



Instituto Federal da Paraíba
Programa de Pós-Graduação em Engenharia Elétrica

Náiron de Vasconcelos Sousa

Automação Residencial por Comandos de Voz para Pessoas com Mobilidade Reduzida

João Pessoa/PB

Março/2018

Dados Internacionais de Catalogação na Publicação – CIP
Biblioteca Nilo Peçanha – IFPB, *Campus* João Pessoa

S725a Sousa, Náiron de Vasconcelos.
Automação Residencial por Comandos de Voz para Pessoas com Mobilidade Reduzida / Náiron de Vasconcelos Sousa. – 2018.
101 f. : il.

Dissertação (Mestrado em Engenharia Elétrica) – Instituto Federal de Educação, Ciência e Tecnologia da Paraíba – IFPB / Coordenação de Pós-Graduação em Engenharia Elétrica, 2018.
Orientador: Prof^a Suzete Élide Nóbrega Correia
Coorientador: Prof^a Silvana Luciene do N. Cunha Costa.

1. Automação residencial. 2. Reconhecimento de fala. 3. Tecnologia assistiva. 4. Sistema embarcado. 5. Automação inclusiva. 6. Sistema biométrico. I. Título.

CDU 681.5:643/645

Ivanise Andrade M. de Almeida
Bibliotecária-Documentalista
CRB-15/0096

Instituto Federal da Paraíba
Programa de Pós-Graduação em Engenharia Elétrica

Náiron de Vasconcelos Sousa

Automação Residencial por Comandos de Voz para Pessoas com Mobilidade Reduzida

Dissertação de mestrado submetida á coordenação do Programa de Pós-Graduação em Engenharia Elétrica do Instituto Federal da Paraíba como requisito necessário à obtenção do grau de Mestre em Ciências no domínio da Engenharia Elétrica.

Orientadora: Prof.^a Dr^a Suzete Élide Nóbrega Correia
Coorientadora: Prof.^a Dr^a Silvana Luciene do Nascimento Cunha Costa

João Pessoa/PB

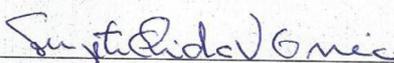
Março/2018

NÁIRON DE VASCONCELOS SOUSA

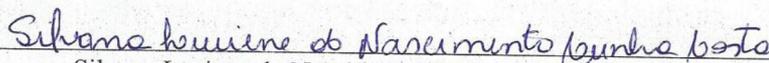
**Automação Residencial por Comandos de Voz para Pessoas
com Mobilidade Reduzida**

Dissertação aprovada pela Banca Examinadora em: 05/03/2018.

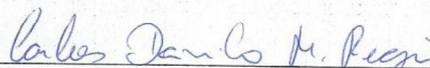
BANCA EXAMINADORA



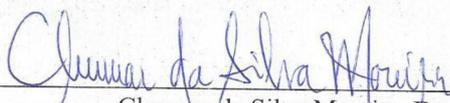
Suzete Élide Nobrega Correia, Dra. – IFPB
Orientadora



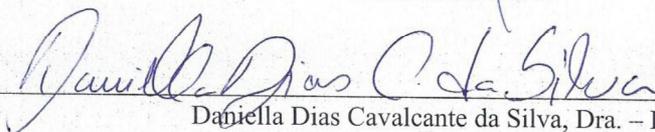
Silvana Luciene do Nascimento Cunha Costa, Dra. – IFPB
Coorientadora



Carlos Danilo Miranda Regis, Dr. – IFPB
Membro da Banca



Cleumar da Silva Moreira, Dr. – IFPB
Membro da Banca



Daniella Dias Cavalcante da Silva, Dra. – IFPB
Membro da Banca

João Pessoa – PB, Março de 2018

©Náiron de Vasconcelos Sousa – naironsousa@yahoo.com.br

Agradecimentos

- ❖ Primeiramente a Deus, por todas as coisas boas que tem acontecido em minha vida, pelas pessoas boas que tem colocado em meu caminho e por me fortalecer nas horas em que mais precisei.
- ❖ Aos meus pais José de Vasconcelos Sousa e Delmira de Jesus Santos Sousa que sempre me apoiaram e incentivaram na minha carreira acadêmica, à minha irmã e ao meu sobrinho.
- ❖ À minha orientadora, Professora Dr^a Suzete Élide Nóbrega Correia, que sempre se mostrou disposta a me ajudar, pelo apoio e por acreditar em mim, que com toda calma e profissionalismo, me guiou da melhor forma nessa jornada.
- ❖ À minha coorientadora, Professora Dr^a Silvana Luciene do Nascimento Cunha Costa pelos puxões de orelha bem merecidos, pelo apoio recebido, paciência, conselhos, que me fizeram refletir e pensar de outra forma.
- ❖ A todos aos meus amigos em especial ao casal maravilha Pablo Henrique Ursulino de Pinho e Maria de Fátima Kallynna Bezerra Couras, a Antônio Xavier, Laís Félix, Francisco Weverton, Bruno Gouveia, Lucas Gomes pelo apoio, confiança e ajuda que foram importantes nessa etapa acadêmica. E a Marcus Brasizza, que apesar de não conhecê-lo pessoalmente sempre se mostrou disposto a ajudar e sanar dúvidas que foram essenciais para os resultados obtidos neste trabalho.
- ❖ Aos professores Carlos Danilo, Ilton Barbacena, Paulo Henrique por todo conhecimento adquirido em suas respectivas disciplinas, os quais foram cruciais para o desenvolvimento desse projeto.
- ❖ À banca examinadora na pessoa do Prof. Cleumar Moreira, Prof. Carlos Danilo e Prof^ª. Daniella Dias, pelas sugestões que foram de suma importância para o enriquecimento desse trabalho de pesquisa.

