

**UNIVERSIDADE DO EXTREMO SUL CATARINENSE - UNESC
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

ARISTOTE KANZA LOPES

**PROTÓTIPO DE UMA APLICAÇÃO WEB USANDO RECONHECIMENTO DE VOZ
PARA CONVERTER VOZ EM TEXTO, ORIENTADO AO APRENDIZADO DE
ESCRITA PARA PESSOAS QUE POSSUEM DIFICULDADES NO USO DO
TECLADO**

CRICIÚMA

2018

ARISTOTE KANZA LOPES

**PROTÓTIPO DE UMA APLICAÇÃO WEB USANDO RECONHECIMENTO DE VOZ
PARA CONVERTER VOZ EM TEXTO, ORIENTADO AO APRENDIZADO DE
ESCRITA PARA PESSOAS QUE POSSUEM DIFICULDADES NO USO DO
TECLADO**

Trabalho de Conclusão de Curso, apresentado para obtenção do grau de Bacharel no curso de Ciência da Computação da Universidade do Extremo Sul Catarinense, UNESC.

Orientador: Prof. Esp. Fabricio Giordani

CRICIÚMA

2018

ARISTOTE KANZA LOPES

**PROTÓTIPO DE UMA APLICAÇÃO WEB USANDO RECONHECIMENTO DE VOZ
PARA CONVERTER VOZ EM TEXTO, ORIENTADO AO APRENDIZADO DE
ESCRITA PARA PESSOAS QUE POSSUEM DIFICULDADES NO USO DO
TECLADO**

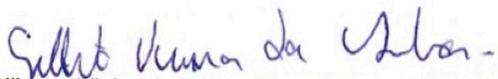
Trabalho de Conclusão de Curso aprovado
pela Banca Examinadora para obtenção do
Grau de Bacharel, no Curso de Ciência da
computação da Universidade do Extremo Sul
Catarinense, UNESC, com Linha de Pesquisa
em Desenvolvimento Web

Criciúma, 28 de novembro de 2018.

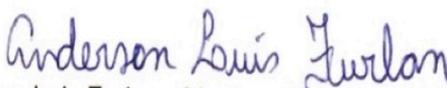
BANCA EXAMINADORA



Prof. Fabrício Giordani - Especialista - (UNESC) - Orientador



Prof. Gilberto Vieira da Silva - Especialista - (UNESC)



Prof. Anderson Luis Furlan - Mestre - (Thomson Reuters/CEDUP)

**Aos meus queridos pais, meus irmãos, meus
amigos que sempre me apoiaram.**

AGRADECIMENTOS

Primeiramente a Deus, que me abençoou durante todo o tempo que estive cursando a faculdade, a minha família que sempre me ajudou e incentivou nos momentos difíceis, meus amigos pelo apoio direta e indiretamente.

Agradeço também ao meu orientador Fabricio Giordani por me orientar e direcionar os caminhos certos para que pudesse desenvolver o trabalho de conclusão de curso, agradeço a todos professores por terem passado seus conhecimentos, e também a todos os colegas que sempre estiveram juntos comigo apoiando e incentivando nos momentos difíceis.

“Para as pessoas sem deficiência, a tecnologia torna as coisas mais fáceis. Para as pessoas com deficiência, a tecnologia torna as coisas possíveis.”

Radabaugh

RESUMO

A aplicação web é uma das tecnologias que possui uma evolução constante, diariamente surgem padrões, tecnologias que facilitam o desenvolvimento das mesmas. Navegadores webs como, Google Chrome, Safari e Mozilla Firefox estão cada vez melhores, facilitando a sua usabilidade, acrescentando recursos e funcionalidades que facilitam seu uso quer ser seja no seio dos desenvolvedores webs ou dos usuários convencionais. Um dos conceitos que tem se beneficiado com essa evolução é a acessibilidade, originando as chamadas Tecnologias Assistivas. Essas tecnologias buscam trazer melhor qualidade de vida e mais acesso a pessoas portadoras de deficiência. Baseado nesse conceito, este trabalho teve como objetivo desenvolver um protótipo de aplicação web para capturar a fala do usuário a partir do microfone no computador e convertê-lo em texto, com intuito de apoiar o aprendizado de escrita de pessoas que possuem dificuldades no uso do teclado convencional. Para isso, foi realizada uma fundamentação teórica referente as técnicas de reconhecimento de fala, responsável por reconhecer a fala do usuário, metodologias de acessibilidade em aplicações web. Com os conhecimentos obtidos a partir da pesquisa, foi possível elaborar um protótipo de aplicação web, utilizando a ferramenta *Web Speech API*, que faz a conversão de voz em texto. Em testes realizados com o protótipo, os resultados apresentados demonstraram que o mesmo é ideal para usuários com alguma deficiência, permitindo que estes convertam a sua fala em texto, fornecendo assim acessibilidade para qualquer usuário.

Palavras-chave: Acessibilidade. Aplicações web. Deficiência nos membros superiores. Reconhecimento de fala.

ABSTRACT

The web application is one of the technologies that has a constant evolution, daily patterns emerge, technologies that facilitate their development. Web browsers such as Google Chrome, Safari and Mozilla Firefox are getting better and better, making it usable to be either web developers or conventional users. Google Chrome is the leader in this technology segment and provides a number of features for developers to create more and more new features. One of the concepts that has benefited from this evolution is the accessibility, originating the so-called Assistive Technologies. These technologies seek to bring about a better quality of life and more access to people with disabilities. Based on this concept, this work aimed to develop a prototype web application to capture the user's speech from the microphone in the computer and convert it into text, in order to support the learning of writing of people who have difficulties in using the conventional keyboard. For that, a theoretical foundation was made regarding the techniques of speech recognition, responsible for recognizing user speech, accessibility methodologies in web applications, among others. With the knowledge obtained from the research, it was possible to elaborate a web application prototype using the Web Speech API tool, which converts from voice to text. In tests performed with the prototype, the results showed that it is ideal for users with a deficiency, allowing them to convert their speech into text, providing accessibility for any user.

Keywords: Accessibility. Web applications. Deficiency in the upper limbs. Speech Recognition.

LISTA DE ILUSTRAÇÕES

Figura 1 - Teclado IntelliKeys	19
Figura 2 - Acionadores com Fios	19
Figura 3 - Mouse Adaptado	20
Figura 4 - Mouse Por Movimento da Cabeça	20
Figura 5 - Reconhecimento de Fala do Windows 7	21
Figura 6 - Reconhecimento de Fala do Google Docs	22
Figura 7 - Tela de Digitação do Dictation Pro	23
Figura 8 - Tela Principal do Speechnotes.....	25
Figura 9 - Arquitetura de Uma Aplicação Web.....	29
Figura 10 - Tecnologias utilizadas na construção de interfaces de voz.....	31
Figura 11- As etapas utilizadas para reconhecimento de fala.	32
Figura 12 - Classes da Web Speech API	44
Figura 13 - Instância do objeto webkitSpeechGrammarList.....	44
Figura 14 - Trecho de código para a inicialização da SpeechRecognition	45
Figura 15 - Método para Alteração da Posição do Cursor.....	47
Figura 16 - Botão sem o mouse por cima & Figura 17 - Botão com o mouse por cima	48

LISTA DE TABELAS

Tabela 1 - Número total de sites por ano.....	24
Tabela 2 - Resultado dos testes de acordo a cada cenário.....	49

LISTA DE ABREVIATURAS E SIGLAS

API	Application Program Interfaces
CAA	Comunicação Aumentativa e Alternativa
CDPD	Convenção Sobre os Direitos da Pessoa com Deficiência
IBDD	Instituto Brasileiro dos Direitos da Pessoa com Deficiência
ITS	Instituto de Tecnologia Social
MVC	Model View Controller
ONU	Organização das Nações Unidas
TA	Tecnologia Assistiva
TIC	Tecnologia de Informação e Comunicação
WSA	Web Speech API
SR	Speech Recognition

SUMÁRIO

1 INTRODUÇÃO	11
1.1 Objetivo Geral	12
1.2 Objetivo Específico	12
1.3 Justificativa	12
1.4 Estrutura do Trabalho	13
2 Deficiência no Contexto social	14
2.1 Sociedade e as pessoas com deficiência.....	14
2.2 ACESSIBILIDADE	15
2.2.1 Acessibilidade e Educação	16
2.2.2 Tecnologia Assistiva Para Deficientes Com Restrições Motoras	17
2.2.3 Acessibilidades em Aplicações Web	23
3 Tecnologia	28
3.1 APLICAções WEB	28
3.1.1 Arquitetura das Aplicações Web	28
3.2 Reconhecimento de Voz.....	30
3.2.1 Etapas do Reconhecimento de Voz	32
3.3 Web Speech API	35
3.3.1 SpeechRecognition	35
4. TRABALHO CORRELATOS	38
4.1 Sistema de Reconhecimento de Fala via Web	38
4.2 Um Software de Reconhecimento de Voz para Português Brasileiro.....	39
4.3 Sistema de Reconhecimento de Voz – Aplicabilidade.....	39
4.4 Tecnologia de ensino e tecnologia assistiva no ensino de crianças com paralisia cerebral.....	40
5 Trabalho Proposto	42
5.1 Metodologia	42
5.1.1 A Tecnologia de Reconhecimento de voz Web Speech API	43
5.1.2 Desenvolvimento do Protótipo	46
5.1.3 Teste de Validação	48
5.2 RESULTADOS OBTIDOS	49
6 CONCLUSÃO	51

1 INTRODUÇÃO

A escrita tem um papel muito importante no desenvolvimento social e intelectual humano. A partir do ingresso na vida escolar, ela se torna imprescindível para o acesso ao conhecimento e a informação. Entretanto, muitas vezes o acesso a esse meio se torna difícil, principalmente para deficientes com coordenação motora reduzida que torna difícil o uso dos membros superiores, seja por falta de recursos tecnológicos, como teclados e mouses especializados, ou algum outro recurso que permita com que ele possa aprender a escrever.

Atualmente, o mercado vem resolvendo vários tipos de problemas por meio digital, e a educação vem acompanhando a tecnologia nesse sentido. Entretanto, existe uma quantidade muito grande de recursos importantes e continuam a ser desenvolvidos recursos para auxiliar no aprendizado ou na possibilidade de escrita de pessoas com mobilidades reduzidas, quer seja recursos físicos ou digital.

Com a evolução das tecnologias, diversos recursos surgem buscando tornar atividades cotidianas mais fáceis, tal como promover alternativas para a resolução de problemas rotineiros. A aplicação desses recursos na promoção da acessibilidade a pessoas portadoras de deficiências recebe o nome de Tecnologia Assistiva. O objetivo dessas tecnologias é a fabricação, uso de equipamentos, recursos ou estratégias utilizadas para capacitar as habilidades funcionais das pessoas com deficiências (BRASIL, 2009).

A popularização e a evolução da tecnologia das aplicações web tornam possíveis a elaboração de novas soluções para auxiliar a execução de tarefas, inclusive na aplicação como Tecnologia Assistiva. Recursos como Reconhecimento de voz e de conversão de áudio em texto, captura de pequenas trechos de fala com precisão, entre outros, tornam possível a elaboração de inúmeros tipos de aplicações, para as mais variadas finalidades.

Este trabalho teve como objetivo utilizar tecnologia de Reconhecimento de voz e conversão de áudio em texto para a elaboração de um protótipo de aplicação web, voltado ao apoio e aprendizado de escrita por parte de usuários com coordenação motoras reduzidas.

1.1 OBJETIVO GERAL

O objetivo deste trabalho é desenvolver um protótipo de aplicação web para o aprendizado de escrita utilizando uma API de reconhecimento de voz e conversão de voz em texto.

1.2 OBJETIVO ESPECÍFICO

Os objetivos específicos aos quais se pretendeu atingir com a elaboração do trabalho são:

- a) compreender metodologias de acessibilidade de pessoas que possuem dificuldades no uso do teclado em computadores;
- b) entender o funcionamento do reconhecimento de voz;
- c) analisar a conversão de voz em texto com a Web Speech API;
- d) investigar formas de interação com o usuário da aplicação;
- e) implementar o protótipo da aplicação;
- f) aplicar teste de eficiência do protótipo.

1.3 JUSTIFICATIVA

Em um sentido amplo percebemos que a evolução tecnológica caminha na direção de tornar a vida mais fácil (BERSCH, 2008). Podendo assim permitir as pessoas com deficiências o acesso aos recursos oferecidos pelas Tecnologias de Informação e Comunicação (TICs), visto que segundo ITS BRASIL (2018), é preciso entender que as pessoas com deficiência querem, antes de tudo, inclusão e direitos. Nesse sentido, em muitos países, as políticas públicas para pessoas com deficiência superaram a visão do chamado “modelo médico” de atendimento e dos enfoques assistencialistas e passaram a adotar os chamados “modelos social”, “modelo dos direitos ou da cidadania” ou “modelo da inclusão ou participação”.

Diversos marcos de declarações de princípios contribuíram para a criação dessa nova sensibilidade mundial. É nesse contexto que as políticas públicas de inserção de pessoas com deficiência em todos os aspectos da vida, com o auxílio da Tecnologia Assistiva (TA), tornam-se extremamente relevantes.

Os *softwares* assistivos para aprendizado existentes, especialmente os de reconhecimento de voz são poucos, difíceis de encontrar, e na sua maioria são feitos

para um tipo de dispositivo (*desktop*) e um tipo de sistema operacional, apesar de aplicativos feitos para desktop terem a vantagem de serem acessados sem a conexão de internet, os aplicativos web são mais eficientes porque podem ser acessados em qualquer parte e em qualquer sistema operacional basta apenas uma conexão de internet e um navegador. Pelo fato de poucos *softwares* voltados para essa área estarem pouco disponíveis, surge a dificuldade na obtenção dos tais *softwares* é o que na maioria dos casos dificulta as instituições de ensino que possui alunos com deficiência ou mesmo os familiares de algumas pessoas com deficiência, onde eles sentem-se obrigados a isolá-lo ou seja colocá-lo de lado nesse aspecto as pessoas que possuem essas dificuldades em interagir com o computador por meio de teclado. Isso podia ser resolvido ou minimizado se existisse uma aplicação web que correspondesse às essas necessidades.

1.4 ESTRUTURA DO TRABALHO

Este trabalho está dividido em uma estrutura com 5 capítulos. Sendo que o primeiro capítulo apresenta uma breve descrição sobre o tema do trabalho, contendo a introdução, o objetivo geral e os específicos e ainda a justificativa do mesmo trabalho.

O segundo capítulo trata dos assuntos relacionados a deficiências no contexto social, englobando a sociedade e a acessibilidade, deficiências com coordenação motora reduzidas, tecnologias assistivas para coordenação motora reduzida e acessibilidade em aplicações web.

O terceiro capítulo aborda o levantamento teórico referente a todas as tecnologias aplicadas no desenvolvimento do trabalho. Dois grandes grupos se destacam nesse capítulo: Aplicações Web, com sua estrutura, arquitetura e funcionamento, e Reconhecimento de Voz, tratando de suas etapas e algumas ferramentas existentes e a Web Speech API, com os seus componentes e utilização.

No quarto capítulo são apresentados trabalhos correlatos ao estudo desenvolvido.

O quinto capítulo traz a estrutura do trabalho desenvolvido, onde aborda-se a metodologia aplicada para o desenvolvimento do protótipo e onde são avaliados os resultados obtidos com a elaboração do mesmo.

2 DEFICIÊNCIA NO CONTEXTO SOCIAL

No Brasil, para que se possa ter uma ideia no contexto social na qual as pessoas com deficiência vivem, teve uma série de propostas de inclusão que foram analisadas pelo governo. A aprovação destas propostas por parte da Assembleia das Nações Unidas¹, da Convenção sobre os Direitos das Pessoas com Deficiência (CDPD) representa uma evolução em todos os setores da política, dos programas e dos projetos que são destinados a um número alto de pessoas com deficiência que vivem ao redor do mundo (GUARALDILAB, 2015). Nesse sentido, os problemas relacionados às pessoas com deficiências acompanham os homens desde os tempos mais remotos da civilização, e hoje com uma inclusão mais presente aos olhos do mundo.

2.1 SOCIEDADE E AS PESSOAS COM DEFICIÊNCIA

A sociedade tem se relacionado com as pessoas com deficiência de maneira diferente, de tal forma que vem sofrendo transformações positivas ao longo dos tempos, principalmente no que diz respeito ao conjunto de atividades, que podem ser os saberes, práticas ou atividades que mediam essa relação (CAMARGO; GOULART JÚNIOR; LEITE, 2017).

A inclusão de pessoas com deficiência na sociedade tem se consagrado no mundo ocidental, especialmente a partir da década de 1980, como um lema motivador de importantes organizações sociais e ações políticas. Na Europa e nos Estados Unidos da América, já nos anos 1970, a inclusão social das pessoas com deficiência destacava-se entre o básico dos direitos sociais expressos em documentos legais e normativos importantes (MACIEL, 2000).

