. (2017). Obtido em julho de 2019, de [https://play.google.com/store/apps/details?id=com.productionsjjmrgmail.consciencia\\_fonologica](https://play.google.com/store/apps/details?id=com.productionsjjmrgmail.consciencia_fonologica)
- Matérias Escolares: Ensino Fundamental e Médio. (2019). Obtido em fevereiro de 2019, de <https://www.materias.com.br/portugues/tipos-de-comunicacao.html>
- OMS. (1946). Obtido em fevereiro de 2019, de <http://www.direitoshumanos.usp.br/index.php/OMS-Organiza%C3%A7%C3%A3o-Mundial-da-Sa%C3%BAde/constituicao-da-organizacao-mundial-da-saude-omswho.html>
- Pequenada. (s.d.). *Terapia da fala para crianças*. Obtido em julho de 2019, de <https://pequenada.com/artigos/terapia-fala-para-criancas>
- Sovie, D., Roberts, G., Murdoch, R., McMahon, L., & Schoelwer, M. (2018). Obtido em fevereiro de 2019, de [https://www.accenture.com/t20180302T094127Z\\_\\_w\\_\\_/us-en/\\_acnmedia/PDF-69/Accenture-2018-Digital-Consumer-Survey-Findings.pdf#zoom=50](https://www.accenture.com/t20180302T094127Z__w__/us-en/_acnmedia/PDF-69/Accenture-2018-Digital-Consumer-Survey-Findings.pdf#zoom=50)
- Terapia da Fala*. (s.d.). Obtido de CUF: <https://www.saudecuf.pt/areas-clinicas/terapia-da-fala>



## Anexo 1 – Formulário / Questionário

### Formulário sobre usabilidade e funcionalidade da aplicação EasyTEFA

Este formulário tenciona recolher dados sobre a usabilidade e funcionalidade da aplicação EasyTEFA

**\*Obrigatório**

Qual é a sua idade? \*

A sua resposta

Qual é o seu género?

- Masculino
- Feminino
- Prefiro não dizer
- Outra:

Na sua opinião, quão experiente é ao utilizar um computador? \*

1 2 3 4 5

Muito pouco experiente

Muito experiente

Tem ou já teve problemas de fala? \*

Selecionar ▼

Na sua opinião, quão fácil foi navegar pelos menus da aplicação? \*

|               |                       |                       |                       |                       |                       |             |
|---------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------|
|               | 1                     | 2                     | 3                     | 4                     | 5                     |             |
| Muito difícil | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Muito fácil |

Na sua opinião, quão intuitivo foi o processo de adicionar uma carta? \*

|                |                       |                       |                       |                       |                       |                 |
|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------|
|                | 1                     | 2                     | 3                     | 4                     | 5                     |                 |
| Nada intuitivo | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Muito intuitivo |

Na sua opinião, quão útil é a funcionalidade de adicionar uma carta? \*

|           |                       |                       |                       |                       |                       |            |
|-----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------|
|           | 1                     | 2                     | 3                     | 4                     | 5                     |            |
| Nada útil | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Muito útil |

Na sua opinião, quão intuitivo foi o processo de adicionar uma lista de jogo? \*

|                |                       |                       |                       |                       |                       |                 |
|----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------|
|                | 1                     | 2                     | 3                     | 4                     | 5                     |                 |
| Nada intuitivo | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Muito intuitivo |

Na sua opinião, quão útil é a funcionalidade de adicionar uma lista de jogo? \*

|           |                       |                       |                       |                       |                       |            |
|-----------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------|
|           | 1                     | 2                     | 3                     | 4                     | 5                     |            |
| Nada útil | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | Muito útil |

Na sua opinião, quão intuitivo foi começar um jogo? \*

1      2      3      4      5

Nada intuitivo                                    Muito intuitivo

Considerou o feedback visual sobre a imagem dita útil? \*

- Sim
- Não

Na sua experiência de jogar o jogo, teve problemas com a deteção de voz? \*

- Sim
- Não

Se respondeu sim na questão anterior, encontrava-se num ambiente com barulho ambiente? \*

- Sim
- Não
- Respondi não na questão anterior

Tendo em conta o contexto educativo e de aprendizagem associado ao jogo, o quão importante considera o painel final com os erros? \*

1      2      3      4      5

Nada importante                                    Muito importante

Recomendaria esta aplicação a alguém?