“O cientista não se destina a um resultado imediato. Ele não espera que suas ideias avançadas sejam prontamente aceitas. Seu dever é estabelecer as bases para aqueles que estão por vir e apontar o caminho.”

(Nikola Tesla)

Resumo

Nos últimos anos, tem se observado um crescimento de tecnologias voltadas à Automação Industrial ou Residencial. O comando de voz para automação residencial é uma área de pesquisa emergente, devido ao surgimento de diversas plataformas embarcadas, simples e de baixo custo, com alto poder de processamento e tamanho reduzido. A utilização de sistemas de automação controlados por voz é indiscutivelmente interessante às pessoas portadoras de necessidades especiais, idosos, nas quais a fala pode substituir algumas ações motoras. Pode ser realizada de forma independente ou dependente do locutor. Nesse sentido, este projeto aborda o desenvolvimento de um sistema de reconhecimento automático de fala, de baixo custo, independente do locutor, que utiliza comandos isolados. O trabalho se divide em duas abordagens. Na primeira, foi empregada a plataforma *Raspberry PI B+*. O sistema foi testado com o *Pocketsphinx* na língua inglesa com uma placa de circuito própria desenvolvida para o acionamento das cargas, com a finalidade de avaliar eficiência e eficácia do *software*. Foi obtida uma taxa de reconhecimento de 90,4%, com resposta rápida do sistema. Na segunda abordagem, o sistema de reconhecimento é baseado na língua portuguesa, utilizando a plataforma embarcada *Raspberry PI Zero W*, ESP8266 12E e o *software Jasper*, tendo a placa de controle dos dispositivos separada do *Raspberry Pi Zero W*. A fala é convertida em texto usando o *Google speech*. Após validado o comando, o mesmo é enviado ao ESP8266 12E através de um roteador que permite ações de controle, como aparelhos eletroeletrônicos ou envio de mensagens SMS, com o intuito de dar autonomia, segurança e comodidade às pessoas com mobilidade reduzida. O sistema proposto foi avaliado através de aspectos de integridade, tais como sensibilidade a ruídos do meio, sensibilidade de aceite da pronúncia das palavras de acordo com o idioma escolhido e detecção de falso-positivo. Nos testes realizados com a segunda abordagem obteve-se uma taxa de acertos de 99,67 %, indicando que a proposta se mostra viável para a criação de uma interface por reconhecimento automático de fala com baixo custo.

Palavras chaves: Automação residencial, Sistema embarcado, Reconhecimento de Fala, Tecnologia Assistiva.

Abstract

In recent years, there has been a growth of technologies aimed at Industrial Automation or Residential. The voice command for home automation is an emerging research area, due to the emergence of several embedded platforms, simple and low cost, with high processing power and small size. The use of voice-controlled automation systems is undoubtedly interesting for people with special needs, the elderly, where speech can replace some motor actions. It can be performed independently or dependent on the speaker. In this sense, this project addresses the development of an automatic speech recognition system, low cost, independent of the speaker, which uses isolated commands. The work is divided into two approaches. In the first one, the Raspberry PI B + platform was used. The system was tested with the Pocketsphinx in the English language with its own circuit board developed for the activation of loads, in order to evaluate efficiency and effectiveness of the software. A recognition rate of 90.4% was obtained with rapid system response. In the second approach, the recognition system is based on the Portuguese language, using the embedded platform Raspberry PI Zero W, ESP8266 12E and the Jasper software, having the device control board separated from Raspberry Pi Zero W. The speech is converted to text using Google speech. Once the command is validated, it is sent to the ESP8266 12E through a router that allows control actions, such as electronic devices or sending SMS messages, in order to provide autonomy, safety and comfort for people with reduced mobility. The proposed system was evaluated through integrity aspects, such as sensitivity to medium noises, sensitivity of the pronunciation of the words according to the chosen language and detection of false positive. In the tests carried out with the second approach, a success rate of 99.67% was obtained, indicating that the proposal proves feasible for the creation of an interface by automatic speech recognition with low cost.

Keywords: Residential automation, Embedded system, Speech recognition, Assistive Technology.