Segundo a Convenção sobre os Direitos das Pessoas com Deficiência promulgada pela Organização das Nações Unidas – ONU em março de 2007, os Estados Partes² assegurarão um sistema educacional inclusivo em todos os níveis de ensino, garantindo assim que as pessoas com deficiência não sejam excluídas sob alegação de possuir qualquer tipo de deficiência. Afim de promover efetiva educação,

¹ a Convenção foi aprovada pela Assembleia Geral das Nações Unidas em 13 de dezembro de 2006. Foi ratificada por 153 países.

² cada um dos países signatários de um determinado acordo e aqueles que a ele aderirem posteriormente.

os alunos deverão receber o apoio necessário de acordo com suas necessidades individuais (BRASIL, 2007a).

A Convenção sobre os Direitos das Pessoas com Deficiência e seu protocolo facultativo foram ratificados pelo Congresso Nacional Brasileiro em 9 de julho de 2008, por meio do Decreto Legislativo no 186/2008, sendo incorporada à nossa legislação com equivalência de emenda constitucional. Cabe, então, ao governo brasileiro implementar as políticas promotoras dos direitos das pessoas com deficiência, tendo como base o texto dessa convenção (BRASIL, 2008). Gradativamente as sociedades democráticas vem divulgando, discutindo e defendendo a inclusão como direito de todos em relação aos diversos espaços e meios sociais (MAZZOTTA; D'ANTINO, 2011).

Segundo o Instituto Brasileiro dos Direitos da Pessoa com Deficiência (IBDD, 2008) os direitos das pessoas com deficiência na sociedade foram personificados na construção do Estado do Bem-Estar Social e foram plantados e colhidos das mais diversas formas em nosso mundo atual. O direito à educação, como necessidade fundamental para a construção do ser humano, indispensável ao homem como ser completo onde conhecimento e cultura transformam cotidianamente suas dimensões de vida.

Pode-se afirmar que aqui no Brasil nos últimos anos houve grande avanço no que se refere ao Direito Educacional, particularmente em relação a grupos reiteradamente excluídos das oportunidades escolares, tais como os negros, os indígenas e as pessoas com deficiências (MAZZOTTA; D'ANTINO, 2011).

Antecederam a Convenção sobre os Direitos das Pessoas com Deficiência da ONU outros documentos internacionais que contribuíram para alavancar ações em prol da educação inclusiva em diversos países. Entre eles estão a Conferência Mundial de Educação para Todos, de 1990, a Declaração de Salamanca, de 1994, e a Convenção da Guatemala, de 1999. Nos vários documentos internacionais citados, encontram-se afirmações como: “identificação e eliminação de obstáculos e barreiras à acessibilidade” (BRASIL, 2007b, p.6);

2.2 ACESSIBILIDADE

Em um sentido amplo percebemos a evolução e a busca pelas melhores condições de vida e a igualdade para todos, entretanto surge a necessidade de tornar

a vida das pessoas com deficiências mais fáceis possíveis, e um dos termos mais utilizados quando se trata de inclusão social é a acessibilidade. Para o Brasil (2011, p. 1) o termo acessibilidade significa:

Incluir a pessoa com deficiência na participação de atividades como o uso de produtos, serviços e informações. Alguns exemplos são os prédios com rampas de acesso para cadeira de rodas e banheiros adaptados para deficientes.

Sendo assim podemos considerar que a acessibilidade é um conjunto de elementos que auxilia às pessoas com deficiências na mobilização e utilização de qualquer espaço quer seja físico ou virtual por vários meios, e dessa forma possibilita às pessoas com deficiências usufruírem dos mesmos recursos que qualquer outra pessoa.

Segundo Ronchetti (2018) acessibilidade significa não apenas permitir que pessoas com necessidades especiais ou mobilidade reduzida pudessem participar de atividades que incluam o uso de produtos, serviços e informação, mas a inclusão e extensão do uso destes por todas as parcelas presentes em uma determinada população, visando sua adaptação e locomoção, eliminando as barreiras.

O objetivo da acessibilidade é de garantir a toda e qualquer pessoa com necessidade especial ou mobilidade reduzida, possa se deslocar ou utilizar recursos sem que seja encontrada barreiras que impossibilitem o convívio, trânsito social ou uso dos recursos ao seu dispor (PAIXAO, 2014).

2.2.1 Acessibilidade e Educação

Abordar sobre os direitos de educação das pessoas com deficiências é falar sobre um conflito antigo e inerente à sociedade capitalista, que é conflito a exclusão social (CAIADO, 2014).

Segundo Fiegenbaum (2009) aqui no Brasil as mudanças nessa perspectiva começaram a ocorrerem nos anos 2000, algumas das quais começaram a ser executadas em diferentes espaços escolares, porém para muitas questões ainda não existem respostas.

Entretanto nessas últimas décadas vem surgindo continuamente soluções que facilitam vários tipos de deficiências, não só o estado, mas a população em si também vem contribuindo de várias maneiras para facilitar cada vez mais o cotidiano das pessoas com deficiências, criando elevadores para pessoas que usam cadeiras de rodas, rampas, softwares para leitura de telas e outros que são desenvolvidos.

Um dos maiores aliados das acessibilidades em diversos tipos de deficiências é a tecnologia. Com o nível de crescimento tecnológico, é de reconhecer que ela vem auxiliando e fornecendo vários meios e alternativas para facilitar o cotidiano das pessoas com deficiências. No que se diz respeito aos equipamentos e recursos tecnológicos usados para o auxílio de pessoas com deficiência um termo é utilizado, este é a Tecnologia Assistiva (TA). Segundo o Comitê de Ajudas Técnicas (CAT, 2007) a TA é uma área do conhecimento, de característica interdisciplinar, que engloba produtos, recursos, metodologias, estratégias, práticas e serviços que objetivam promover a funcionalidade, relacionada à a vida de e participação, de pessoas com deficiência, incapacidades ou mobilidade reduzida, visando sua autonomia, independência, qualidade de vida e inclusão social.

Existem vários tipos de deficiências que utilizam a TA como auxílio para o seu cotidiano, dentre elas pode se destacar os deficientes que possuem restrição motoras (dificuldades no uso do teclado ou mouse), e para estes tipos de deficiências um dos maiores desafios é a escrita. Várias tecnologias como teclados virtuais e softwares de reconhecimento de voz estão sendo e já foram criadas para que esse grupo possa ser incluído no mundo digital e até mesmo para facilitar o seu cotidiano.

2.2.2 Tecnologia Assistiva Para Deficientes Com Restrições Motoras

Como Radabaugh (1993) disse: “Para as pessoas sem deficiência, a tecnologia torna as coisas mais fáceis. Para as pessoas com deficiência, a tecnologia torna as coisas possíveis”.

Entretanto, constata-se que mesmo na presença de tanta diversidade de recursos tecnológicos para auxiliar na ação de escrever, ainda assim não é possível considerar que está diversidade atenda a todas as necessidades, uma vez que há uma série de diferentes deficiências, de modo que cada uma delas se apresenta em cada indivíduo e dos contextos que envolvem esta atividade, como no mercado de

trabalho ou no ambiente educativo. Desse modo, uma tecnologia assistiva pode funcionar para uns indivíduos e não para os outros, mesmo que apresentem a mesma classificação nosológica ³ (DUSIK, 2013).

A TA cada vez mais procura resolver maior número possível de problemas ligados a pessoas com deficiência, dentre eles podemos citar o cuidado pessoal, mobilidade, conhecimento, produção de escrita e outros.

A área de TA que trata especificamente a ampliação de habilidades de comunicação é chamada de Comunicação Aumentativa e Alternativa (CAA) ela é focada a pessoas que possuem deficiências que dificultam ou impossibilitam a fala ou a escrita ou ainda defasagem entre a suas necessidades de se comunicar e a sua habilidade em falar e/ou escrever (BERSCH; SCHIRMER, 2005).

A CAA trabalha com dois tipos de sistemas comuns, os sistemas com e sem ajuda. Sistemas sem ajuda compreendem assinatura e gestos, que não exigem materiais ou equipamentos especiais. Sistemas “com ajuda” dependem de materiais e equipamentos especiais, como imagens, livros e computadores. Os métodos de CAA são variados e devem ser personalizado para atender às necessidades de cada indivíduo. Muitas formas de CAA incluem componentes de tecnologia assistiva que vão desde os mais básicos até os de alta tecnologia (LEITE, 2014). Destes componentes de TA podem se destacar alguns que facilitam ou possibilitam a escrita de pessoas com restrição motora como:

- a) teclado IntelliKeys: um produto desenvolvido pela AbleNet⁴, o teclado IntelliKeys USB introduz um novo conceito: um teclado que muda de aparência em segundos, permitindo acesso físico, visual e cognitivo conforme a figura 1, para pessoas portadoras de uma ampla gama de dificuldades. (CLICK, 2018). O teclado é fácil de ser instalado ele vem acompanhado de um CD, após a instalação basta conectar em um PC ou Mac ligando o cabo USB, só para que conste ele não tira as atividades do mouse e do teclado convencional;

³ área da Medicina que se dedica ao estudo, descrição e classificação das diferentes doenças

⁴ é uma empresa líder mundial em desenvolvimentos tecnologia assistencial, possui um bom currículo e serviços para ajudar pessoas com deficiência (ABLENET INC, 2018).

Figura 1 - Teclado IntelliKeys



Fonte: Tecno Accesible (2018).

- b) acionadores: são o centro de tecnologia de acesso. O que pode parecer para alguns um simples “botão” como é ilustrado na figura 2, pode - apropriadamente selecionado e instalado - abrir mundos de acesso para dispositivo de comunicação, controles de ambiente, software de computador, dispositivos de celular, uma vez conectado ele funciona como os dois botões de clicks do mouse convencional (TECASSISTIVA, 2018). Mais uma vez a Ablenet se destaca nessa tecnologia por possuir acionadores com fácil uso e simples instalação para os seus usuários;

Figura 2 - Acionadores com Fios



Fonte: Tec Assistiva (2018)

- c) plugmouse: É um mouse convencional conforme a figura 3, adaptado para receber um ou mais acionadores especiais que funcionam como o clique esquerdo e clique direito, facilitando o uso do mouse e permitindo exercícios motores (CLICK, 2018);

Figura 3 - Mouse Adaptado



Fonte: Click Tecnologia Assistiva (2018)

- d) headmouse é um dispositivo que permite os usuários movimentarem o cursor do computador, fazendo apenas alguns movimentos com a cabeça. O programa trabalha em conjunto com qualquer webcam como ilustra a figura 4, que detecta estes movimentos e os transmite para o mouse (GONÇALVES, 2015);

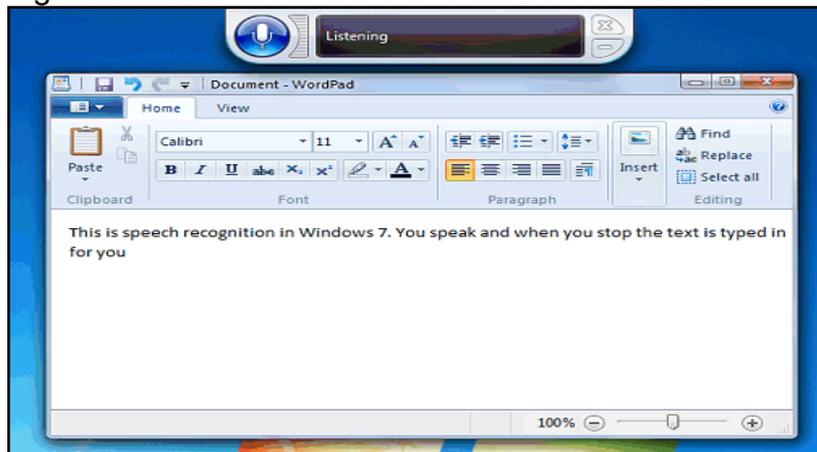
Figura 4 - Mouse Por Movimento da Cabeça



Fonte: Desktop Technology Services Limited (2018)

e) reconhecimento de voz do windows: o reconhecimento de voz está a utilizar a sua voz para controlar o computador. As versões mais antigas do pacote Office, que é conjunto de aplicações para escritórios, incluíam o reconhecimento de voz, mais logo esta funcionalidade começou a fazer parte das funcionalidades do Windows (Microsoft, 2018). Possui uma fácil ativação, no Windows 7 para ativá-lo basta digitar “Reconhecimento” na barra de pesquisa do menu iniciar do Windows e aparecerá “Reconhecimento de Fala do Windows”, entre e configure e após a configuração quando cessado aparecerá na parte superior da tela como ilustrado na figura 5, mas lembre-se de que o software de reconhecimento de fala do Windows 7 não possui suporte ao português. O Suporte a língua portuguesa só foi adicionado na versão do Windows 10. O reconhecimento de voz do Windows torna opcional o uso do teclado e do mouse. Você pode controlar o PC com sua voz e ditar o texto visto que no Windows o reconhecimento de voz pode ser configurado para inicializar junto com o sistema, este recurso é muito útil na tecnologia assistiva;

Figura 5 - Reconhecimento de Fala do Windows 7

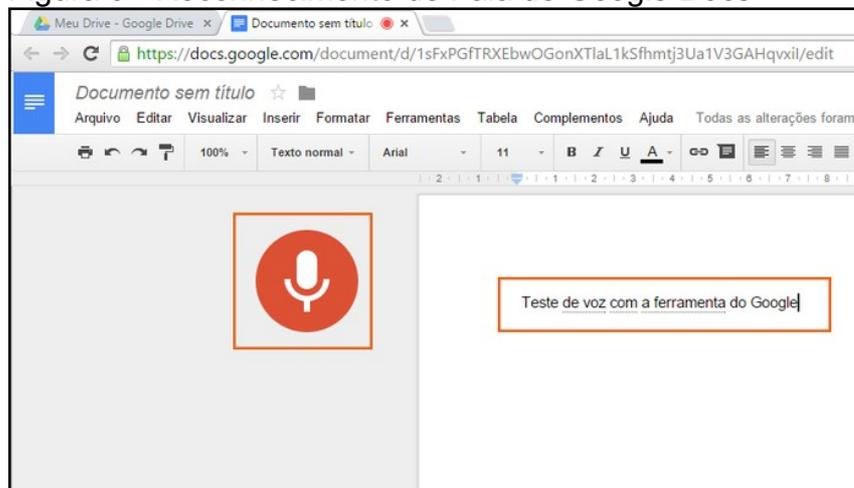


Fonte: RAW Computing (2016).

f) reconhecimento de voz do google docs: por meio desse recurso é possível capturar toda a narração e transformá-la em texto. Esse tipo de funcionalidade é excelente como ferramenta de acessibilidade, estudo, ou em reuniões de trabalho, quando acessado o recurso aparece do lado esquerdo da tela como pode ser visto na figura 6. A digitação por voz está disponível somente para o navegador Google Chrome, ou para os

aplicativos em dispositivos móveis. Existem dezenas de idiomas suportados pelo reconhecimento de voz, entretanto para poder formatar o texto e inserir pontuação, por enquanto, só é possível através de comandos narrados em inglês. O idioma do usuário também precisa estar configurado na linguagem americana. Não foi divulgada uma previsão de data para o lançamento dos recursos de formatação de texto pelo do comando de voz em outros idiomas. (Google, 2018a);

Figura 6 - Reconhecimento de Fala do Google Docs

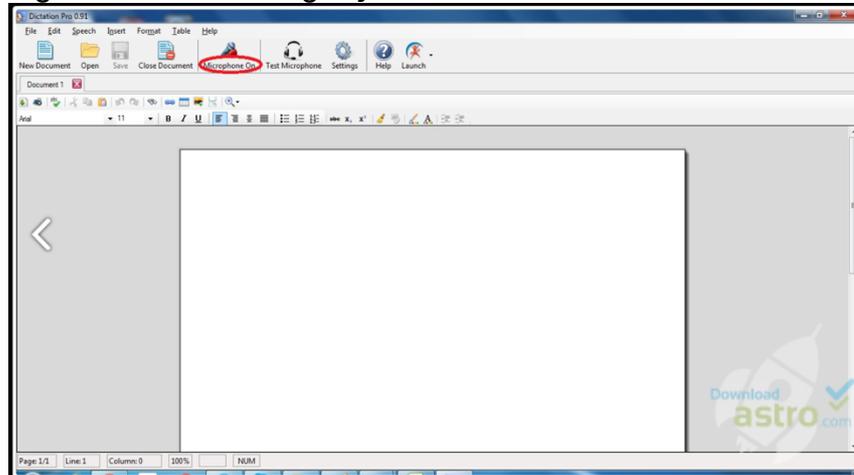


Fonte: Techtudo (2018)

- g) dictationpro: é um dos que fazem parte do grupo de vários *softwares* de reconhecimento de fala gratuito, ele possui uma facilidade de uso, por ser um programa com uma interface amigável como ilustrado na figura 7, se encarrega de digitar tudo o que você fala, desde que o usuário possui um microfone habilitado em sua máquina. Desta forma, você economiza tempo, descansa as mãos ao mesmo tempo e da possibilidade a pessoas com restrições motoras de escreverem sem usar um teclado convencional ou especial. Caso necessite utilizar gírias, basta incorporá-las ao dicionário do *software* antes de começar o trabalho ou durante os comandos. Os comandos são fáceis de memorizar e reconhecidos rapidamente pelo *software*. O comando de voz ainda inclui recursos para copiar texto, apagar palavras, apagar caracteres, alterar a fonte, criar espaço, criar parágrafo e pular linha.