- Sim
- Não



# **ANEXO 2 - Avaliação de usabilidade e funcionalidade do EasyTEFA – Solução para apoio à terapia da fala**

Estudo da usabilidade da aplicação

**Rúben Magina Pereira Vieira Ribeiro**

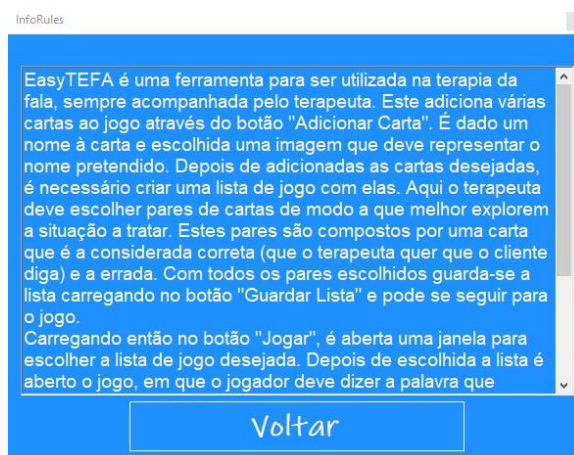
## Início da aplicação

1. Inicie a aplicação clicando no executável EasyTEFA.exe na pasta já aberta deste computador;
2. É aberto o ecrã inicial da aplicação.



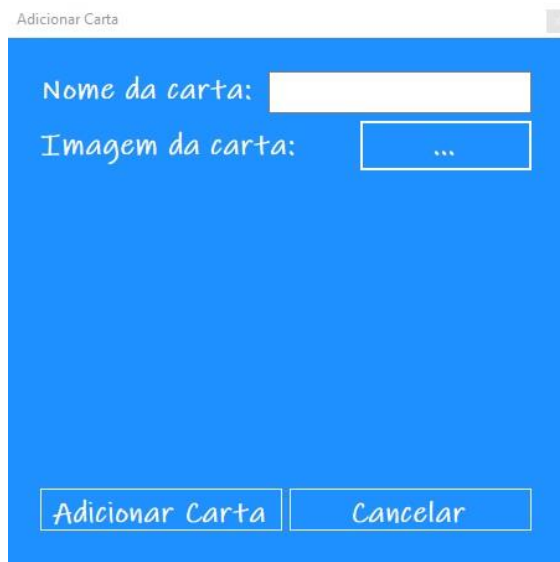
## Ecrã de informação e regras

1. Carregue no botão que diz "Infor e Regras"
2. Será aberta uma janela com a informação e regras da aplicação
3. Depois de as ler, carregue no botão "Voltar"



## Adicionar carta

1. Carregue no botão que diz “Adicionar Carta”
2. Será aberta uma nova janela com a seguinte interface



3. Nesta janela insira o nome para a carta, que pretende que seja o nome a ser reconhecido durante o jogo.

Para efeitos de teste escolha um dos seguintes:

- Gato
- Cato
- Bolacha
- Borracha
- Faca
- Vaca
- Pão
- Cão

4. Depois carregue no botão “...” para escolher uma imagem correspondente ao nome.
5. Depois de escolhida carregue no botão “Adicionar Carta”

## Adicionar lista de jogo

1. Carregue no botão “Adicionar Lista de Jogo”
2. Será aberta uma nova janela com a seguinte interface

Criar Lista

Nome da Lista:

Todas as cartas:

-> Corretas

-> Erradas

Cartas corretas:  X

Cartas erradas:  X

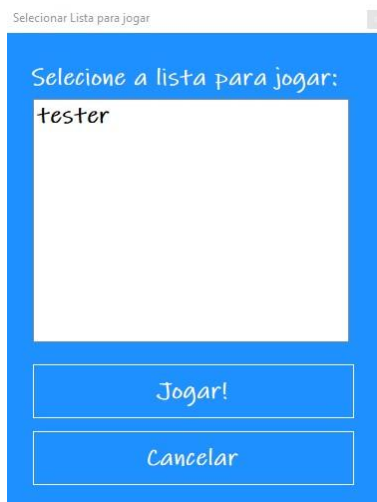
Guardar Lista

Cancelar

1. Insira um nome para a lista
2. Depois, selecione uma carta da lista “Todas as cartas” e clique em “-> Corretas”. A carta selecionada aparecerá na lista “Cartas corretas”
3. Repita o processo de carregar no botão “-> Corretas”
4. Agora, na lista “Cartas corretas”, selecione uma das cartas repetidas e carregue no botão “X” em cima da lista.
5. Selecione outra das cartas da lista “Todas as cartas” e carregue em “-> Erradas”. A carta selecionada aparecerá na lista “Cartas erradas”
6. Repita os pontos 4.5 e 4.6 para a lista de “Cartas erradas”
7. Finalmente, carregue no botão “Guardar Lista”

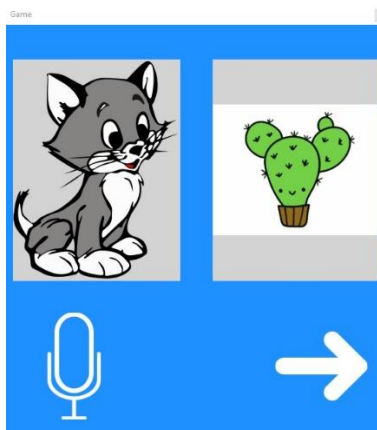
## Ecrã de seleção de lista de jogo

1. Carregue no botão “Jogar”
2. Será aberta uma nova janela com a seguinte interface



## Ecrã do jogo

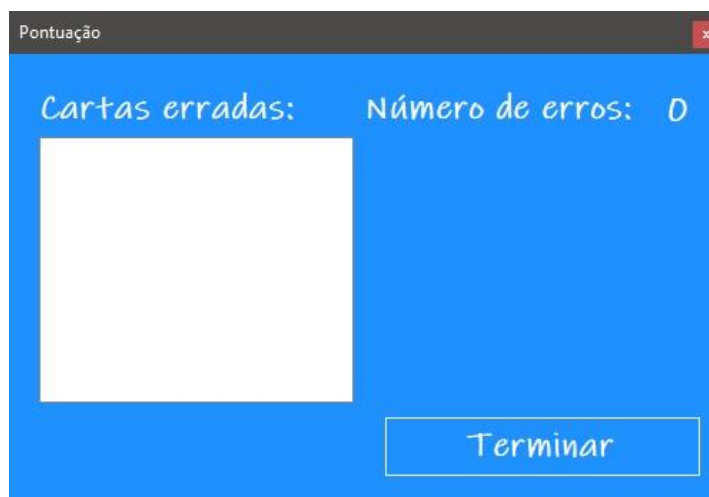
1. Na janela de seleção de lista de jogo, selecione a lista “Teste” e carregue no botão “Jogar!”
2. Será aberta uma nova janela com a seguinte interface



3. Diga a palavra que descreve uma das imagens
4. Depois de ter sido detetada corretamente, carregue no botão com a seta
5. Repita os passos 6.3 e 6.4 até ser aberta uma nova janela

## Ecrã de erros

1. Terminando o jogo, deverá ter sido aberta uma janela com uma interface idêntica à seguinte



2. Aqui é demonstrado as cartas erradas. Selecione uma da lista e será apresentada ao lado
3. Carregue em terminar

## Ecrã de créditos

1. No menu inicial novamente, carregue em “Créditos”
2. Será aberta uma janela como a seguinte



3. Leia e carregue em voltar