Lista de Figuras

Figura 1 - Tecnologias assistivas.....	26
Figura 2 - Aparelho fonador.....	32
Figura 3 - Pregas vocais.....	33
Figura 4 - <i>Hardware</i> RAF.....	34
Figura 5 - <i>Software</i> RAF.....	34
Figura 6 - Módulo V2.....	37
Figura 7 - <i>Shield EasyVR3</i>	37
Figura 8 - Alto-falante <i>Echo</i>	38
Figura 9 - <i>Raspberry Pi</i> , (a) vista de cima, (b) vista de baixo.....	40
Figura 10 - Pinagem GPIO's <i>Raspberry Pi B+</i> e <i>Zero W</i>	41
Figura 11 - <i>Raspberry Pi Zero W</i>	42
Figura 12 - ESP8266 12E.....	43
Figura 13 - Pinagem ESP8266 12E.....	44
Figura 14 - Modelo Geral da Metodologia Aplicada com o <i>Pocketsphinx</i>	48
Figura 15 - Placa de circuito.....	49
Figura 16 - Estrutura para acoplamento com o <i>Raspberry Pi</i>	50
Figura 17 - Adaptador de Áudio.....	50
Figura 18 - Microfone plugue P2 com filtro “POP”.....	51
Figura 19 - Placa de controle dos dispositivos centralizada.....	51
Figura 20 - Fluxograma do algoritmo desenvolvido em <i>Python</i>	53
Figura 21 - Placa de controle dos dispositivos descentralizada.....	55
Figura 22 - Modelo Geral da Metodologia Aplicada com o <i>Google Speech API</i>	56
Figura 23 - Metodologia detalhada da segunda abordagem.....	57
Figura 25 - Mensagem enviada através do <i>Twilio API</i>	59
Figura 24 - Fluxograma geral segunda abordagem.....	60
Figura 26 - Esquema eletrônico de conexão entre <i>Buzzer</i> e o <i>Raspberry Pi</i>	61
Figura 27 - Microfone utilizado na segunda abordagem.....	62
Figura 28 - Esquema eletrônico acionamento de cargas.....	62
Figura 29 - (a) Caixa patota onde o circuito foi montado e fixado na parede (b) Montagem eletrônica.....	63
Figura 30 - Montagem para testes da segunda abordagem.....	67

Figura 31 - Taxa de acertos individual e média geral.	69
Figura 32 - Exemplo corte de palavra pelo sistema.	70
Figura 33 - Janela principal <i>SDFormatter</i>	80
Figura 34 - Configurações <i>SDFormatter</i>	81
Figura 35 - Janela principal <i>Disk Imager</i>	81
Figura 36 - Janela de configurações do <i>Raspbian</i>	82
Figura 37 - Janela de confirmação.....	83
Figura 38 - Interface network <i>Raspberry Pi</i>	84
Figura 39 - Painel de controle ALSA	86

Lista de Tabelas

Tabela 1 - Trabalhos empregando a fala para automatização.....	22
Tabela 2 - Recursos de tecnologias assistivas.	26
Tabela 3 - Especificações da placa <i>Raspberry Pi B+</i>	39
Tabela 4 - Especificações da placa <i>Raspberry Pi Zero W</i>	42
Tabela 5 - Comandos definidos para os módulos e suas respectivas ações.....	52
Tabela 6 - Custo para a implementação da primeira abordagem.....	54
Tabela 7 - Indicativo de comunicação entre as setas.....	58
Tabela 8 - Comandos definidos a serem reconhecidos pelo sistema.....	58
Tabela 9 - Custo para a implementação da segunda abordagem.	64
Tabela 10 - Tabela feita para anotação dos resultados	65
Tabela 11 - Resultados taxas de acertos e média geral com <i>Pocketsphinx</i>	66
Tabela 12 - Número de acertos para cada locutor.	68

Lista de abreviaturas e siglas

IBGE	Instituto Brasileiro de Geografia e Estatística.
VoIP	<i>Voice over Internet Protocol</i>
GSM	<i>Global System for Mobile Communications</i>
AT	<i>Attention</i>
LaPS	Laboratório de Processamento de Sinais
XML	<i>Extensible Markup Language</i>
API	<i>Application Programming Interface</i>
IBM	<i>International Business Machines</i>
PC	<i>Personal Computer</i>
ARM	<i>Acorn RISC Machine</i>
USB	<i>Universal Serial Bus</i>
LCD	<i>Liquid Crystal Display</i>
LED	<i>Light Emitting Diode</i>
HDMI	<i>High-Definition Multimedia Interface</i>
I/O	<i>Input/Output</i>
GPIO	<i>General Purpose Input/Output</i>
LAN	<i>Local Area Network</i>
SD	<i>Secure Digital</i>
ALSA	<i>Advanced Linux Sound Architecture</i>
STT	<i>Speech To Text</i>
TTS	<i>Text To Speech</i>
CMU	<i>Carnegie Mellon University</i>
IHM	Interação Homem Máquina
SMS	<i>Short Message Service</i>
IDE	<i>Integrated Development Environment</i>
TTL	<i>Transistor-Transistor – Logic</i>
OLED	<i>Organic Light-Emitting Diode</i>
PCB	<i>Printed Circuit Board</i>
RAM	<i>Random Access Memory</i>
IoT	<i>Internet of Things</i>
SPI	<i>Serial Peripheral Interface</i>

I2C	<i>Inter-Integrated Circuit</i>
TCP/IP	<i>Transmission Control Protocol/Internet Protocol</i>
UART	<i>Universal Asynchronous Receiver/Transmitter</i>
ADC	<i>Conversor Analógico-Digital</i>
HTTP	<i>HyperText Transfer Protocol</i>
SMS	<i>Short Message Service</i>
SSH	<i>Secure Shell</i>
ROM	<i>Read Only Memory</i>