Além disso, você conta com revisor gramatical, que memoriza os erros para que não se repitam posteriormente (DICTATION PRO, 2018a);

Figura 7 - Tela de Digitação do Dictation Pro



Fonte: Dictation Pro (2018b)

2.2.3 Acessibilidades em Aplicações Web

As aplicações web, são hoje consideradas umas das tecnologias mais populares. A sua taxa de usabilidade vem crescendo, segundo a Internet Live Stats (2018, tradução nossa) existem mais de 1,5 bilhão de sites no World Wide Web hoje. Destes, menos de 200 milhões estão ativos. O marco de 1 bilhão de websites foi alcançado pela primeira vez em setembro de 2014, conforme confirmado pela NetCraft em sua Web Server Survey de outubro de 2014 e primeiramente estimado e anunciado pela Internet Live Stats. O número diminuiu subsequentemente, revertendo para um nível abaixo de 1 bilhão (devido às flutuações mensais na contagem de sites inativos) antes de atingir novamente e estabilizar acima da marca de 1 bilhão a partir de março de 2016. Durante 2016, o número total de sites cresceu significativamente, de 900 milhões em janeiro de 2016 para 1,7 bilhão em dezembro de 2016. De 1 site em 1991 para 1 bilhão em 2014, a tabela 1 mostra o número total de sites por ano ao longo da história:

Tabela 1 - Número total de sites por ano

Ano	Websites	Mudanças	Usuários de Internet	Usuários P/ Ano	Website lançados
2017	1,766,926,408	69%			
2016	1,045,534,808	21%			
2015	863,105,652	-11%	3,185,996,155	3.7	
2014	968,882,453	44%	2,925,249,355	3.0	
2013	672,985,183	-3%	2,756,198,420	4.1	
2012	697,089,489	101%	2,518,453,530	3.6	
2011	346,004,403	67%	2,282,955,130	6.6	
2010	206,956,723	-13%	2,045,865,660	9.9	Pinterest
2009	238,027,855	38%	1,766,206,240	7.4	
2008	172,338,726	41%	1,571,601,630	9.1	Dropbox
2007	121,892,559	43%	1,373,327,790	11.3	Tumblr
2006	85,507,314	32%	1,160,335,280	13.6	Twitter
2005	64,780,617	26%	1,027,580,990	16	Youtube, Reddit
2004	51,611,646	26%	910,060,180	18	Facebook, Flickr
2003	40,912,332	6%	778,555,680	19	WordPress, LinkedIn
2002	38,760,373	32%	662,663,600	17	
2001	29,254,370	71%	500,609,240	17	Wikipédia
2000	17,087,182	438%	413,425,190	24	Baidu
1999	3,177,453	32%	280,866,670	88	PayPal
1998	2,410,067	116%	188,023,930	78	Google
1997	1,117,255	334%	120,758,310	108	Yandex
1996	257,601	996%	77,433,860	301	
1995	23,500	758%	44,838,900	1,908	Altavista, Amazon, AuctionWeb
1994	2,738	2006%	25,454,590	9,297	Yahoo
1993	130	1200%	14,161,570	108,935	
1992	10	900%			
1991	1				World Wide Web Project

Fonte: Internet Live Stats (2018).

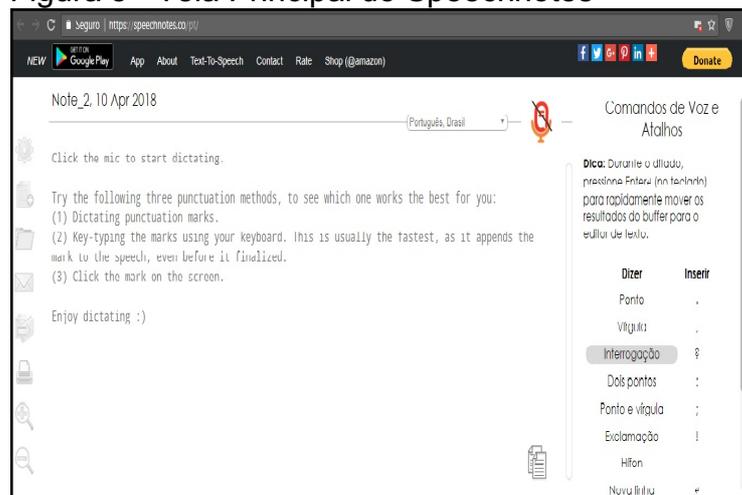
Junto no crescimento do uso de aplicações web, veio à necessidade de criar um meio acessível a qualquer tipo de usuário, a ponto de proporcionar a execução de diferentes tarefas que facilitem o dia-a-dia de seus usuários. Dessa maneira são usadas várias técnicas e estratégias para formar uma melhor interação entre o utilizador com suas capacidades e o dispositivo.

Para usuários com deficiência que impossibilita o uso do teclado convencional, ferramentas especiais como teclados inteligentes, mouses adaptados

são umas das formas mais utilizadas para o acesso à web, mas com o passar dos tempos alguns sites já foram criados com acessibilidades para as pessoas deficiências. Em 2004, um Decreto Federal (nº 5.296) torna obrigatório que todos os portais e sites dos órgãos da administração pública atendam aos padrões de acessibilidade digital. Depois disso, vários decretos, portarias e até uma lei – a Lei de Acesso à Informação Nº 12.527, de novembro de 2011 – trataram do tema, abrangendo todos os sites e não apenas os governamentais. No dia 06 de julho de 2015 foi sancionada a Lei Brasileira de Inclusão da Pessoa com Deficiência (Lei 13.146) que torna obrigatória a acessibilidade nos sítios da internet mantidos por empresas com sede ou representação comercial no País ou por órgãos de governo (BRASIL, 2015).

Um exemplo de aplicação web que facilita a escrita dos usuários com restrição motora é *Speechnotes*, é um software que a voz converte em texto que atende as necessidades de pessoas com restrição motora. Esse software possui uma janela de texto na página principal e um botão no canto superior direito onde o usuário aciona o reconhecimento de voz conforme figura 8, sendo ao lado direito do site possui uma aba onde é selecionada a pontuação do texto. Dessa forma os usuários podem usar os mouses convencionais ou adaptados para navegar entre o aplicativo, sendo que possui o recurso de salvar os dados escritos, enviar por e-mail ou colocar no google drive.

Figura 8 - Tela Principal do Speechnotes



Fonte: Speechnotes (2018)

Alguns aplicativos são feitos exclusivamente para deficientes com restrição motoras ou mobilidade reduzida, como é o caso do Motrix. Esta aplicação é específica para atender a pessoas com esse tipo de deficiência.

O Motrix é um aplicativo para que pessoas com deficiências motoras graves tivessem acesso a computadores. O programa cumpre a função de um mouse ou teclado, por exemplo, ao acionar o computador com a voz do usuário a seta do mouse se movimentar consoante a direção dita pelo usuário. Ao pronunciar determinados comandos em inglês em um microfone, a pessoa é capaz de exercer funções simples no aparelho, como mexer o cursor do mouse e acionar aplicações do Windows. Além disso, consegue soletrar para o computador, mas o processo é lento, de letra por letra (PRODOCTOR, 2017).

Aplicações como esta que permite que pessoas com restrição motora ou mobilidade reduzida tenham acesso ao computador e aos recursos comuns que o mesmo pode oferecer. Apesar de ser uma aplicação gratuita, um dos problemas mais comum é a necessidade de microfones de qualidade mediana para uma experiência sem maiores transtornos (THOMAZ, 2016).

Com o surgimento da tecnologia de reconhecimento de voz, que a princípio era mais comum ver a serem usadas em aplicações desktop, nos últimos tempos constata-se um crescimento de aplicações web e mobile usando a tecnologia de reconhecimento de voz nas suas aplicações. A aplicação desta tecnologia nos websites pode melhorar drasticamente a usabilidade dos websites (NETO, 2014).

Seguindo a mesma ideia de navegação de aplicações via voz, grandes empresas de desenvolvimento de aplicações web estão criando API com o uso de reconhecimento de voz que funcione como o reconhecimento de voz tradicional são instalados nos computadores. Dois das maiores API's de aplicações web do mundo todo voltado para reconhecimento de voz. São estes, a Web Speech API e o Google Cloud Speech API.

A Web Speech API é desenvolvida pelo Grupo de Comunidade Speech API da W3C. A iniciativa é aberta a todos, mas parece que é em grande parte impulsionado por duas empresas: Google e *Openstream*. Na altura foi uma iniciativa recente e era uma versão experimental que aguardavam compromissos finais, e foi publicado apenas em outubro de 2012, e o único navegador que suportava era o Chrome.

A *Google Cloud Speech API* desenvolvido pela Google, permite que os desenvolvedores convertam áudio em texto aplicando poderosos modelos de rede

neural em uma API de fácil uso pelo desenvolvedores. A API reconhece 120 idiomas e variantes para oferecer suporte à sua base de usuários global. Você pode ativar o comando e controle de voz, transcrever áudio de call centers e muito mais. Ele pode processar streaming em tempo real ou áudio pré-gravado usando a tecnologia de aprendizado de máquina do Google (GOOGLE, 2018b, tradução nossa). A Cloud Speech API possui uma desvantagem com relação a Web Speech API da W3C que é o fator de não ser totalmente de graça, ela possui uma cobrança mensal de acordo com a quantidade de áudio processada efetivamente pelo serviço, medida em incrementos arredondados até 15 segundos. Os níveis de preços são baseados na quantidade total de áudio processada pelo serviço por mês (GOOGLE CLOUD, 2018).

Estas duas tecnologias apresentadas demonstram a evolução das TAs em aplicações web, comprovando que com o uso de técnicas existentes e a exploração de novas tecnologias que é sempre possível inovar. Fatores como o alto poder de processamento e a popularidade dos websites atuais, tornam possíveis o desenvolvimento de soluções mais práticas e eficientes, principalmente na área da acessibilidade e a educação que tem tanto a evoluir.

3 TECNOLOGIA

As Tecnologias Assistivas ocorrem tanto no momento que for criado ou utilizado e transformado os recursos de tecnologias já existentes. Variadas vezes elas são criadas a partir da combinação de diferentes técnicas que, quando projetadas em um sistema, são feitas para atingir novas finalidades ou melhorar os processos já existentes.

Este capítulo traz uma breve explicação dos recursos de tecnologias utilizadas para a elaboração do trabalho proposto, como as Aplicações Web, reconhecimento de voz via web e conversor de voz em texto.

3.1 APLICAÇÕES WEB

Uma aplicação web ou "Web App" é um programa de software que é executado em um servidor da web. Ao contrário dos aplicativos de desktop tradicionais, que são lançados para um sistema operacional, os aplicativos da Web devem ser acessados por meio de um navegador da web. Como eles são executados em navegadores da Web, os desenvolvedores não precisam desenvolver aplicativos da Web para várias plataformas. Por exemplo, um único aplicativo que é executado no Chrome funcionará no Windows, OS X e no Linux. Os desenvolvedores não precisam distribuir atualizações de software aos usuários quando o aplicativo da Web é atualizado. Ao atualizar o aplicativo no servidor, todos os usuários têm acesso à versão atualizada (CHRISTENSSON, 2014, tradução nossa).

3.1.1 Arquitetura das Aplicações Web

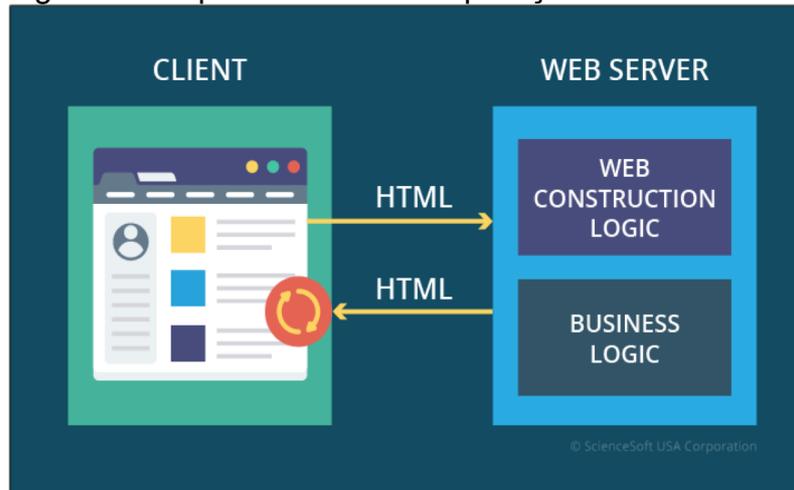
Para desenvolver uma aplicação web, como em um qualquer outro tipo de aplicação, é importante conhecer a arquitetura e a maneira como estão distribuídas as camadas neste modelo de desenvolvimento, para entender de que forma ocorrem as interações dos processos dentro dessa plataforma. Desta feita é imprescindível um breve entendimento de sua disposição.

A arquitetura de aplicativos da Web é um padrão de interação entre componentes de aplicativos da Web. No entanto, essa definição é um pouco ambígua, já que quando falamos sobre componentes de um aplicativo da web, podemos significar duas coisas. O primeiro deles apresenta componentes do aplicativo da Web

UI / UX⁵, mas não tem nada a ver com uma arquitetura de aplicativo da web que pretendemos abordar aqui, mas sim com um plano de layout de interface (SCIENCE SOFT, 2018, tradução nossa).

Conforme demonstra a figura 9, a arquitetura das aplicações web estão divididas em dois componentes principais: Lado do Cliente e o Lado do Servidor.

Figura 9 - Arquitetura de Uma Aplicação Web



Fonte: ScienceSoft (2018).

Para o melhor uso e organização desses componentes, padrões de arquiteturas web foram criados para auxiliar, padrões estes como é o caso do Modelo-Visão-Controlador, do inglês *Model-View-Controller* (MVC), e a arquitetura em três camadas (Three-Tier Architecture). Estes padrões conduzem os desenvolvedores a criar aplicações web que facilitam a sua manutenção, evolução e integração de novos membros no desenvolvimento da mesma, uma vez que cada parte da aplicação possui um papel bem definido e independente de acompanhamento facilitando, assim, a manutenção da aplicação (LUCIANO; ALVES, 2011).

3.1.1.1 Client Side

O componente mais comum nos olhos do usuário, ou seja, é a camada da aplicação que interage com os usuários das aplicações. Quem irá processar as tecnologias que rodam nesse lado não é o servidor, mas o sim o browser como

⁵ o design da interface entre o usuário e o produto / a experiência geral do usuário de usar o produto (ZHOU, 2017, tradução nossa)

(Chrome, IE, Firefox, entre outros). Esse é denominado o "lado do cliente", ou seja, é o tipo de aplicação que roda no navegador do usuário sem necessitar o processamento ou uma requisição no servidor para execução de suas tarefas.

Nos dias de hoje basicamente quando é abordado sobre cliente-side na maioria das vezes está se falando de ECMAScript (sabendo que também está no lado do servidor com o Node.js) e Ajax, estas são duas das tecnologias de *client-side* que mais vem crescendo na área de desenvolvimento web. Quando a linguagem está rodando somente no cliente-side deve se ter um cuidado porque o ECMAScript precisa estar ativado no Navegador para funcionar, por isso devemos preparar bem o nosso server-side para possíveis falhas de ECMAScript.

3.1.1.2 Server Side

O lado do servidor, também conhecido como *back-end* é a parte mecânica da aplicação, ou seja, é onde são usadas aquelas linguagens que são reconhecidas, executadas e interpretadas pelo servidor e que se enviam ao cliente em um formato reconhecido por ele. Este lado normalmente consiste em pelo menos mais duas partes: lógica da aplicação ou centro de controle principal e banco de dados, onde todos os dados são armazenados.

Para os navegadores da web se comunicarem com os servidores web usam o Hyper Text Transfer Protocol (HTTP). Quando o usuário faz um click em um link da página web, envia um formulário ou executa uma pesquisa, uma solicitação HTTP é enviada do navegador para o servidor de destino e lhe é retornado um resultado (MOZILLA DEVELOPER, 2018a, tradução nossa). Algumas das tecnologias mais usadas no *server-side* são essas C#, PHP, ASP, Java, Ruby on Rails e ECMAScript (Node.js) que de uns tempos pra cá vem ganhando um grande espaço em desenvolvimento de aplicações web.