Sumário

1. INTRODUÇÃO	16
1.1. Formulação do problema	16
1.2. Justificativa	17
1.3. Estado da Arte	18
1.4. Objetivos do Trabalho	23
1.4.1. Objetivo Geral	24
1.4.2. Objetivos Específicos	24
1.5. Organização do trabalho	24
2. FUNDAMENTAÇÃO TEÓRICA	25
2.1. Tecnologias assistivas	25
2.3. Automação inclusiva	29
2.4. Sistemas biométricos	30
2.5. Reconhecimento de fala	31
2.6. Dispositivos eletrônicos e <i>Software</i> de assistência social voltados para o reconhecimento de fala	36
2.7. <i>Raspberry Pi B+</i>, <i>Raspberry Pi Zero W</i> e <i>ESP8266 12E</i>	38
2.7.1. <i>Raspberry Pi B+</i>	38
2.7.2. <i>Raspberry Pi Zero W</i>	41
2.7.3. <i>ESP 8266 12E</i>	43
2.8. <i>Jasper</i>	45
2.9. Sistema Operacional <i>Raspbian Wheezy</i>	46
3. MATERIAIS E MÉTODOS	47
3.1. Primeira abordagem – Reconhecimento na língua inglesa.	47
3.1.1. <i>Software</i>	47
3.1.2. <i>Hardware</i>	48

3.2. Segunda abordagem – Reconhecimento na língua portuguesa	54
3.2.1. <i>Software</i>	55
3.2.2. <i>Hardware</i>	61
4. RESULTADOS	65
4.1. Resultados para a primeira abordagem	65
4.1.1. Discussão.....	66
4.2. Resultados para a segunda abordagem	67
4.2.1 - Discussão	69
5. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS.....	71
5.1. Contribuições do trabalho.....	71
5.2. Sugestões para trabalhos futuros	72
6. REFERENCIAIS BIBLIOGRÁFICAS.....	73

APÊNDICES

APÊNDICE A – GUIA DE GRAVAÇÃO DO SISTEMA OPERACIONAL NO CARTÃO SD E CONFIGURAÇÕES DO <i>RASPBERRY PI</i>.	80
A.1 – Formatação (<i>SD Formatter</i>)	80
A.2 - Gravação (<i>Win32DiskImager</i>)	81
A.3 – Configurações iniciais <i>Raspbian Wheezy</i>.....	82
A.3.1 – Expandindo o sistema de arquivos	82
APÊNDICE B – GUIA DE INSTALAÇÃO DO <i>SOFTWARE JASPER</i>.	85
B.1 – Atualização do sistema e configurações <i>ALSA</i>.....	85
B.2 – Instalação <i>Jasper</i> e <i>STT</i>.....	86

ANEXOS

ANEXO A – TERMO DE COMPROMISSO LIVRE E ESTABELECIDO (TCLE).	93
ANEXO B – PARECER DO COMITÊ DE ÉTICA APROVANDO O RELATÓRIO FINAL DE PESQUISA.	96

1. INTRODUÇÃO

1.1. Formulação do problema

Dados do IBGE no Brasil mostram que cerca de 23,9% da população, o que equivale a mais de 45,6 milhões de pessoas, possuem algum tipo de deficiência, seja física ou psicológica, entre elas visuais (18,60%), auditiva (5,10%), motora (7,00%) e mental ou intelectual (1,40%) (VIEIRA, 2012).

Sasaki (2003) define uma pessoa com deficiência como sendo aquela que possui um comprometimento físico, sensorial ou mental, o qual lhe traz limitação, incapacidade ou desvantagem em relação à categoria das pessoas sem nenhum tipo de deficiência. Conforme o artigo 1º do Decreto nº 186/2008 da Convenção sobre os Direitos das Pessoas com Deficiência, pode-se dizer que “pessoas com deficiência são aquelas que têm impedimentos de natureza física, intelectual ou sensorial, os quais, em interação com diversas barreiras, podem obstruir sua participação plena e efetiva na sociedade com as demais pessoas”.

Já o conceito de pessoa com mobilidade reduzida é aquela pessoa que “não é portadora de deficiência, mas tem dificuldade de movimentar-se, de flexibilidade, de coordenação motora e de percepção” (BRASIL, 2004, p.14). Sendo assim, neste conceito podem ser incluídas as pessoas idosas. Dados do Instituto Brasileiro de Geografia e Estatísticas (IBGE, 2015), mostraram que a expectativa de vida ao nascer no Brasil era de 75 anos, 5 meses e 26 dias em 2015. Em 2016, houve um aumento de 3 meses e 14 dias em relação ao ano anterior. Um estudo de Sales (2016), que faz uma análise das condições de vida da população brasileira, mostrou que entre 2005 e 2015, a proporção de idosos de 60 anos ou mais, na população do país, passou de 9,8% para 14,3%.

Pessoas com necessidades especiais ou pessoas idosas muitas vezes encontram dificuldades em exercer atividades corriqueiras simples, dentro de casa ou na rua, tornando-as, assim, dependentes de terceiros para ajudá-las. A ausência de adaptações apropriadas, em residências, prejudica a acessibilidade e o desempenho de pessoas com deficiência. Com o notável crescimento do número de idosos, o desenvolvimento de sistemas com o intuito de auxiliá-los em atividades básicas do dia a dia, em sua residência, que os traga autonomia, comodidade e segurança, é de suma importância.

1.2. Justificativa

Entende-se como automação qualquer sistema baseado em computador, utilizado em indústrias, em serviços ou para o bem-estar das pessoas que, substituindo o trabalho humano, reduz os custos e aumenta a qualidade e a segurança dos processos e dos produtos (MORAES, 2007).

Dispositivos inteligentes são tema central na construção de ambientes e residências autônomas. Qualquer equipamento que tenha autonomia para desenvolver uma tarefa básica, efetuar troca de informações com outros dispositivos, possibilitar comando remoto e ter capacidade de tomada de decisão pode ser considerado um dispositivo inteligente (BOLZANI, 2004).

O comando de voz para automação residencial é uma área de pesquisa emergente, devido ao surgimento de diversas plataformas embarcadas, simples e de baixo custo, com alto poder de processamento e tamanho reduzido. Estas plataformas podem ser facilmente adaptadas e incorporadas aos sistemas já existentes, podendo ser empregadas como ferramentas de auxílio ao controle do ambiente doméstico (domótica) (PIMENTEL, 2014).

A utilização de sistemas de automação controlados por voz é indiscutivelmente interessante às pessoas portadoras de necessidades especiais e idosos, nas quais a fala pode substituir algumas ações motoras. A voz carrega informações únicas de cada indivíduo, podendo ser empregada para realizar o controle de dispositivos, a exemplos de controle de cadeira de rodas (BORGES, 2015), o controle de funcionalidades em um automóvel (ZHANG, 2015), eletrodomésticos (OLIVEIRA, 2015), controle de acesso a ambientes (segurança) (LIMA, 2015), entre outros. Pode, ainda, ser incorporada a tarefa de reconhecimento do locutor, para personalizar o controle das funções desejadas. Nesse caso, os sistemas ou dispositivos seriam controlados somente por indivíduos autorizados, os quais utilizariam a voz como senha de acesso.

Diversos trabalhos sobre automação por comandos de voz, voltados às tecnologias assistivas, têm sido desenvolvidos ao longo dos últimos anos, visando dar autonomia às pessoas com mobilidade reduzida, seja por uma deficiência física ou idade avançada. A ideia é o desenvolvimento de um dispositivo que seja de fácil aquisição, fácil utilização e *hardware* de baixo custo.

Projetos de automação com o intuito de auxiliar pessoas com algum tipo de deficiência utilizando sistemas embarcados e voz já foram testados, utilizando placas principalmente como Arduino, *Raspberry Pi* e *BeagleBone Black*, com *softwares* de código aberto,

apresentando taxas de reconhecimento acima de 80% (OLIVEIRA e OTTOBONI, 2015). As duas últimas plataformas apresentam um melhor custo benefício que o Arduino, pois além de possuírem processadores que comportam os *softwares* de reconhecimento e mesmo tamanho físico, utilizam um sistema operacional para aplicação e boa quantidade de memória RAM, tornando-as mais eficientes.

Nesse contexto, este projeto aborda o desenvolvimento de um sistema de reconhecimento automático de fala, de baixo custo, independente do locutor, que utiliza comandos isolados. O trabalho se divide em duas abordagens. Na primeira, foi empregada a plataforma *Raspberry PI B+*. O sistema foi testado com o *Pocketsphinx* na língua inglesa com uma placa de circuito própria desenvolvida para o acionamento das cargas, com a finalidade de avaliar eficiência e eficácia do *software*. Na segunda abordagem, o sistema de reconhecimento é baseado na língua portuguesa, utilizando a plataforma embarcada *Raspberry PI Zero W*, ESP8266 12E e o *software Jasper*, tendo a placa de controle dos dispositivos separada do *Raspberry Pi Zero W*. A conversão fala-texto é realizada pelo *Google speech*. Após a validação do comando, o mesmo é enviado ao ESP8266 12E através de um roteador que permite ações de controle, como aparelhos eletroeletrônicos ou envio de mensagens SMS, o que possibilita mais autonomia, segurança e comodidade às pessoas com mobilidade reduzida.

1.3. Estado da Arte

Chiele e Zerbetto (2010) desenvolveram um controlador de cadeira de rodas que emprega o *software* Via Voice, da empresa IBM para o reconhecimento de comandos de voz, associada a uma interface feita em *Visual Basic*. O dispositivo após receber o comando e validá-lo, envia-o ao microcontrolador PIC16F877A, responsável por se comunicar com a placa de circuito da cadeira de rodas. Os resultados obtidos foram satisfatórios para o reconhecimento de palavras em português, com índices de acertos em torno de 95,2%. Os autores concluíram que é possível e viável a construção de uma cadeira de rodas elétrica por comando de voz de baixo custo, oferecendo ao usuário um conforto extra e liberdade em movimentos simples do cotidiano.

Perico et al. (2012) tiveram como objetivo controlar uma plataforma elevatória para cadeiras de rodas, disponibilizada pelo PROTA (Programa de Tecnologia Assistiva), implementando um sistema de reconhecimento de fala utilizando a plataforma *Raspberry PI*. O comando de voz era transmitido para o *Raspberry Pi* através do programa de

reconhecimento de fala, que utilizava o decodificador *Julius*. O modelo acústico utilizado foi o LapSAM e o dicionário fonético o UFPAdic 3.0, ambos do grupo de pesquisa Fala Brasil. Os estudos foram executados com a participação de cinco voluntários. Obteve-se um grau de confiança acima de 70% com um bom tempo de resposta. O sistema era totalmente independente de conexão com a Internet ou qualquer outra fonte de dados. Concluíram dessa maneira a viabilidade de execução de um trabalho voltado para pessoas com necessidades especiais, utilizando-se processamento via PC, comandado por voz e de baixo custo.