3.2 RECONHECIMENTO DE VOZ

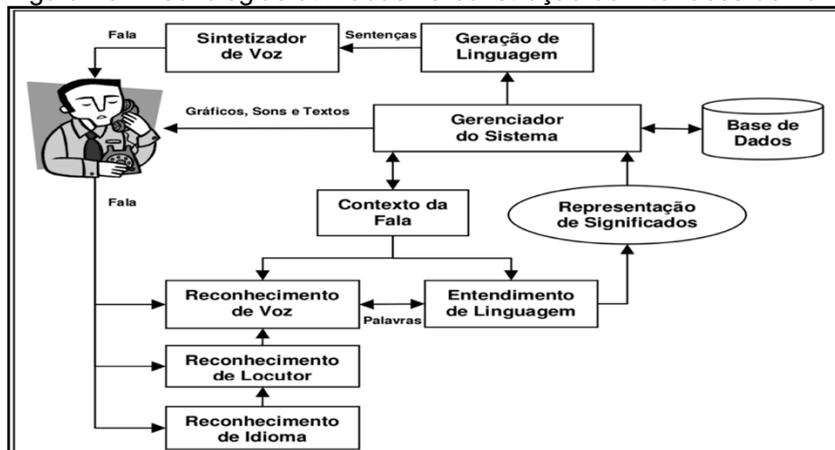
Segundo Viliati (2000) o estudo feito por cima do reconhecimento de voz, ele está dividido em 3 tipos básicos de sistemas: análise, síntese e reconhecimento. O primeiro software de reconhecimento de voz foi criado pela a empresa de pesquisa dos Estados Unidos, Bell Laboratories (fundada por Alexander Graham Bell, inventor

do telefone) no ano de 1952 e o mesmo se chamava AUDREY. O tal software poderia interpretar com precisão mais de 97% dos dígitos falados a ele. Hoje ele é sem dúvida um primitivo pelos padrões de hoje, mas estabeleceu as bases para o reconhecimento por voz, uma tecnologia que foi amplamente utilizada e até hoje abrange o mercado tecnologia.

O reconhecimento de voz faz uso de uma serie de etapas para as quais é necessário o processo entre a emissão do som de uma voz e seu respectivo reconhecimento pelo computador ou dispositivo. Utilizando o microfone como o dispositivo de entrada, para captar a fala e transformá-la em sinais elétricos, o processamento do som se dá através de um conversor analógico-digital responsável por captar vibrações da voz pronunciada e convertê-lá em dados digitais que posteriormente é aplicada uma medida para cada onda captada da voz e este som digitalizado é filtrado para separá-lá dos ruídos e interferências (TEIXEIRA; SOUZA, 2012).

As tecnologias de processamento da voz em conjunto com técnicas avançadas de linguística, inteligência artificial e recuperação de dados, proporcionam a modelagem de um sistema de interface falada entre homens e máquinas conforme apresentado por Cole (1997) em “*Survey of the State of the Art in Human Language Technology*” e representado de forma breve na Figura 10. Nesta figura, observa-se o dever incumbido a um bloco gerenciador do sistema que, associado a um banco de dados com inteligência artificial e conhecimento linguístico, permitiria um diálogo em linguagem natural entre homem e máquina através das suas interfaces de reconhecimento e síntese de voz (TEVAH, 2006).

Figura 10 - Tecnologias utilizadas na construção de interfaces de voz



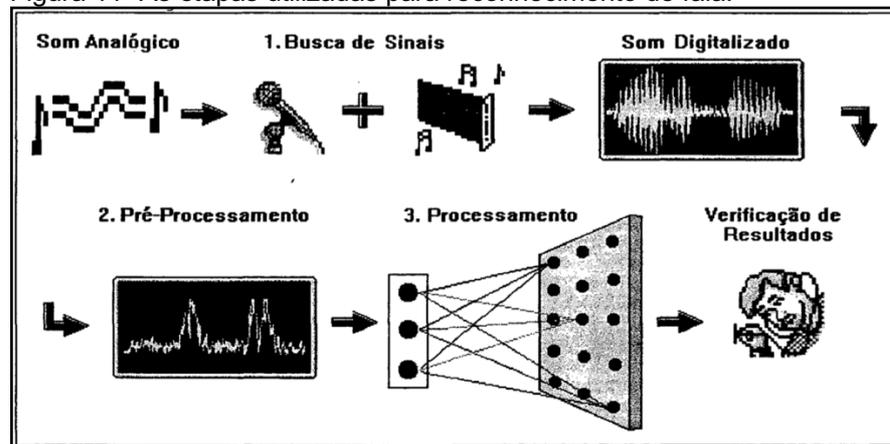
Fonte: (TEVAH, 2006)

A figura acima mostra o funcionamento de uma interface de voz e o seu funcionamento bem estruturado com uma base de dados sólida, a evolução desta área está mais ou menos resumida no subtema abaixo.

3.2.1 Etapas do Reconhecimento de Voz

Segundo Hugo (1995) a função de reconhecer palavras faladas pode ser repartida em três etapas principais: Busca de Sinais, Pré-processamento e Processamento. Cada qual realiza alterações nos dados recebidos, como pode ser visto na figura 11.

Figura 11- As etapas utilizadas para reconhecimento de fala.



Fonte: Hugo (1995)

3.2.1.1 Busca dos sinais

A primeira etapa, considera-se Busca dos Sinais ou Captação, faz a captação do som e sua alteração em sinal digital para utilização por um computador digital.

"Na sua forma natural os sons são análogos. Eles aumentam e diminuem, alterando a frequência e amplitude em suaves progressões. Obviamente, o computador não pode armazenar o som como ele é. Para armazenar o som, o formato da informação precisa ser alterado para que o computador possa armazenar ou manipular os sons na memória. Em outras palavras, a informação precisa ser digitalizada. Ela precisa ser convertida em números digitais com os quais o computador pode trabalhar" (MICROSOFT, 2018, tradução nossa).

O som é captado através de um microfone, no qual está ligado a uma placa digitalizadora. Placas de som, conforme a *Sound Blaster* ou *Fusion CD 16*, fazem operações de digitalização de sinais sonoros. Particularidades sobre o som e a sua digitalização, podem ser consultados em Matras (1991) e Stolz (1993, tradução nossa) respectivamente.

3.2.1.2 Pré-processamento

O pré-processamento de sinais que é considerada a segunda etapa, é uma maneira de regular ou regularizar os sinais captados buscando torná-los passíveis de tratamento na fase seguinte do processo de reconhecimento de fala.

Segundo Bezdek e Pal (1992) esta fase pode ser dividida no que chamam de:

- a) pré-processamento: utilização de técnicas de "limpeza" do sinal, como normalização, escala, suavização, etc.;
- b) extração de características: eliminação de sinais repetitivos ou desnecessários através de uma seleção ou transformação.

3.2.1.3 Processamento

Esta fase é responsável por realizar o reconhecimento do padrão, normalmente por agrupamento ou classificação (BEZDEK; PAL, 1992, tradução nossa). A diferença entre eles, é que no agrupamento utiliza-se frequentemente de técnicas de aprendizado não supervisionado para criar grupos de padrões dentro do universo a ser reconhecido, enquanto que a classificação frequentemente utiliza-se de técnicas de aprendizado supervisionado para rotular, se for o caso, todos os padrões do universo a ser reconhecido.

Uma vez feito o processamento, pode-se considerar a ação associada ao tipo adotado, de acordo com a aplicação.

Em relação às diversas formas de enunciação do locutor, pode-se classificar os reconhecedores de vozes das seguintes formas (FURUI, 1989 apud MARTINS, 1997):

- a) reconhecimento de palavras isoladas: surgido nos anos 70, este tipo de reconhecimento as palavras são enunciadas de forma independentes. Existem diversos sistemas comerciais amplamente utilizados, tanto na versão em software como na versão em hardware, cuja taxa de reconhecimento, independente do locutor, é de aproximadamente 100% para vocabulários muito pequenos (até 10 palavras) e de aproximadamente 95%, para vocabulários pequenos e médios (de 10 até 1000 palavras);
- b) reconhecimento de palavras conectadas: surgido nos anos 80, neste tipo de reconhecimento as palavras são enunciadas de forma contínua e são selecionadas na base de dados do reconhecedor uma por uma. O reconhecimento de palavras contínuas difere do de palavras isoladas, pois neste o locutor não precisa fazer pausas entre as palavras. Deste modo, a demarcação das fronteiras entre as palavras e do número destas na expressão deve ser feita pelo sistema (YNOGUTI, 1999);
- c) reconhecimento de fala contínua: surgido nos anos 90, esse tipo de reconhecimento costuma a ser bem mais complexa que as outras duas, pois nesse tipo de reconhecimento as palavras são reconhecidas em grupos e apresenta efeitos de coarticulação. Este tipo de reconhecimento e bem mais difícil de serem implementados, pois eles devem ser capazes de lidar com todas as características e formas naturais de fala, dentre eles podem ser destacados: durações de palavras desconhecidas, efeitos de coarticulação (por exemplo, o /z/ de deu zebra e de belíssimas zebras) e pronúncias com pouco cuidado (ALENCAR, 2005, p. 18).

Atualmente o reconhecimento de fala contínua tem sido o que segue o estudo mais aprofundado e sendo assim é ampliado o seu vocabulário, universidades e indústria têm tentado resolver problemas realistas da área, de forma a possibilitar o reconhecimento da fala natural. O propósito é a criação de sistemas que possam ser utilizados sem a necessidade de treinamento intensivo por parte do usuário e com mínimas taxas de erros. A última década testemunhou um progresso significativo na tecnologia de reconhecimento de fala. Em 1997 já se reportava taxas de erro da ordem de 10% para sistemas de reconhecimento de fala contínua com vocabulários amplos

(20.000 palavras ou mais) nos testes elaborados normalmente nos EUA por meio do Departamento de Defesa (TEVAH, 2006). Existem hoje uma boa quantidade de tecnologias e recursos, quer seja APIs ou mesmo tecnologias embutidas em linguagens de programação que possibilitam hoje uma fácil criação desses sistemas na web, desktop ou mesmo mobile.

3.3 WEB SPEECH API

Essa especificação define uma API desenvolvida em ECMAScript criada em 2012, com a finalidade de permitir que os desenvolvedores da Web integrem reconhecimento e síntese de fala em suas páginas da Web. Ele permite que os desenvolvedores usem scripts para gerar saída de texto em voz e usar o reconhecimento de fala como uma entrada para formulários, ditado e controle contínuos. A API permite que as páginas da web controlem a ativação e o tempo e manipulem resultados e alternativas (SHIRES, 2014, tradução nossa). A Web Speech API está dividida em duas partes: *SpeechSynthesis* (Síntese de Voz) e *SpeechRecognition* (Reconhecimento de Fala) que é o recurso da API que usaremos no nosso protótipo.

3.3.1 SpeechRecognition

O reconhecimento de fala da aplicação é acessado por meio da interface *SpeechRecognition*, que fornece a capacidade de reconhecer o contexto de voz a partir de uma entrada de áudio (normalmente por meio do serviço de reconhecimento de fala padrão do dispositivo) e responder apropriadamente. Geralmente, você usará o construtor da interface para criar um novo objeto *SpeechRecognition*, que possui vários manipuladores de eventos disponíveis para detectar quando a fala é inserida através do microfone do dispositivo que são abordados nos subtemas abaixo. Uma das interfaces chave é a interface *SpeechGrammar* que representa um contêiner para um determinado conjunto de gramática que o seu aplicativo deve reconhecer. A gramática é definida usando o formato de gramática JSpeech (JSGF.) (MOZILLA DEVELOPERS, 2018b, tradução nossa).

3.3.1.1 Métodos da interface reconhecimento de fala

Os métodos da interface mencionados pela Mozilla Developers (2018b, tradução nossa) são:

- a) *SpeechRecognitionEvent* o objeto de evento para os eventos *result* e *nomatch* e contém todos os dados associados a um resultado de reconhecimento de fala provisório ou final;
- b) *SpeechGrammar* as palavras ou padrões de palavras que queremos que o serviço de reconhecimento reconheça;
- c) *SpeechGrammarList* representa uma lista de objetos *SpeechGrammar*;
- d) *SpeechRecognitionResult* representa uma única correspondência de reconhecimento, que pode conter vários objetos *SpeechRecognitionAlternative*;
- e) *SpeechRecognitionResultList* representa uma lista de objetos *SpeechRecognitionResult* ou um único se os resultados estiverem sendo capturados no modo contínuo.

3.3.1.2 Atributos da interface reconhecimento de fala

Os atributos da interface mencionados por Shires (2012, tradução nossa) são as seguintes:

- a) *grammars attribute* o atributo armazena a coleção de objetos *SpeechGrammar* que representam as gramáticas que estão ativas para esse reconhecimento;
- b) *lang attribute* esse atributo definirá o idioma do reconhecimento para a solicitação, usando uma tag de idioma BCP 47 válida. [BCP47] Se não definido, ele permanecerá indefinido para entrar no script, mas usará como padrão o idioma do elemento raiz do documento html e da hierarquia associada. Esse valor padrão é calculado e usado quando a solicitação de entrada abre uma conexão com o serviço de reconhecimento;
- c) *continuous attribute* quando o atributo contínuo está configurado como *false*, o agente do usuário não deve retornar mais do que um resultado

final em resposta ao reconhecimento inicial, por exemplo, um padrão de interação de um único turno. Quando o atributo contínuo é configurado como true, o agente do usuário deve retornar zero ou mais resultados finais representando vários reconhecimentos consecutivos em resposta ao reconhecimento inicial, por exemplo, um ditado. O valor padrão deve ser falso. Observe que essa configuração de atributo não afeta os resultados provisórios;

- d) *interimResults attribute* controla se os resultados provisórios são retornados. Quando definido como true, os resultados provisórios devem ser retornados. Quando definido como falso, os resultados provisórios NÃO devem ser retornados. O valor padrão deve ser falso. Observe que essa configuração de atributo não afeta os resultados finais;
- e) *maxAlternatives attribute* este atributo irá definir o número máximo de `SpeechRecognitionAlternatives` por resultado. O valor padrão é 1;

- f) *serviceURI attribute* ele especifica o local do serviço de reconhecimento de fala que o aplicativo da Web deseja usar. Se esse atributo não for definido no momento da chamada do método de início, o agente do usuário deverá usar o serviço de fala padrão do agente do usuário. Observe que o `serviceURI` é um URI genérico e pode, portanto, apontar para serviços locais por meio do uso de um URN com significado para o agente do usuário ou especificando um URL que o agente do usuário reconheça como um serviço local. Além disso, o padrão do agente do usuário pode ser local ou remoto e pode incorporar opções do usuário final por meio de interfaces fornecidas pelo agente do usuário, como parâmetros de configuração do navegador.

4. TRABALHO CORRELATOS

Diversas pesquisas e desenvolvimentos são realizados a todo o momento no mundo inteiro na área de tecnologia. Algumas áreas como tecnologia móvel e recursos Web são muito exploradas, devido a grande quantidade de usuários a qual essas plataformas atingem e a variedade de recursos para o desenvolvimento que elas disponibilizam.

Nos tópicos subsequentes são abordadas algumas dessas pesquisas relacionadas ao presente trabalho.

4.1 SISTEMA DE RECONHECIMENTO DE FALA VIA WEB

Em 2015, o acadêmico Joao Daniel da Silva Cardoso, do programa de Dissertação de Mestrado em Engenharia Electrotécnica e de Computadores na Universidade de Coimbra, em Portugal, elaborou um estudo sobre reconhecimento de voz. Com o conhecimento obtido, desenvolveu uma aplicação com Julius Speech Recognizer e NodeJs o sistema web com reconhecimento de voz.

O reconhecedor é integrado a um servidor e este gere os acessos dos clientes, feitos através de páginas HTML. Para que tal aconteça, o usuário deverá enviar o áudio juntamente com a lista de possíveis palavras a reconhecer. O microfone é acedido através da MediaStream API, será analisado pelo WebAudio API. Cada segmento de áudio gerado é enviado para o servidor, via streaming.

O sistema permite satisfazer pedidos de reconhecimento em simultâneo e oriundos de qualquer sítio da internet. Também é compatível com os principais browsers de internet. Para isso, basta que as páginas HTML sigam um protocolo simples de indicação de objetos que podem ter uma entrada por fala, além dos habituais dispositivos de entrada, rato e teclado.