Santana (2014) desenvolveu um trabalho que visava o controle de um robô Lego NXT por reconhecimento de fala contínua. Os comandos eram enviados através de conexão *Bluetooth*, onde os modelos acústicos utilizados foram os do Laboratório de Processamento de Sinais (LaPS) da Universidade Federal do Pará (UFPA). Uma API também foi desenvolvida pelo LaPS, que encapsula o *Julius*, para todas as linguagens previstas para a plataforma .NET, onde a ideia era controlar o *Julius* em tempo real, e como modelo de linguagem o XML. Para a implementação do reconhecimento de fala, foi utilizado como base o código fornecido pelo LaPS Coruja 2.0, programa que reconhece palavras isoladas e pode ser facilmente adaptado para fala contínua. Na análise de dados obteve um grau de confiança de 80%, observando que a captura da fala com o mínimo de ruído é muito importante para resultados mais satisfatórios, chegando à conclusão de que reconhecimento por voz de uma linguagem, é feita de forma simples.

Barros e Ianhas (2014) propuseram um dispositivo de automação residencial para o controle de equipamentos domésticos através de comandos por voz. O dispositivo era interativo e tinha como principal objetivo trazer comodidade e segurança aos usuários. Para o protótipo, foi utilizada a plataforma Arduino UNO e o módulo de voz *EasyVr Shield*, onde foram gravados os bancos de vozes. Adicionalmente, foram utilizados módulos relés e LED's de infravermelhos para o controle das cargas, como ventiladores, lâmpadas e aparelhos de TV. Após os testes, concluíram que os resultados foram satisfatórios, atendendo todas as perspectivas, tanto em desempenho quanto em segurança.

Pimentel (2014) desenvolveu uma plataforma de reconhecimento automático de fala para o auxílio de pessoas com mobilidade reduzida, com tecnologias eficientes, portáteis e de baixo custo para a interação humano-computador, através de comandos na língua inglesa. O projeto integrava componentes de *hardware*, *software* e de código livre, sendo os comandos de voz fornecidos ao sistema através de um *smartphone* com *softphone* VoIP (Voz sobre IP) instalado, que se registrava em um servidor de comunicação *Asterisk*. Integrada ao servidor estava a ferramenta de reconhecimento da fala, *Julius*. Todos esses componentes foram

embarcados na plataforma Beaglebone Black, utilizando o sistema operacional Ubuntu Linux, com um ambiente de desenvolvimento *Python*, onde o treinamento dos comandos para validação foi realizado anteriormente no próprio módulo. Foram obtidas respostas adequadas aos comandos, mesmo pronunciadas por indivíduos diferentes, com uma taxa média global de acertos de 95,9%, mesmo em ambientes ruidosos.

Obaid et al. (2014) desenvolveram um protótipo de automação residencial para o auxílio de pessoas com mobilidade reduzida ou idosas, utilizando a plataforma Arduino UNO, módulo Wi-Fi *ZigBee* e o *software Lab View*. O conversor STT foi utilizado o *Microsoft Speech (SAPI)*, que após receber o comando de voz através de uma interface no PC, envia ao *ZigBee* conectado ao Arduino UNO um conjunto de números binários através da rede *Wireless* estabelecida. Os testes foram feitos com três eletrodomésticos que eram comutados por relés. Os resultados foram satisfatórios para os autores, com a maior vantagem de que o sistema só é treinado uma única vez. Os autores chegaram à conclusão de que é possível o controle de aparelhos domésticos utilizando o *ZigBee* de forma fácil por interação da voz através de uma interface simples de usar como o *Lab View*.

Gobinathan e Narayanadass (2014) desenvolveram um sistema de automação residencial por reconhecimento de fala para pessoas idosas ou com deficiência, de baixo custo utilizando-se a plataforma embarcada *Raspberry Pi modelo B*, dois módulos GSM/GPRS SIM300 e o *Sphinx-4 speech recognition*. O comando após ser reconhecido através de uma interface feita em Java para fazer a interação entre o microfone e o decodificador *Sphinx-4*, era enviada em forma de texto via SMS do *Raspberry Pi* com um dos módulos SIM300 ao segundo módulo SIM300 conectado a uma placa microcontroladora para o acionamento dos dispositivos. O sistema apresentou uma taxa de acerto de 93,55% fazendo com que os autores chegassem a conclusão que o sistema pode contribuir para capacitar os idosos e deficientes, atendendo suas necessidades.

Kronbauer e Costa (2015) desenvolveram um dispositivo para permitir a interação de um usuário com os AmIs (Ambientes de Inteligência), por comando de voz, onde seria possível a integração de sensores, atuadores e sistema de inferência. A plataforma escolhida foi o *Raspberry Pi* com o sistema operacional *Raspbian*. O *software* para servir de base para as interpretações locais dos comandos foi o *Jasper*. Para a tradução de voz para texto foi escolhido o *AT&T Speech API*. Na avaliação inicial, na visão dos usuários, a infraestrutura foi capaz de realizar as tarefas propostas na maioria dos casos. Assim, foi possível concluir que a abordagem pode se tornar uma solução viável para a criação de interfaces de interação via voz.

Suresh e Rao (2015), desenvolveram uma implementação automática para a segurança de residências empregando o *Raspberry Pi*, conectado a um *Hub* via *Wi-Fi*. Para a captura e reconhecimento da voz foi utilizado um módulo *Voice Recognition V3* que é uma placa de fácil controle. Nos testes obtiveram uma taxa de 93,55% de acertos, chegando à conclusão que é possível criar um dispositivo de segurança de baixo custo, via voz, empregando o *Raspberry Pi*.

Lingaria (2015) desenvolveu uma tecnologia assistiva residencial por comando de voz utilizando a plataforma Arduino, para criar uma interface entre o módulo *Voice Recognition V3* e o módulo GSM SIM900. O módulo GSM SIM900 se comunicava com as operadoras locais de celular, onde o usuário poderia realizar uma chamada de voz, inclusive para serviços de emergência. O *Shield* GSM permite que a placa Arduino receba chamadas de voz, para se comunicar com a internet e para enviar e receber mensagens SMS, onde também é possível se comunicar via comandos AT. Concluíram que os resultados foram efetivos, obtiveram em torno de 80% de acertos em ambiente silencioso e, assim, atendendo a proposta de um sistema para chamadas de emergência através de *Hardware*s efetivamente baratos.

Kumar e Shimi (2015) desenvolveu uma proposta de automação residencial de baixo custo para pessoas com paralisia, com a intenção de auxiliá-las via comandos de voz. Na proposta ele utilizou um Arduino UNO, o módulo de reconhecimento de voz V3. Além de controle de lâmpadas, um *Buzzer*, caso a pessoa precise avisar o seu cuidador e também o sistema foi implementado em uma cama de madeira ajustável. O módulo de reconhecimento V3 após o reconhecimento do comando realizava a ação pré-definida seja na iluminação ou ajuste do ângulo da cama. Com isso foi possível se utilizar de dispositivos embarcados e do módulo de reconhecimento de voz que tem uma boa taxa de aceitação de 99% em ambientes silenciosos e uma média de 60% em ambientes ruidosos, para o auxílio de pessoas com paralisia ou idosas.

Ramlin et al (2016), desenvolveram uma proposta de automação residencial por comando de voz se utilizando da plataforma embarcada Arduino UNO R3, módulo *Bluetooth* HC-06 e o *smartphone* para o envio dos comandos através do APP, *AMR_Voice*. Quando pressionado o botão virtual do APP o sistema capturava o comando e o convertia para texto. Depois de convertido era enviado ao Arduino UNO via *Bluetooth*, então o sistema combinava o texto enviado com os comandos definidos no sistema realizando a ação determinada. Para os testes foram usados um conjunto de LEDs para simular uma lâmpada e uma ventoinha para simular um ventilador, comutados por um módulo de relés. Os testes apresentaram acertos de 100% para todos os comandos, seis no total, exceto para a palavra “*fan*” que às vezes era

traduzida como “*fen*” ou “*fend*”. Chegaram à conclusão que é possível o desenvolvimento de automação por reconhecimento de fala de baixo custo, voltado para pessoas com deficiência ou que estão acamadas.

Anusha et al. (2016) desenvolveram um dispositivo para automação residencial por voz voltado para pessoas com mobilidade reduzida, utilizando o *Raspberry Pi B+*, dois módulos *ZigBee*, a plataforma ARMLPC2103 e o API da Google para a conversão de voz para texto através de uma interface desenvolvida com o *QT Creator*. A voz, após capturada, é convertida para texto e salva com arquivo “*rss.text*” através do servidor, então é comparada com as palavras predefinidas para reconhecimento do sistema. O *Raspberry Pi* envia a plataforma ARMLPC2103 através do *ZigBee* a *string* referente a cada GPIO da plataforma ARMLPC2103 quer por fim ativa as cargas. Os resultados foram satisfatórios obtendo o sistema uma alta taxa de acerto, o que levou aos autores a conclusão que é possível se utilizar do *Raspberry Pi* e o módulo *ZigBee* para automação residencial por *Wireless* de baixo custo.

Na Tabela 1 é apresentado um resumo do estado da arte com os trabalhos mais relevantes para a pesquisa encontrados na literatura.

Tabela 1- Trabalhos empregando a fala para automação.

(continua)

Autor/ Ano/ Proposta	Metodologia	Resultados e Conclusões
Chiele e Zerbetto (2010) Controlador para cadeira de rodas.	<ul style="list-style-type: none"> • Emprega <i>software</i> Via <i>Voicer</i>; • Interface em <i>Visual Basic</i>; • PIC16F877 para o controle da placa de circuito. 	<ul style="list-style-type: none"> • Índices de acertos de 95,2%. • Conclusão de ser possível e viável a construção de uma cadeira de rodas controlada por voz de baixo custo.
Perico et al. (2012) Controlar uma plataforma elevatória para cadeiras de rodas.	<ul style="list-style-type: none"> • Utilizou o <i>Raspberry Pi</i>; • Programa de reconhecimento de fala o decodificador <i>Julius</i>; • Modelo acústico utilizado foi o LapSAM. 	<ul style="list-style-type: none"> • Obteve-se um grau de confiança acima de 70% com um bom tempo de resposta. Concluíram ser viável o desenvolvimento de automação via voz utilizando-se <i>hardware</i> de baixo custo.
Santana (2014) Desenvolveu um trabalho que visava o controle de um robô Lego NXT.	<ul style="list-style-type: none"> • Comandos enviados do PC ao robô via <i>Bluetooth</i>; • Modelos acústicos do LaPS; <i>Julius</i> como conversor. 	<ul style="list-style-type: none"> • Obteve um grau de confiança de 80%; • Concluiu que o reconhecimento por voz de uma linguagem é feita de forma simples.