4.2 UM SOFTWARE DE RECONHECIMENTO DE VOZ PARA PORTUGUÊS BRASILEIRO

Em 2012, o acadêmico Carlos Patrick Alves da Silva do Programa de Pós-Graduação em Engenharia elétrica da Universidade Federal de Pará (UFP) desenvolveu um software para desktop de reconhecimento de voz fortemente voltado ao português Brasileiro. Esta aplicação foi desenvolvida com o sistema de reconhecimento de voz contínua por ser voltada a grandes vocabulários.

Como parte deste projeto foram produzidos e disponibilizados vários recursos: modelos acústicos e de linguagem, novos *corpora* de voz e texto. O *corpus*⁶ de texto vem sendo construído através da extração e formatação automática de textos de jornais na Internet. Além disso, foram produzidos dois *corpora* de voz, um baseado em audiobooks e outro produzido especificamente para simular testes em tempo-real.

O trabalho também propõe a utilização de técnicas de adaptação de locutor para resolução de problemas de descasamento acústico entre *corpora* de voz. Por último, é apresentada uma interface de programação de aplicativos que busca facilitar a utilização do decodificador Julius. Testes de desempenho são apresentados, comparando os sistemas desenvolvidos e um software comercial.

Foram realizados testes de desempenho utilizando gramáticas. Para a realização dos testes se utilizou o decodificador Julius, pois a sua gramática é bastante simples de implementar. Inicialmente, foi construída uma gramática composta por 31 frases que correspondem a possíveis comandos dados a um computador.

4.3 SISTEMA DE RECONHECIMENTO DE VOZ – APLICABILIDADE

O sistema de reconhecimento de voz foi desenvolvido pelo acadêmico Luis Gustavo de Carvalho Uzai da Faculdade de Tecnologia de Garça, em São Paulo, o sistema é apto a processar voz contínua e retorna a resposta imediata pelo programa.

⁶ é a coletânea ou conjunto de documentos sobre determinado tema; repertório ou conjunto da obra científica, técnica e/ ou artística de uma pessoa ou a ela atribuída (SILVA; SILVA, 2013).

O então aluno elaborou um estudo sobre o reconhecimento de voz através do computador. O mesmo estudo foi para especificar uma plataforma de hardware e implementar o software para síntese, identificação e reconhecimento de voz, baseado nas técnicas de processamento digital de sinais a partir de dados estatísticos e verificação, para isso foi necessário a realização um estudo das técnicas de reconhecimento de padrões de voz.

Esse sistema foi desenvolvido para ambiente Windows, utilizando a linguagem C#.NET, ferramenta que faz parte do Microsoft Visual Studio 2010 Express Edition e o pacote complementar Microsoft SpeechSDK 5.1. Os comandos e suas funções do sistema são fornecidos pelo usuário, adequando-se a suas necessidades práticas.

Foram realizados vários testes no sistema, e mostrou-se que possui uma boa performance no reconhecimento de voz, porém o sistema não obteria desempenho necessário para se tornar aplicável em um sistema de ditar a voz, pela alta taxa de erro na duplicidade de palavras (37% aproximados) e taxa de falha no reconhecimento de palavras simultâneas (28% aproximados), porém não apresenta uma baixa no desempenho sendo utilizado na língua portuguesa, em relação ao mesmo sistema em língua inglesa, em quesito de reconhecimento para execução de tarefas, onde palavras individuais são usadas para ativar os comandos o sistema se mostrou otimizado.

4.4 TECNOLOGIA DE ENSINO E TECNOLOGIA ASSISTIVA NO ENSINO DE CRIANÇAS COM PARALISIA CEREBRAL

Na Universidade Federal de São Carlos, um grupo de pesquisadores realizou um estudo sobre o uso da tecnologia assistiva na escola regular pelas pessoas que possuem paralisia cerebral e deficiências físicas.

Ao decorrer da pesquisa foi constatado que a escolarização da criança com comprometimento motor na escola comum tem sido discutida juntamente com as possíveis estratégias que possam colaborar para a participação e aprendizagem desse aluno. Uma das estratégias indicadas tem sido o uso de recursos de tecnologia assistiva para favorecer a execução das tarefas pedagógicas, porém pouco se sabe

sobre a implementação, a eficácia ou as contribuições desses recursos no processo de escolarização da criança com deficiência física.

O foco desta pesquisa foi identificar, a partir do ponto de vista do aluno com paralisia cerebral, as contribuições, dificuldades e o cotidiano implicado no uso de recursos de tecnologia assistiva no contexto da escolarização no ensino regular. Foram selecionados 5 alunos com diagnóstico de paralisia cerebral nível motor IV e V segundo o Sistema de Classificação da Função Motora Grossa para paralisia cerebral – GMFCS.

As percepções das crianças foram relevantes e elas mostraram-se capazes de apontar demandas e identificar quando e como os recursos podem ser limitantes dentro do contexto. Assim, considerar as opiniões e a participação da criança na identificação de recursos de tecnologia pode contribuir na compreensão do impacto, das demandas e das possíveis contribuições no processo de implementação e utilização do recurso de tecnologia assistiva.

5 TRABALHO PROPOSTO

Este trabalho tem como o objetivo desenvolver um protótipo de aplicação web usando o reconhecimento de voz voltado ao aprendizado de escrita voltado para pessoas que possuem dificuldades no uso do teclado. O reconhecimento de voz que será utilizado neste protótipo será proveniente de uma API desenvolvida pela World Wide Web Consortium (W3C) que terá como a finalidade de reconhecer a voz do usuário quando acionado e converter a voz em texto para auxiliar no aprendizado de escrita de pessoas que possuem de coordenação motora reduzidas nos membros superiores e dificultam o uso do teclado.

5.1 METODOLOGIA

Para a concretização deste trabalho foi feito um levantamento bibliográfico relacionado a acessibilidade para deficientes com coordenação motora reduzidas, dificuldades nos membros superiores, reconhecimento de voz e conversão de voz em texto. Sendo assim, foi proposto o desenvolvimento de um protótipo de uma aplicação Web para conversão de voz em texto. Foram aplicadas ainda formas de interação com o usuário do aplicativo, também adquiridas no processo de fundamentação teórica.

As operações relacionadas ao reconhecimento de voz e a conversão de voz em texto que será responsabilidade da Web Speech API, será implementado usando a linguagem de programação ECMAScript. Primeiramente será feito uma análise do funcionamento dos sistemas de reconhecimento de voz existentes e em seguida será feito uma outra análise com relação a conversão de voz em texto com a Web Speech API, após essa análise será definido a forma de interação do usuário com aplicação e dando início ao do desenvolvimento do prototipo.

O teste será realizado em ambiente habitacional para poder avaliar a eficiência do protótipo, usando microfone nativo do notebook e o navegador Google Chrome na versão 70.0.

5.1.1 A Tecnologia de Reconhecimento de voz Web Speech API

A tecnologia escolhida para a realização do reconhecimento de voz foi a Web Speech API, pois além de ser *open-source* com um excelente histórico de desenvolvimento, tendo na sua lista de contribuintes empresa como a *Google*, ainda apresentou uma pré-disposição para a utilização em aplicações web. Na sua página oficial, os navegadores Chrome, Edge, Firefox são citadas dentre as suportadas pela tecnologia (MOZILLA DEVELOPER, 2018, tradução nossa).

A tecnologia Web Speech API tem como princípio básico de funcionamento receber uma fala, submetê-la ao processo de reconhecimento de vozes e então retornar a fala convertida em texto. A Web Speech API tem duas partes: Síntese de Fala (*Text-to-Speech*) e Reconhecimento de Fala (*Asynchronous Speech Recognition*).

O reconhecimento de fala é acessado por meio da interface *SpeechRecognition*, que fornece a capacidade de reconhecer o contexto de voz a partir de uma entrada de áudio (normalmente por meio do serviço de reconhecimento de fala padrão do dispositivo) e responder apropriadamente. Geralmente, você usará o construtor da interface para criar um novo objeto *SpeechRecognition*, que possui vários manipuladores de eventos disponíveis para detectar quando a fala é inserida através do microfone do dispositivo. A interface *SpeechGrammar* representa um contêiner para um determinado conjunto de gramática que seu aplicativo deve reconhecer. A gramática é definida usando o formato de gramática *JSpeech* (JSGF).

5.1.1.1 Captura de Fala

Foram feitas pesquisas sobre a captura de fala em computadores, e foi constatada duas formas de realizar essa tarefa, utilizando um microfone nativo do computador ou microfones adicionais.

O reconhecimento de fala envolve o recebimento de fala por meio do microfone de um dispositivo, que é verificado por um serviço de reconhecimento de fala em relação a uma lista de gramática (basicamente, o vocabulário que você deseja reconhecer em um aplicativo específico). É retornado como resultado (ou lista de resultados) como uma cadeia de texto, e outras ações podem ser iniciadas como resultado.

5.1.1.2 Utilizando a Web Speech API

Com o código pronto para funcionar usando somente o teclado convencional, foi possível iniciar o desenvolvimento do protótipo utilizando a API. A princípio foi criada uma classe, ou seja, um *service* denominado *SpeechRecognitionService* (SRS), visto que usaríamos o serviço de reconhecimento de voz da API declaramos os métodos necessários no desenvolvimento do nosso protótipo como podemos ver na figura 12.

Figura 12 - Classes da Web Speech API

```
interface IWindow extends Window {  
    |   webkitSpeechRecognition: any;  
    |   SpeechRecognition: any;  
    |  
    }  
}
```

Fonte - Do Autor

A Web Speech API possui duas classes principais, uma delas é a *webkitSpeechRecognition* ou a *SpeechRecognition* (SR) e a classe *SpeechGrammar* ou *webkitSpeechGrammarList*. Dessa forma, ao instanciar um objeto dos tipos citados acima, muitos dos recursos originais da API já se tornam disponíveis como observamos na figura 13.

Figura 13 - Instância do objeto *webkitSpeechGrammarList*

```
const { webkitSpeechRecognition }: IWindow = <IWindow>window;  
this.speechRecognition = new webkitSpeechRecognition();  
//this.speechRecognition = SpeechRecognition;
```

Fonte: Do Autor

Para realizar um reconhecimento de fala com a Web Speech API, se faz necessário seguir alguns passos básicos. Com a instância *speechRecognition* criada, é necessário chamar o método *speechRecognition.start()*, definir os atributos mais importantes como o tipo de reconhecimento de fala que será usado, a linguagem desejada. É necessário também definir o número máximo de alternativas. O reconhecimento de fala pode ser feita através do método *onResult()* onde tem toda uma lógica para transformar a voz em texto, e aí temos outros métodos que poderão

auxiliar no processo de reconhecimento de fala. O procedimento descrito pode ser observado no trecho de código-fonte conforme figura 14.

Figura 14 - Trecho de código para a inicialização da *SpeechRecognition*

```

this.speechRecognition.continuous = true;
//this.speechRecognition.interimResults = true;
this.speechRecognition.lang = 'pt-br';
this.speechRecognition.maxAlternatives = 1;

this.speechRecognition.onresult = speech => {
  let term: string = "";
  if (speech.results) {
    var result = speech.results[speech.resultIndex];
    var transcript = result[0].transcript;
    if (result.isFinal) {
      if (result[0].confidence < 0.3) {
        console.log("Unrecognized result - Please try again!");
      }
      else {
        term = _.trim(transcript);
        console.log("Did you said? -> " + term + " , If not then say something else...");
      }
    }
  }
  this.zone.run(() => {
    observer.next(term);
  });
};

this.speechRecognition.onerror = error => {
  observer.error(error);
};

this.speechRecognition.onend = () => {
  observer.complete();
};

this.speechRecognition.start();
console.log("Say something - We are listening !!!");
};

```

Fonte: Do Autor

5.1.1.3 Conversão de Voz em Texto

Para conversão de texto em voz, a Web Speech API disponibiliza o recurso de *SpeechRecognition* como pode ser observado na figura 14, que possui uma série de recursos para gerenciar os mecanismos de saída de texto na tela após a fala.

Para que uma instância do objeto *SpeechRecognition* possa reconhecer uma fala, é necessário que este já tenha concluído sua inicialização. Dessa forma, uma classe que utiliza SR deve implementar o método *start()* da interface *SpeechRecognition* para ser notificado quando a inicialização estiver completa.

Alguns atributos também são muito importantes quanto ao gerenciamento do SR, como o *lang*, *continuous*, *maxAlternatives* no qual o primeiro define a linguagem a ser utilizada, o segundo define se o tipo de reconhecimento é contínuo ou não, e o terceiro define o número máximo de alternativas, e o atributo *onend* que é para a leitura em execução, após que o reconhecimento de fala é iniciado, há muitos manipuladores de eventos que podem ser usados para recuperar resultados e outras

informações adjacentes. O mais comum que você provavelmente usará é *SpeechRecognition.onresult*, que é acionado quando um resultado bem-sucedido é recebido.

5.1.2 Desenvolvimento do Protótipo

Com base na aplicação desenvolvida em uma primeira etapa sem a implementação da tecnologia de reconhecimento, iniciou-se o desenvolvimento do protótipo. Baseado na experiência obtida no uso normal do protótipo, foram elencados os passos para a criação do protótipo, iniciando uma nova etapa, onde foram aplicadas as implementações utilizadas de reconhecimento de fala, agora aplicando conceitos de usabilidade e modo de interação com o usuário.

Alguns dos principais itens elencados foram: movimentação do cursor entre os espaços vazios, botões acessíveis, e disponibilizar saídas na tela por meio de texto.

5.1.2.1 Movimentação do Cursos Entre os Espaços Vazios

Para que o processo de aprendizado de escrita possa ter um bom resultado são usados vários métodos ou métricas de ensino, trouxemos alguns métodos aqui dentro do nosso protótipo usando o reconhecimento de fala, um dos métodos escolhidos é o preenchimento de palavras inacabadas. Neste método as palavras vêm com espaços vazios entre as letras e o usuário completa as palavras com as falas.

Neste caso, foi desenvolvido o método `palavraChange()` que permitisse que o cursor se deslocasse entre os espaços vazios da palavra enquanto o usuário digita ou fala a letra que acha a correta como podemos ver na figura 15.

Figura 15 - Método para Alteração da Posição do Cursor

```
palavraChange(palavra: String, letra: String){  
  
    for (let index = 0; index < palavra.length; index++) {  
  
        if (palavra[index] == " "){  
  
            var texto = palavra[index];  
            palavra = palavra.replace(texto, letra[0]);  
  
            return palavra;  
        }  
  
    }  
  
    return palavra;  
}
```

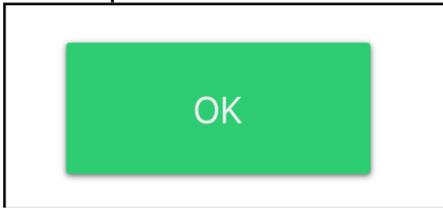
Fonte: Do Autor

5.1.2.2 Acessibilidade nos Botões

A acessibilidade é um dos elementos chaves do nosso trabalho, dessa forma, visando apoiar os usuários deficientes ou seja, usuários com restrição motora nos membros superiores na execução das atividades do protótipo, buscou-se a utilização de um recurso disponível na maioria das aplicações que são a utilização de botões acessíveis, permitindo assim o usuário que possui dificuldade em utilizar os teclados e mouses convencionais a usarem o mesmo.

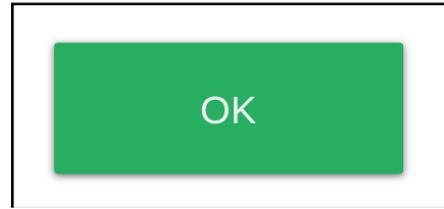
Como recurso visual, foi desenvolvido com uma borda *radius* de 2px para os contornos, podendo esta estar na cor verde clara, quando desfocado, e verde mais escuro, quando em foco. Tal características foi implementado no arquivo de estilo do componente em que o mesmo se encontra, o qual é responsável por mudar as cores de acordo com a movimentação do mouse por cima do botão, como ilustrado nas figuras 16 e 17.

Figura 16 - Botão sem o mouse por cima
mouse por cima



Fonte 1: Do Autor

Figura 17 - Botão com o



Fonte 2: Do Autor

5.1.3 Teste de Validação

Após a conclusão do desenvolvimento do protótipo da aplicação web, iniciaram-se os testes com intuito de avaliar o mesmo. Para isso, foi elencado um método distinto de validação de dado de forma objetiva.

Para uma avaliação objetiva, o protótipo foi submetido a testes, em diferentes cenários de utilização, a fim de medir a eficiência.

5.1.3.1 Teste de Eficiência

Durante o desenvolvimento do protótipo, em testes informais, foram percebidos fatores importantes que influenciavam diretamente no bom desempenho do reconhecimento de fala. Dentre os mais relevantes estão os ruídos. Sendo assim, durante a elaboração dos cenários utilizados para os testes, foi adicionado o fator ruídos em diferentes condições, buscando medir sua influência sobre o processo de reconhecimento.

Os testes de eficiência se basearam em dois diferentes cenários, sendo que em cada um deles havia ruídos em ambiente diferentes. No cenário B, o ruído é reduzido, ou seja, dentro de um recinto com as portas fechadas, podendo assim capturar apenas a fala do usuário. No cenário A, o ruído é mais presente, pois nesse recinto era um ambiente aberto, possuía varias pessoas falando ao mesmo tempo, e nesse exato momento o usuário tenta interagir com o protótipo.

Na execução dos testes, foram realizadas dez leituras para cada cenário e cada leitura teve seu tempo cronometrado e seus resultados foram colocadas na tela e na console do navegador. Cada fala reconhecida foi comparado a fala normal da mesma, para a apuração do percentual de erros ocasionados no processo.

Com os resultados coletados da execução dos testes, elaborou-se a tabela 2, de forma a tabular as informações e facilitar a avaliação dos resultados.

Tabela 2 - Resultado dos testes de acordo a cada cenário

Teste	Cenário A		Cenário B	
	% Erro	Tempo Pro.	% Erro	Tempo Pro
Teste 1	25,00%	00:00: 10,65	10,83%	00:00:4,25
Teste 2	19,64%	00:00: 10,53	09,42%	00:00:3,95
Teste 3	19,19%	00:00: 8,36	08,09%	00:00:3,88
Teste 4	19,73%	00:00: 6,98	08,73%	00:00:3,83
Teste 5	16,67%	00:00: 6,82	07,87%	00:00:3,75
Teste 6	13,34%	00:00: 6,36	07,34%	00:00:3,00
Teste 7	13,06%	00:00: 5,73	07,17%	00:00:2,90
Teste 8	12,89%	00:00: 5,62	05,99%	00:00:2,75
Teste 9	12,66%	00:00: 5,31	04,75%	00:00:2,71
Teste 10	11,77%	00:00: 4,62	04,70%	00:00:2,00
Media	31,84%	00:00:7,09	07,57%	00:00:3,30

Fonte: Do Autor

Com base nos resultados obtidos, é possível perceber a direta influência do ruído para com a eficiência do reconhecimento da fala. No cenário A, avaliando cada cenário, é possível perceber que os erros de reconhecimento são predominantes devido ruídos, ocasionando assim uma média de 31,84% de erros contabilizados nesse cenário.

No cenário B os resultados já são bem mais satisfatórios, apurando-se uma média de apenas 07,57% de erros. Porém o tempo de processamento da fala foi menor em relação ao outro cenário, com tempo médio de duração de 3,30 segundos.

5.2 RESULTADOS OBTIDOS

Com a pesquisa realizada neste trabalho e a aplicação dos fundamentos obtidos através das pesquisas bibliográficas, foi possível obter um protótipo de aplicação web, capaz de realizar a conversão da voz em texto orientado ao aprendizado de escrita, utilizando a Web Speech API.

Através da realização de testes de eficiência, foi possível constatar que o a API de reconhecimento de voz usada possui uma ótima precisão e funciona perfeitamente em conjunto com a plataforma Web. Porém foi possível detectar uma sensibilidade à ruídos, o que é comum a qualquer aplicação com reconhecimento de voz. Sendo assim, a aplicação de técnicas de pro e pré-processamento do reconhecimento das falas, conforme pesquisas descritas no sub-capítulo 3.2.1, pode ser uma alternativa para amenizar tais problemas gerados no processo do reconhecimento da fala.

Sendo assim, o protótipo apresentou-se ideal para uma possível utilização em ambientes de aprendizagem de escrita, como escolas, instituições, entre outros, onde um usuário efetua uma fala simples e então protótipo disponibiliza a informação em formato de texto para o usuário.

Os recursos de acessibilidade pesquisados foram muito importantes para que isso fosse possível. Funções como posicionar o cursor nos campos vazios das palavras, isso facilitaria os usuários com dificuldades nos membros superiores.

O protótipo se mostrou uma solução de simples aplicação, fornecendo acessibilidade com baixo custo. É importante ressaltar a facilidade de acesso a tecnologia, por se tratar de uma aplicação web, e um microfone embutido ou externo. Sendo assim, quase todos os objetivos propostos foram atingidos, resultando em um protótipo de aplicativo, com diversas possibilidades de aprimoramento e em uma linha de pesquisa que permite a aplicação de diversos recursos tecnológicos diferentes e ainda assim muito carente de novos recursos.

6 CONCLUSÃO

A execução do presente trabalho de conclusão de curso proporcionou adquirir conhecimentos em diversas áreas. Foi possível perceber a evolução dos navegadores web e o potencial das aplicações web atualmente, tal como uma breve disposição dos recursos por ela suportada. A acessibilidade foi o principal conceito que direcionou essa pesquisa, levando a aplicação das diversas outras técnicas abordadas neste trabalho.

As técnicas e tecnologias abordadas nesse trabalho, foram aplicadas buscando atender um objetivo principal: criar um protótipo de aplicação web para que usuários possam obter um aprendizado de escrita ou das letras, de forma acessível, visando auxiliar pessoas com deficiência nos membros superiores ou pessoas com restrições na coordenação motora nos mesmos.

Os objetivos específicos elencados para a execução desse trabalho foram todos atingidos, os quais foram responsáveis por orientar a pesquisa e o desenvolvimento, além de permitir que o objetivo geral do trabalho também fosse alcançado.

Com o estudo das metodologias de acessibilidade para pessoas com deficiências nos membros superiores, para ambiente *desktop*, adquiriu-se embasamento teórico imprescindível para a determinação dos métodos de interação com o usuário aplicados ao protótipo. Uma principal tecnologia esteve diretamente ligada ao objetivo geral, sendo ela a Web Speech API (WSA).

Como foi abordado no capítulo 3.3 do presente trabalho, a WSA possui duas funcionalidades principais, sendo elas a conversão da voz em texto e a conversão do texto em voz, que foi usada neste protótipo foi a conversão da voz em texto, sendo que existem outras ferramentas e tecnologias que fazem a mesma tarefa, a WSA mostrou-se uma tecnologia propícia para o desenvolvimento em aplicações web, visto que ela é grátis e possui uma ótima documentação e uma boa comunidade. Além disso, em testes de reconhecimento, a ferramenta apresentou resultados muito satisfatórios, com boa precisão em diferentes cenários.

Durante o levantamento bibliográfico, buscou-se compreender o funcionamento da classe *SpeechRecognition* um recurso da API para conversão de voz em texto. A mesma se mostrou muito robusta e eficiente, pois ela no seu melhor cenário apresentou uma média de 3,30% segundos no tempo de resposta,

disponibilizando uma série de recursos que tornam possível trazer acessibilidade ao desenvolvimento de diversas aplicações.

As pesquisas realizadas sobre as formas de interação com os usuários do protótipo permitiram que técnicas fossem aplicadas a este, prevendo possíveis cenários e situações decorrentes de sua utilização, com base nas metodologias de acessibilidade para pessoas com deficiência nos membros superiores. Foram também fundamentais os trabalhos correlatos nessa linha de pesquisa, agregando muito conhecimento ao trabalho desenvolvido.

Desenvolveu-se um protótipo de aplicativo utilizando a plataforma Web, o qual fez a integração com a API, permitindo assim o reconhecimento de fala e a conversão de voz em texto, com base nos conceitos de acessibilidade e suas metodologias, sendo assim capaz de realizar a captura da fala para apoio a pessoas com deficiências nos membros superiores ou coordenação motora reduzidas.

Foi avaliada a eficiência do protótipo por nós, através de testes em dois cenários, que mostraram um bom desempenho no processo de reconhecimento de fala, com tempo médio de execução de 3,30 segundos e média de 7,57% de erro no melhor cenário. Os testes ainda demonstraram a sensibilidade que o processo demonstra com relação a fatores provenientes da etapa de reconhecimento da fala, como ruídos, os quais podem ser amenizados com aplicação de técnicas de pré-processamento do reconhecimento de voz, segundo teorias e praticas pesquisadas durante o levantamento bibliográfico.

Durante o processo de desenvolvimento do protótipo, algumas dificuldades foram encontradas, porém compensadas com o embasamento teórico adquirido no levantamento bibliográfico realizado. As objeções no decorrer do desenvolvimento da plataforma web, foi compensada por uma boa documentação disponível referente o uso da tecnologia. No instante da captura da fala houve dificuldades em algumas questões, como pegar a fala do usuário, mas que conseguiram ser superadas no próprio protótipo, aplicando recursos e conceitos disponíveis na documentação e em artigos encontrados na internet.

Como sugestão para trabalhos futuros, pode-se aplicar técnicas de funcionamento em dispositivos moveis, com intuito de tornar disponíveis para todos possíveis usuários. Pode-se ainda atribuir diversas funcionalidades ao aplicativo, usando novas soluções de forma criativa. Alguns exemplos de funcionalidades passíveis de implementação:

- a) o retorno da fala reconhecida por meio de libras, não só por meio de texto;
- b) criar opção de exportar todas atividades que ele realiza no protótipo por meio de fala em um arquivo de texto;
- c) aplicação de técnicas de pré-processamento do reconhecimento de voz para amenizar os ruídos no processo de reconhecimento da fala;
- d) aplicar novos métodos de ensino de escrita por meio de voz, como aprendizado de sílabas por exemplo.

_____, Lei nº 12.527, de 18 de novembro. **Avaliador e Simulador de Acessibilidade em Sítios**: Acessibilidade. 2011. Disponível em: <<http://asesweb.governoeletronico.gov.br/ases/acessibilidade>>. Acesso em: 14 de abr 2018.

_____. Lei nº 13.146, de 6 de julho de 2015. Capítulo II, **Do acesso à informação e à comunicação**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm?tspd_101_r0=da10733bbe9090d87fdafad3b92fbca4z2500000000000000be58bbbffff000000000000000000000000000005b06dbae00d145faef>. Acesso em: 24 maio. 2018.

_____. Subsecretaria Nacional de Promoção dos Direitos da Pessoa com Deficiência. Comitê de Ajudas Técnicas. Tecnologia Assistiva. – Brasília: CORDE, 2009. 138 p.

CAIADO, Katia Regina Moreno. **Aluno com deficiência visual na escola; Lembranças e Depoimentos**. - 3. ed. - Campinas, SP: Autores Associados, 2014.

CAMARGO, Mário Lázaro; GOULART JÚNIOR, Edward; LEITE, Lúcia Pereira. O Psicólogo e a Inclusão de Pessoas com Deficiência no Trabalho. **Psicologia: Ciência e Profissão**, [s.l.], v. 37, n. 3, p.799-814, set. 2017. Fap UNIFESP (SciELO). <http://dx.doi.org/10.1590/1982-3703003232016>.

CARVALHO, José Roberto e TURECK, Lucia Terezinha Zanato. **Algumas reflexões sobre a inclusão escolar de alunos com deficiência**. Pessoa com deficiência na sociedade contemporânea: problematizando o debate/Org. do programa de ações relativas às pessoas com necessidades especiais– PEE – Cascavel: Edunioeste, 2006, p. 63-90.

CAT - Comitê de Ajudas Técnicas. **Ata da Reunião VII, de dezembro de 2007 do Comitê de Ajudas Técnicas**. Secretaria Especial dos Direitos Humanos da Presidência da República (CORDE/SEDH/PR), 2007. Disponível em: <<http://www.mj.gov.br/corde/comite.asp>> Acesso em: 16 set. 2009.

CHRISTENSSON, Per. **Web Applications Definition**. TechTerms. Disponível em: <https://techterms.com/definition/web_application> Acesso em: 05 maio. 2018.
 COLE, R. A., MARIANI, J., USZKOREIT, H., ZAENEN, A. and ZUE, V., “Survey of the State of the Art in Human Language Technology”, *Cambridge University Press*, Cambridge, UK, 1997 ([http://cslu.cse.ogi.edu/HL T survey](http://cslu.cse.ogi.edu/HL_T_survey)).
 CLICK. Tecnologia Assistiva. **Produtos para o uso do computador mais acessível e independente**. 2018. Disponível em: <http://www.clik.com.br/clik_01.html>. Acesso em: 02 abr. 2018.

DESKTOP TECHNOLOGY SERVICES LIMITED. **Mouse Por Movimento de Cabeça**. 2018. Disponível em: <<http://assistive.dtsl.co.nz/products/15687-headmouse-extreme.aspx>>

DICTATION PRO. **Sobre nós**. 2018a. Disponível em: <https://pt.downloadastro.com/f/sobre_nós/> Acesso em 20 abril 2018.

_____. **Tela de Digitação do Dictation Pro**. 2018b. Disponível em: <
https://dictation_pro.pt.downloadastro.com>

DUSIK, Cláudio. **Teclado Virtual Silábico-Alfabético**: Tecnologia Assistiva para Pessoas com Deficiência Física. 2013. 192 p. Dissertação (Mestrado)- Faculdade de Educação - FAGED, Universidade Federal do Rio Grande, Porto Alegre, 2013. Disponível em:
 <http://bdt.d.ibict.br/vufind/Record/URGS_52fc992b9f8792fd14138a6fba85ddb5>. Acesso em: 29 mar. 2018.

FIGENBAUM, Joseane. **Acessibilidade no contexto escolar**: tornando a inclusão possível. Porto Alegre, 2009.

FUNDAÇÃO INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Normas de apresentação tabular**. 3. ed. Rio de Janeiro: IBGE, 1993. Disponível em: <
<http://biblioteca.ibge.gov.br/visualizacao/livros/liv23907.pdf>>. Acesso em: 18 maio 2016.

GONÇALVES, Bento. **Manual Do Headmouse**: Acessibilidade Virtual, Rio Grande do Sul, v.1, n.1, p.1-9, nov. 2015. Disponível em:
 <<http://cta.ifrs.edu.br/files/doc/04aa2c200e123eff7680d12e488a81a3.pdf>>. Acesso em: 20 mar. 2018.

GOOGLE. **Ajuda do Editores do Documentos**: Digitar com sua voz. Disponível em: <
<https://support.google.com/docs/answer/4492226?hl=pt-BR>>. Acesso em: 23 abr. 2018a.

_____. **Reconhecimento Poderoso de Fala**. Disponível em: <
<https://cloud.google.com/speech-to-text/>>. Acesso em: 03 maio. 2018b.

GOOGLE CLOUD. **Preços**. Disponível em: <
<https://cloud.google.com/speech/pricing?hl=pt-br>>. acesso em: 03 maio. 2018.

GUARALDILAB. **Deficiência e Desenvolvimento**: Como Incluir as Pessoas Com Deficiência na Cooperação para o desenvolvimento. 2015. Disponível em: <
http://www.ridsnetwork.org/files/2015/06/disabilitaesviluppo_por.pdf>. Acesso em: 17 abr. 2018.

HUGO, Marcel. **Uma Interface de Reconhecimento de Voz para o Sistema de Gerenciamento de Central de Informação de Fretes**. 1995. 60 f. Dissertação (Mestrado) - Curso de Engenharia de Produção, Programa de Pós-graduação, Universidade Federal de Santa Catarina, Florianópolis, 1995. Cap. 3. Disponível em: <
<https://repositorio.ufsc.br/bitstream/handle/123456789/157911/101794.pdf?sequence=1>>. Acesso em: 18 jun. 2018.

IBDD. **Inclusão Social da Pessoa com Deficiência**: medidas que fazem a diferença. 1ª edição. Rio de Janeiro, 2008.

INTERNET LIVE STATS. **Total Number of Websites**. Disponível em: <<http://www.internetlivestats.com/total-number-of-websites/>>. Acesso em: 12 mar. 2018.

ITS BRASIL. **Tecnologia assistiva nas escolas**: recursos básicos de acessibilidade sócio digital para pessoas com deficiência. Juiz de Fora - Minas Gerais, 2008. Disponível em: <http://www.ufjf.br/ acessibilidade/files/2009/07/Cartilha_Tecnologia_Assistiva_nas_escolas_-_Recursos_basicos_de_acessibilidade_socio-digital_para_pessoal_com_deficiencia.pdf>. Acesso em: 08 jun. 2018.

JUNIOR, Edson Alves de Oliveira; FORTES, Renata Pontin de Mattos. **Arquitetura de Software na Web Atual**: Processamento no Servidor. Disponível em: <http://conteudo.icmc.usp.br/CMS/Arquivos/arquivos_enviados/BIBLIOTECA_113_ND_78.pdf>

LUCIANO, Josué; ALVES, Wallison Joel Barberá. Padrão de Arquitetura MVC: Model-View-Controller. **Revista Epeq Fafibe**, Bebedouro -sp, v. 1, n. 3, p.102-107, 2011. Semanal. Original. Disponível em: <<http://www.unifafibe.com.br/revistasonline/arquivos/revistaepqfafibe/sumario/20/16112011142249.pdf>>. Acesso em: 5 maio 2018.

MACIEL, MARIA REGINA CAZZANIGA. Portadores de deficiência: a questão da inclusão social. **São Paulo Perspec.**, São Paulo , v. 14, n. 2, p. 51-56, June 2000 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-88392000000200008&lng=en&nrm=iso>. access on 12 Dec. 2018. <http://dx.doi.org/10.1590/S0102-88392000000200008>.

MARTINS, José Antônio. **Avaliação de diferentes técnicas para reconhecimento de fala**. Campinas, SP: Universidade Estadual de Campinas – UNICAMP, 1997, (Tese, Doutorado). Disponível em: <http://www.decom.fee.unicamp.br/lpdf/teses_pdf/Tese-Doutorado- Jose_Martins.pdf>. Acesso em 13/03/2011.

MATRAS, Jean-Jacques. **O som**. São Paulo: Martins Fontes, 1991.

MAZZOTTA, Marcos José da Silveira; D'ANTINO, Maria Eloísa Famá. Inclusão social de pessoas com deficiências e necessidades especiais: cultura, educação e lazer. **Saude soc.**, São Paulo , v. 20, n. 2, p. 377-389, June 2011 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-12902011000200010&lng=en&nrm=iso>. access on 12 Dec. 2018. <http://dx.doi.org/10.1590/S0104-12902011000200010>.

MENDES, E. G. **A radicalização do debate sobre inclusão escolar no Brasil**. Revista Brasileira de Educação, Rio de Janeiro, v. 11, n. 33, p. 387-405, 2006. Disponível em: <<http://www.scielo.br/scielo =S1413-24782006000300002&lng=pt&nrm=iso>>. Acesso em: 15 mar. 2008.

MICROSOFT. Windows 11 SDK: multimedia programmer's guide. Microsoft Development Library, Disponível em: < [https://msdn.microsoft.com/en-us/library/windows/desktop/dd757576\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/dd757576(v=vs.85).aspx) >. Acesso em 18 Apr. 2018.

MICROSOFT. O que é o Pacote Office?. 2018, Disponível em: <<https://products.office.com/pt-br/home>>. Acesso em 18 Apr. 2018.

MOZILLA DEVELOPER. **Introdução ao lado do servidor**. 2018a. Disponível em: <https://developer.mozilla.org/en-US/docs/Learn/Server-side/First_steps/Introduction>. Acesso em: 19 maio 2018.

_____. **Web Speech API**. 2018b. Disponível em: <https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API>. Acesso em: 29 maio 2018.

NETCRAFT. **March 2018 Web Server Survey**. Disponível em: <<https://news.netcraft.com/archives/category/web-server-survey/>>. Acesso em: 10 mar. 2018.

NETO, Clovis. **Web Speech API: Reconhecimento de voz com JavaScript**. Disponível em: <<https://tableless.com.br/web-speech-api-reconhecimento-de-voz-com-javascript/>>. Acesso em: 19 mar. 2018.

OLIVEIRA, Ana Irene Alves de; GAROTTI, Marilice Fernandes; SÁ, Nonato Márcio Custódio A Maia. Tecnologia de ensino e tecnologia assistiva no ensino de crianças com paralisia cerebral. **Ciências & Cognição**, Belém, Pará, v. 13, n. 3, p.243-262, 10 dez. 2008.

PAIXÃO, Luciana. **A importância do Projeto de Acessibilidade no Projeto Arquitetônico**, 2014. Disponível em: <<https://www.aarquiteta.com.br/blog/projetos-de-arquitetura/acessibilidade-importancia/>>. Acesso em: 22 mar. 2018.

PALLETT, D.S., et al., "1997 Broadcast News Benchmark Test Results: English and Non- English," *Proc. of the Broadcast News Transcription and Understanding Workshop*, 1998, Lansdowne, Virginia, Morgan Kaufmann Publishers.

PESSOTI, Isaías. **Deficiência mental: da superstição a ciência**. São Paulo: T. A. Queiroz - Editora da Universidade de São Paulo, 1984.

PRODOCTOR. **Dia do Deficiente Físico**. Disponível em: <<https://prodoctor.net/blog/2016/10/aplicativos-melhoram-qualidade-de-vida-de-deficientes/>>. Acesso em: 12 mar. 2018.

RAW COMPUTING. **Reconhecimento de Fala do Windows 7**. 2016. Disponível em: <<http://rc.rawinfopages.com/windows7/windows-speech-recognition.html>>

RONCHENTTI, Eduardo. **A importância da Acessibilidade**. Disponível em: <<https://eduardoronchetti.wordpress.com/2013/07/31/a-importancia-de-acessibilidade/>>. Acesso em: 14 mar. 2018.

SCIENCESOFT. **Web application architecture: components, models and types**. Disponível em: <<https://www.scnsoft.com/blog/web-application-architecture>>. Acesso em: 05 mai. 2018.

SHIRES, Glen. (EUA). Speech Api Community Group (Ed.). **Web Speech API Specification**. 2014. Disponível em: <<https://w3c.github.io/speech-api/speechapi.html>>. Acesso em: 29 maio 2018.

SILVA, Otto, M. **A epopéia ignorada; a pessoa deficiente na história do mundo de ontem e de hoje**. São Paulo: Cedas, 1987.

SILVA, Thaysa Danyella Lira da; SILVA, Edcleide Maria da. **Mas o que é mesmo Corpus?** Alguns Apontamentos sobre a Construção de Corpo de Pesquisa nos Estudos em Administração. In: ENANPAD, 27., 2013, Rio de Janeiro. **Anais...** Rio de Janeiro: Anpad, 2013. v. 1, p. 1 - 15. Disponível em: <http://www.anpad.org.br/admin/pdf/2013_EnANPAD_EPQ1021.pdf>. Acesso em: 4 jun. 2018.

SUGITA, Kaoru; YOKOTA, Masao. Implementation of Voice Guidance Function for Sightseeing Contents Using Web Speech API. **2016 10th International Conference On Innovative Mobile And Internet Services In Ubiquitous Computing (imis)**, [s.l.], p.133-138, jul. 2016. IEEE. <http://dx.doi.org/10.1109/imis.2016.106>. Disponível em: <<https://ieeexplore.ieee.org/document/7794452/>>. Acesso em: 29 maio 2018. SPEECHNOTES. **Tela Principal do Speeachnotes**. 2018. 1 Captura de Tela.

STOLZ, Alex. The Sound Blaster book. USA: Abacus, 2 Edição. 1993.

TECASSISTIVA. **Acionadores Especiais**. Disponível em: <<http://www.tecassistiva.com.br/component/spidercatalog/showproduct/492/73>>. Acesso em: 15 mar. 2018.

TECHTUDO. **Reconhecimento de Fala do Google Docs**. 2018. Disponível em: <<http://www.techtudo.com.br/dicas-e-tutoriais/noticia/2015/09/como-escrever-textos-usando-voz-no-google-docs-em-portugues.html>>

TECNO ACESIBLE. **Teclado IntelliKeys**. 2018. Disponível em: <<https://tecnoaccesible.net/content/intellikeys>>

TEIXEIRA, Màyra Mundim; SOUZA, Ricardo Luíz Mendonça. JAVA SPEECH COMO PROPOSTA DE IMPLEMENTAÇÃO DE TECNOLOGIA ASSISTIVA PARA PORTADORES DE DEFICIÊNCIA AUDITIVA. **E-rac**, Uberlândia - Mg, v. 2, n. 1, p.1-14, mar. 2012. Disponível em: <<http://www.computacao.unitri.edu.br/erac/index.php/e-rac/article/view/90/0>>. Acesso em: 25 maio 2018.

TEVAH, Rafael Teruszkin. **Implementação De Um Sistema De Reconhecimento De Fala Contínua Com Amplo Vocabulário Para O Português Brasileiro**. 2006. 91 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2006. Disponível em: <<http://pee.ufrj.br/teses/textocompleto/2006053001.pdf>>. Acesso em: 29 maio 2018.

THOMAZ, Raphael Henriques. **Tutorial do Software Motrix – Instalação e Utilização**. Rio ResvISta, Faculdade De Educação Tecnológica Do Estado Do Rio De Janeiro, 7ª Edição, 2016.

VALIATI, Joao Francisco. **Reconhecimento de voz para comandos de direcionamento por meio de redes neurais**. 2000. 129 f. Dissertação (Mestrado) - Curso de Ciência da Computação, Instituto de Informática, Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, 2000. Disponível em: <<http://hdl.handle.net/10183/2947>>. Acesso em: 25 maio 2018.

YNOGUTI, Carlos Alberto. **Reconhecimento de Fala Contínua Usando Modelos Ocultos de Markov**. 1999. 138 f. Tese (Doutorado) - Curso de Engenharia Elétrica, Departamento de Comunicações, Universidade Estadual de Campinas, Campinas, 1999. Disponível em: <http://www.decom.fee.unicamp.br/lpdf/teses_pdf/Tese-Doutorado-Carlos_Alberto_Ynoguti.pdf>. Acesso em: 28 maio 2018.

ZHOU, Yolanda. **Web Design & UI/UX**. 2017. 57 slides. Material Apresentado no Workshop 6148 do Instituto De Tecnologia De Massachusetts (MIT).

APÊNDICE(S)

APÊNDICE A – ARTIGO CIENTÍFICO

PROTÓTIPO DE UMA APLICAÇÃO WEB USANDO RECONHECIMENTO DE VOZ PARA CONVERTER VOZ EM TEXTO, ORIENTADO AO APRENDIZADO DE ESCRITA PARA PESSOAS QUE POSSUEM DIFICULDADES NO USO DO TECLADO**Aristote Kanza Lopes¹, Fabricio Giordani²**

¹ Curso de Ciência da Computação
Universidade do Extremo Sul Catarinense (UNESC) – Criciúma, SC – Brazil

{aristo.lopez, fgiordani}@gmail.com

Abstract. *This article aims to develop a prototype web application to capture the user's speech from the microphone on the computer and convert it into text, in order to support the learning of writing of people who have difficulties using the conventional keyboard. For that, a theoretical foundation was made regarding the techniques of speech recognition, responsible for recognizing user speech, accessibility methodologies in web applications, among others. With the knowledge obtained from the research, it was possible to elaborate a web application prototype using the Web Speech API tool, which converts from voice to text.*

Resumo. *Este artigo tem como objetivo desenvolver um protótipo de aplicação web para capturar a fala do usuário a partir do microfone no computador e convertê-lo em texto, com intuito de apoiar o aprendizado de escrita de pessoas que possuem dificuldades no uso do teclado convencional. Para isso, foi realizada uma fundamentação teórica referente as técnicas de reconhecimento de fala, responsável por reconhecer a fala do usuário, metodologias de acessibilidade em aplicações web, dentre outros. Com os conhecimentos obtidos a partir da pesquisa, foi possível elaborar um protótipo de aplicação web, utilizando a ferramenta Web Speech API, que faz a conversão de voz em texto.*

1. INTRODUÇÃO

A escrita tem um papel muito importante no desenvolvimento social e intelectual humano. A partir do ingresso na vida escolar, ela se torna imprescindível para o acesso ao conhecimento e a informação. Porém, muitas vezes o acesso a esse meio se torna difícil, principalmente para deficientes com coordenação motora reduzidas que torna difícil o uso dos membros superiores, seja por falta de recursos tecnológicos, como teclados e mouses especializados, ou algum outro recurso que permita com que ele possa aprender a escrever.

Atualmente, o mercado vem resolvendo vários tipos de problemas por meio digital, e a educação vem acompanhando a tecnologia nesse sentido. Porém, existe uma quantidade muito grande de recursos importantes e continuam a ser desenvolvidos recursos para auxiliar no aprendizado ou na possibilidade de escrita de pessoas com mobilidades reduzidas, quer seja recursos físicos ou digital.

Com a evolução das tecnologias, diversos recursos surgem buscando tornar atividades cotidianas mais fáceis, tal como promover alternativas para a resolução de problemas rotineiros.

A aplicação desses recursos na promoção da acessibilidade a pessoas portadoras de deficiências recebe o nome de Tecnologia Assistiva. O objetivo dessas tecnologias é a fabricação, uso de equipamentos, recursos ou estratégias utilizadas para capacitar as habilidades funcionais das pessoas com deficiências (Brasil, 2009).

A popularização e a evolução da tecnologia das aplicações web tornam possíveis a elaboração de novas soluções para auxiliar a execução de tarefas, inclusive na aplicação como Tecnologia Assistiva. Recursos como Reconhecimento de voz, *Application Program Interfaces* (API) de conversão de áudio em texto, captura de pequenas trechos de fala com precisão, entre outros, tornam possível a elaboração de inúmeros tipos diferentes de aplicações, para as mais variadas finalidades.

Este trabalho teve como objetivo utilizar tecnologia de Reconhecimento de voz e conversão de áudio em texto para a elaboração de um protótipo de aplicação web, voltado para o apoio de escrita e aprendizado de escrita por parte de usuários com coordenações motoras reduzidas.

2. METODOLOGIA

Para a concretização deste trabalho foi feito um levantamento bibliográfico relacionado a acessibilidade para deficientes com coordenação motora reduzidas, dificuldades nos membros superiores, reconhecimento de voz e conversão de voz em texto. Sendo assim, foi proposto o desenvolvimento de um protótipo de uma aplicação Web para conversão de voz em texto. Foram aplicadas ainda formas de interação com o usuário do aplicativo, também adquiridas no processo de fundamentação teórica.

As operações relacionadas ao reconhecimento de voz e a conversão de voz em texto que será responsabilidade da Web Speech API, será implementado usando a linguagem de programação ECMAScript. Primeiramente será feito uma análise do funcionamento dos sistemas de reconhecimento de voz existentes e em seguida será feito uma outra análise com relação a conversão de voz em texto com a Web Speech API, após essa análise será definido a forma de interação do usuário com aplicação e dando início ao do desenvolvimento do protótipo.

O teste será realizado em ambiente habitacional para poder avaliar a eficiência do protótipo, usando microfone nativo do notebook e um navegador.

2.1 A TECNOLOGIA DE RECONHECIMENTO DE VOZ WEB SPEECH API

A tecnologia escolhida para a realização do reconhecimento de voz foi a Web Speech API, pois além de ser *open-source* com um excelente histórico de desenvolvimento, tendo na sua lista de contribuintes empresa como a *Google*, ainda apresentou uma pré-disposição para a utilização em aplicações web. Na sua página oficial, os navegadores Chrome, Edge, Firefox são citadas dentre as suportadas pela tecnologia (MOZILLA DEVELOPER, 2018, tradução nossa).

A tecnologia Web Speech API tem como princípio básico de funcionamento receber uma fala, submetê-la ao processo de reconhecimento de vozes e então retornar a fala convertida em texto. A Web Speech API tem duas partes: Síntese de Fala (*Text-to-Speech*) e Reconhecimento de Fala (*Asynchronous Speech Recognition*).

O reconhecimento de fala é acessado por meio da interface *SpeechRecognition*, que fornece a capacidade de reconhecer o contexto de voz a partir de uma entrada de áudio (normalmente por meio do serviço de reconhecimento de fala padrão do dispositivo) e responder apropriadamente. Geralmente, você usará o construtor da interface para criar um

novo objeto *SpeechRecognition*, que possui vários manipuladores de eventos disponíveis para detectar quando a fala é inserida através do microfone do dispositivo. A interface *SpeechGrammar* representa um contêiner para um determinado conjunto de gramática que seu aplicativo deve reconhecer. A gramática é definida usando o formato de gramática *JSpeech* (JSGF).

2.1.1 Captura de Fala

Foram feitas pesquisas sobre a captura de fala em computadores, e foi constatada duas formas de realizar essa tarefa, utilizando um microfone nativo do computador ou microfones adicionais.

O reconhecimento de fala envolve o recebimento de fala por meio do microfone de um dispositivo, que é verificado por um serviço de reconhecimento de fala em relação a uma lista de gramática (basicamente, o vocabulário que você deseja reconhecer em um aplicativo específico). É retornado como resultado (ou lista de resultados) como uma cadeia de texto, e outras ações podem ser iniciadas como resultado.

2.1.2 Utilizando a Web Speech API

Com o código pronto para funcionar usando somente o teclado convencional, foi possível iniciar o desenvolvimento do protótipo utilizando a API. A princípio foi criada uma classe, ou seja, um *service* denominado *SpeechRecognitionService* (SRS), visto que usaríamos o serviço de reconhecimento de voz da API declaramos os métodos necessários no desenvolvimento do nosso protótipo como podemos ver na figura 12.

Figura 13 - Classes da Web Speech API

```
interface IWindow extends Window {
    webkitSpeechRecognition: any;
    SpeechRecognition: any;
}
```

Fonte - Do Autor

A Web Speech API possui duas classes principais, uma delas é a *webkitSpeechRecognition* ou a *SpeechRecognition* (SR) e a classe *SpeechGrammar* ou *webkitSpeechGrammarList*. Dessa forma, ao instanciar um objeto dos tipos citados acima, muitos dos recursos originais da API já se tornam disponíveis como observamos na figura 13.

Figura 13 - Instância do objeto webkitSpeechGrammarList

```
const { webkitSpeechRecognition }: IWindow = <IWindow>window;
this.speechRecognition = new webkitSpeechRecognition();
//this.speechRecognition = SpeechRecognition;
```

Fonte: Do Autor

Para realizar um reconhecimento de fala com a Web Speech API, se faz necessário seguir alguns passos básicos. Com a instância *speechRecognition* criada, é necessário chamar o método *speechRecognition.start()*, definir os atributos mais importantes como o tipo de reconhecimento de fala que será usado, a linguagem desejada. É necessário também definir o número máximo de alternativas. O reconhecimento de fala pode ser feita através do método *onResult()* onde tem toda uma lógica para transformar a voz em texto, e aí temos outros métodos que poderão auxiliar no processo de reconhecimento de fala. O procedimento descrito pode ser observado no trecho de código-fonte conforme figura 14.

Figura 14 - Trecho de código para a inicialização da *SpeechRecognition*

```

this.speechRecognition.continuous = true;
//this.speechRecognition.interimResults = true;
this.speechRecognition.lang = 'pt-br';
this.speechRecognition.maxAlternatives = 1;

this.speechRecognition.onresult = speech => {
  let term: string = "";
  if (speech.results) {
    var result = speech.results[speech.resultIndex];
    var transcript = result[0].transcript;
    if (result.isFinal) {
      if (result[0].confidence < 0.3) {
        console.log("Unrecognized result - Please try again");
      }
      else {
        term = _.trim(transcript);
        console.log("Did you said? -> " + term + " , If not then say something else...");
      }
    }
  }
  this.zone.run(() => {
    observer.next(term);
  });
};

this.speechRecognition.onerror = error => {
  observer.error(error);
};

this.speechRecognition.onend = () => {
  observer.complete();
};

this.speechRecognition.start();
console.log("Say something - We are listening !!!");
};

```

Fonte: Do Autor

2.1.2 Conversão da Voz em Texto

Para conversão de texto em voz, a Web Speech API disponibiliza o recurso de *SpeechRecognition* como pode ser observado na figura 14, que possui uma série de recursos para gerenciar os mecanismos de saída de texto na tela após a fala.

Para que uma instância do objeto *SpeechRecognition* possa reconhecer uma fala, é necessário que este já tenha concluído sua inicialização. Dessa forma, uma classe que utiliza SR deve implementar o método *start()* da interface *SpeechRecognition* para ser notificado quando a inicialização estiver completa.

Alguns atributos também são muito importantes quanto ao gerenciamento do SR, como o *lang*, *continuous*, *maxAlternatives* no qual o primeiro define a linguagem a ser utilizada, o segundo define se o tipo de reconhecimento é contínuo ou não, e o terceiro define o número máximo de alternativas, e o atributo *onend* que é para a leitura em execução, após que o reconhecimento de fala é iniciado, há muitos manipuladores de eventos que podem ser usados para recuperar resultados e outras informações adjacentes. O mais comum que você provavelmente usará é *SpeechRecognition.onresult*, que é acionado quando um resultado bem-sucedido é recebido.

2.2 DESENVOLVIMENTO DO PROTÓTIPO

Com base na aplicação desenvolvida em uma primeira etapa sem a implementação da tecnologia de reconhecimento, iniciou-se o desenvolvimento do protótipo. Baseado na experiência obtida no uso normal do protótipo, foram elencados os passos para a criação do protótipo, iniciando uma nova etapa, onde foram aplicadas as implementações utilizadas de

reconhecimento de fala, agora aplicando conceitos de usabilidade e modo de interação com o usuário.

Alguns dos principais itens elencados foram: movimentação do cursor entre os espaços vazios, botões acessíveis, e disponibilizar saídas na tela por meio de texto.

2.2.1 Movimentação do Cursos Entre os Espaços Vazios

Para que o processo de aprendizado de escrita possa ter um bom resultado são usados vários métodos ou métricas de ensino, trouxemos alguns métodos aqui dentro do nosso protótipo usando o reconhecimento de fala, um dos métodos escolhidos é o preenchimento de palavras inacabadas. Neste método as palavras vêm com espaços vazios entre as letras e o usuário completa as palavras com as falas.

Neste caso, foi desenvolvido o método `palavraChange()` que permitisse que o cursor se deslocasse entre os espaços vazios da palavra enquanto o usuário digita ou fala a letra que acha a correta como podemos ver na figura 15.

Figura 15 - Método para Alteração da Posição do Cursor

```
palavraChange(palavra: String, letra: String) {
    for (let index = 0; index < palavra.length; index++) {
        if (palavra[index] == " "){
            var texto = palavra[index];
            palavra = palavra.replace(texto, letra[0]);
            return palavra;
        }
    }
    return palavra;
}
```

Fonte: Do Autor

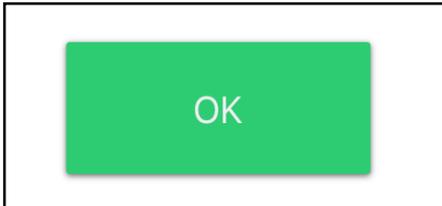
2.2.2 Acessibilidade nos Botões

A acessibilidade é um dos elementos chaves do nosso trabalho, dessa forma, visando apoiar os usuários deficientes ou seja, usuários com restrição motora nos membros superiores na execução das atividades do protótipo, buscou-se a utilização de um recurso disponível na maioria das aplicações que são a utilização de botões acessíveis, permitindo assim o usuário que possui dificuldade em utilizar os teclados e mouses convencionais a usarem o mesmo.

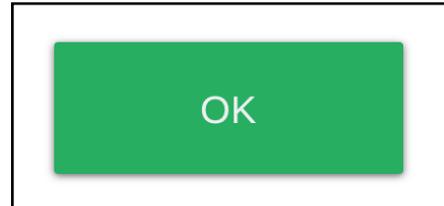
Como recurso visual, foi desenvolvido com uma borda *radius* de 2px para os contornos, podendo esta estar na core verde clara, quando desfocado, e verde mais escuro, quando em foco. Tal características foi implementado no arquivo de estilo do componente em que o mesmo se encontra, o qual é responsável por mudar as cores de acordo com a movimentação do mouse por cima do botão, como ilustrado nas figuras 16 e 17.

Figura 16 - Botão sem o mouse por cima
mouse por cima

Figura 17 - Botão com o



Fonte 3: Do Autor



Fonte 4: Do Autor

3. TESTES

Após a conclusão do desenvolvimento do protótipo da aplicação web, iniciaram-se os testes com intuito de avaliar o mesmo. Para isso, foi elencado um método distinto de validação de dado de forma objetiva.

Para uma avaliação objetiva, o protótipo foi submetido a testes, em diferentes cenários de utilização, a fim de medir a eficiência.

3.1. TESTE DE EFICIÊNCIA

Durante o desenvolvimento do protótipo, em testes informais, foram percebidos fatores importantes que influenciavam diretamente no bom desempenho do reconhecimento de fala. Dentre os mais relevantes estão os ruídos. Sendo assim, durante a elaboração dos cenários utilizados para os testes, foi adicionado o fator ruídos em diferentes condições, buscando medir sua influência sobre o processo de reconhecimento.

Os testes de eficiência se basearam em dois diferentes cenários, sendo que em cada um deles havia ruídos em ambiente diferentes. No cenário A, o ruído é reduzido, ou seja, dentro de um recinto com as portas fechadas, podendo assim capturar apenas a fala do usuário. No cenário B, o ruído é mais presente, pois nesse recinto era um ambiente aberto, possuía varias pessoas falando ao mesmo tempo, e nesse exato momento o usuário tenta interagir com o protótipo.

Na execução dos testes, foram realizadas dez leituras para cada cenário e cada leitura teve seu tempo cronometrado e seus resultados foram colocadas na tela e na console do navegador. Cada fala reconhecida foi comparado a fala normal da mesma, para a apuração do percentual de erros ocasionados no processo.

Com os resultados coletados da execução dos testes, elaborou-se a tabela 2, de forma a tabular as informações e facilitar a avaliação dos resultados.

Tabela 3 - Resultado dos testes de acordo a cada cenário

Teste	Cenário A		Cenário B	
	% Erro	Tempo Pro.	% Erro	Tempo Pro
Teste 1	25,00%	00:00: 10,65	10,83%	00:00:4,25
Teste 2	19,64%	00:00: 10,53	09,42%	00:00:3,95
Teste 3	19,19%	00:00: 8,36	08,09%	00:00:3,88
Teste 4	19,73%	00:00: 6,98	08,73%	00:00:3,83
Teste 5	16,67%	00:00: 6,82	07,87%	00:00:3,75
Teste 6	13,34%	00:00: 6,36	07,34%	00:00:3,00
Teste 7	13,06%	00:00: 5,73	07,17%	00:00:2,90
Teste 8	12,89%	00:00: 5,62	05,99%	00:00:2,75
Teste 9	12,66%	00:00: 5,31	04,75%	00:00:2,71
Teste 10	11,77%	00:00: 4,62	04,70%	00:00:2,00

Media	31,84%	00:00:7,09	07,57%	00:00:3,30
-------	--------	------------	--------	------------

Fonte: Do Autor

Com base nos resultados obtidos, é possível perceber a direta influência do ruído para com a eficiência do reconhecimento da fala. No cenário A, avaliando cada cenário, é possível perceber que os erros de reconhecimento são predominantes devido ruídos, ocasionando assim uma média de 31,84% de erros contabilizados nesse cenário.

No cenário B os resultados já são bem mais satisfatórios, apurando-se uma média de apenas 07,57% de erros. Porém o tempo de processamento da fala foi menor em relação ao outro cenário, com tempo médio de duração de 3,30 segundos.

4. RESULTADOS

Com a pesquisa realizada neste trabalho e a aplicação dos fundamentos obtidos através das pesquisas bibliográficas, foi possível obter um protótipo de aplicação web, capaz de realizar a conversão da voz em texto orientado ao aprendizado de escrita, utilizando a Web Speech API.

Através da realização de testes de eficiência, foi possível constatar que o a API de reconhecimento de voz usada possui uma ótima precisão e funciona perfeitamente em conjunto com a plataforma Web. Porém foi possível detectar uma sensibilidade à ruídos, o que é comum a qualquer aplicação com reconhecimento de voz. Sendo assim, a aplicação de técnicas de pré e pré-processamento do reconhecimento das falas, conforme pesquisas descritas no sub-capítulo 3.2.1, pode ser uma alternativa para amenizar tais problemas gerados no processo do reconhecimento da fala.

Sendo assim, o protótipo apresentou-se ideal para uma possível utilização em ambientes de aprendizagem de escrita, como escolas, instituições, entre outros, onde um usuário efetua uma fala simples e então protótipo disponibiliza a informação em formato de texto para o usuário.

Os recursos de acessibilidade pesquisados foram muito importantes para que isso fosse possível. Funções como posicionar o cursor nos campos vazios das palavras, isso facilitaria os usuários com dificuldades nos membros superiores.

O protótipo se mostrou uma solução de simples aplicação, fornecendo acessibilidade com baixo custo. É importante ressaltar a facilidade de acesso a tecnologia, por se tratar de uma aplicação web, e um microfone embutido ou externo. Sendo assim, quase todos os objetivos propostos foram atingidos, resultando em um protótipo de aplicativo, com diversas possibilidades de aprimoramento e em uma linha de pesquisa que permite a aplicação de diversos recursos tecnológicos diferentes e ainda assim muito carente de novos recursos.

5. CONCLUSÃO

A execução do presente trabalho de conclusão de curso proporcionou adquirir conhecimentos em diversas áreas. Foi possível perceber a evolução dos navegadores web e o potencial das aplicações web atualmente, tal como uma breve disposição dos recursos por ela suportada. A acessibilidade foi o principal conceito que direcionou essa pesquisa, levando a aplicação das diversas outras técnicas abordadas neste trabalho.

As técnicas e tecnologias abordadas nesse trabalho, foram aplicadas buscando atender um objetivo principal: criar um protótipo de aplicação web para que usuários possam obter um aprendizado de escrita ou das letras, de forma acessível, visando auxiliar pessoas com

deficiência nos membros superiores ou pessoas com restrições na coordenação motora nos mesmos.

Os objetivos específicos elencados para a execução desse trabalho foram todos atingidos, os quais foram responsáveis por orientar a pesquisa e o desenvolvimento, além de permitir que o objetivo geral do trabalho também fosse alcançado.

Com o estudo das metodologias de acessibilidade para pessoas com deficiências nos membros superiores, para ambiente *desktop*, adquiriu-se embasamento teórico imprescindível para a determinação dos métodos de interação com o usuário aplicados ao protótipo. Uma principal tecnologia esteve diretamente ligada ao objetivo geral, sendo ela a Web Speech API (WSA).

Como foi abordado no capítulo 3.3 do presente trabalho, a WSA possui duas funcionalidades principais, sendo elas a conversão da voz em texto e a conversão do texto em voz, que foi usada neste protótipo foi a conversão da voz em texto, sendo que existem outras ferramentas e tecnologias que fazem a mesma tarefa, a WSA mostrou-se uma tecnologia propícia para o desenvolvimento em aplicações web, visto que ela é grátis e possui uma ótima documentação e uma boa comunidade. Além disso, em testes de reconhecimento, a ferramenta apresentou resultados muito satisfatórios, com boa precisão em diferentes cenários.

Durante o levantamento bibliográfico, buscou-se compreender o funcionamento da classe *SpeechRecognition* um recurso da API para conversão de voz em texto. A mesma se mostrou muito robusta e eficiente, pois ela no seu melhor cenário apresentou uma média de 3,30% segundos no tempo de resposta, disponibilizando uma série de recursos que tornam possível trazer acessibilidade ao desenvolvimento de diversas aplicações.

As pesquisas realizadas sobre as formas de interação com os usuários do protótipo permitiram que técnicas fossem aplicadas a este, prevendo possíveis cenários e situações decorrentes de sua utilização, com base nas metodologias de acessibilidade para pessoas com deficiência nos membros superiores. Foram também fundamentais os trabalhos correlatos nessa linha de pesquisa, agregando muito conhecimento ao trabalho desenvolvido.

Desenvolveu-se um protótipo de aplicativo utilizando a plataforma Web, o qual fez a integração com a API, permitindo assim o reconhecimento de fala e a conversão de voz em texto, com base nos conceitos de acessibilidade e suas metodologias, sendo assim capaz de realizar a captura da fala para apoio a pessoas com deficiências nos membros superiores ou coordenação motora reduzidas.

Foi avaliada a eficiência do protótipo por nós, através de testes em dois cenários, que mostraram um bom desempenho no processo de reconhecimento de fala, com tempo médio de execução de 3,30 segundos e média de 7,57% de erro no melhor cenário. Os testes ainda demonstraram a sensibilidade que o processo demonstra com relação a fatores provenientes da etapa de reconhecimento da fala, como ruídos, os quais podem ser amenizados com aplicação de técnicas de pré-processamento do reconhecimento de voz, segundo teorias e praticas pesquisadas durante o levantamento bibliográfico.

Durante o processo de desenvolvimento do protótipo, algumas dificuldades foram encontradas, porém compensadas com o embasamento teórico adquirido no levantamento bibliográfico realizado. As objeções no decorrer do desenvolvimento da plataforma web, foi compensada por uma boa documentação disponível referente o uso da tecnologia. No instante da captura da fala houve dificuldades em algumas questões, como pegar a fala do usuário, mas que conseguiram ser superadas no próprio protótipo, aplicando recursos e conceitos disponíveis na documentação e em artigos encontrados na internet.

Como sugestão para trabalhos futuros, pode-se aplicar técnicas de funcionamento em dispositivos moveis, com intuito de tornar disponíveis para todos possíveis usuários. Pode-se ainda atribuir diversas funcionalidades ao aplicativo, usando novas soluções de forma criativa. Alguns exemplos de funcionalidades passíveis de implementação: O retorno da fala reconhecida por meio de libras, não só por meio de texto, criar opção de exportar todas atividades que ele realiza no protótipo por meio de fala em um arquivo de texto, aplicação de técnicas de pré-processamento do reconhecimento de voz para amenizar os ruídos no processo de reconhecimento da fala, aplicar novos métodos de ensino de escrita por meio de voz, como aprendizado de sílabas por exemplo.

6. REFERÊNCIAS

Brasil. (2009), “Subsecretaria Nacional de Promoção dos Direitos da Pessoa com Deficiência. Comitê de Ajudas Técnicas. Tecnologia Assistiva”, – Brasília: CORDE, 138 p.

Mozilla Developer. 2018b, **Web Speech API**. Disponível em: <https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API>. Acesso em: 29 maio 2018.