(continuação)

Autor/ Ano/ Proposta	Metodologia	Resultados e Conclusões
<p>Barros e Ianhas (2014)</p> <p>Controle de equipamentos domésticos através de comandos por voz.</p>	<ul style="list-style-type: none"> • Usou o Arduino UNO; • Módulo de voz <i>EasyVr Shield</i>; • Foram utilizados módulos relés e LED's de infravermelhos para o controle das cargas. 	<ul style="list-style-type: none"> • Concluíram que os resultados foram satisfatórios, atendendo todas as perspectivas, tanto em desempenho quanto em segurança.
<p>Pimentel (2014)</p> <p>Plataforma de reconhecimento automático de fala para o auxílio de pessoas com mobilidade reduzida.</p>	<ul style="list-style-type: none"> • Comandos de voz fornecidos ao sistema através de um <i>smartphone</i> com <i>softphone</i> VoIP; • Utilizou <i>Beaglebone Black</i>; • Ferramenta de reconhecimento da fala, <i>Julius</i>. 	<ul style="list-style-type: none"> • Taxa média global de acertos de 95,9%. • Concluiu que existe confiabilidade necessária para o emprego da plataforma em diversas aplicações.
<p>Obaid et al. (2014)</p> <p>Desenvolveram um protótipo de automação residencial para o auxílio de pessoas com mobilidade reduzida ou idosas.</p>	<ul style="list-style-type: none"> • Utilizaram a plataforma Arduino UNO, módulo Wi-Fi <i>ZigBee</i> e o <i>software Lab VieW</i>; • O conversor STT utilizado foi o <i>Microsoft Speech</i>. 	<ul style="list-style-type: none"> • Os resultados foram satisfatórios, com a maior vantagem de que o sistema só é treinado uma única vez; • Chegaram à conclusão de que é possível o controle de aparelhos domésticos, utilizando <i>hardware</i> de baixo custo.
<p>Gobinathan e Narayanadass (2014)</p> <p>Desenvolveram um sistema de automação residencial por reconhecimento de fala para pessoas idosas ou com deficiência.</p>	<ul style="list-style-type: none"> • Utilizaram a plataforma embarcada <i>Raspberry Pi</i>; • Módulos GSM/GPRS SIM300; • A ferramenta <i>Sphinx-4 speech recognition</i> para conversão da fala para texto. 	<ul style="list-style-type: none"> • Apresentou uma taxa de acerto de 93,55%; • Concluíram que o sistema pode contribuir para capacitar os idosos e deficientes.
<p>Kronbauer e Costa (2015)</p> <p>Desenvolveram um dispositivo para permitir a interação de um usuário com os AmIs (Ambientes de Inteligência), por comando de voz.</p>	<ul style="list-style-type: none"> • Utilizaram a plataforma <i>Raspberry Pi</i>; • <i>Software</i> de base <i>Jasper</i>; • STT <i>AT&T Speech API</i>. 	<ul style="list-style-type: none"> • A infraestrutura foi capaz de realizar as tarefas propostas na maioria dos casos. • Concluíram que a abordagem é viável para a criação de interfaces via voz.
<p>Anusha et al. (2016)</p> <p>Desenvolveram um dispositivo para automação residencial por voz voltado para pessoas com mobilidade reduzida.</p>	<ul style="list-style-type: none"> • Utilizaram o <i>Raspberry Pi B+</i>; • Módulos <i>ZigBee</i>, a plataforma <i>ARMLPC2103</i>; • API da Google para a conversão de voz para texto. 	<ul style="list-style-type: none"> • Os resultados foram satisfatórios, o sistema obteve uma alta taxa de acerto. • Concluíram ser possível uma automação residencial com <i>hardware</i> de baixo custo.

Fonte: Autoria própria.

1.4. Objetivos do Trabalho

1.4.1. Objetivo Geral

Desenvolver um dispositivo para automação residencial por comando de fala com *hardware* de baixo custo empregando o *Raspberry Pi*, voltado para pessoas com algum tipo de necessidade especial com o intuito de auxiliá-las em atividades básicas do dia a dia, em sua residência como o controle de aparelhos eletroeletrônicos.

1.4.2. Objetivos Específicos

- a) Construir um protótipo de automação residencial com *hardware* de baixo custo, para o controle de dispositivos domésticos, controlado por voz;
- b) Estudar o desempenho e a praticidade do sistema para o reconhecimento automático da fala;
- c) Avaliar o protótipo desenvolvido e o módulo de acionamento dos dispositivos e transmissão Wi-Fi, com a intenção de validação do sistema.

1.5. Organização do trabalho

O trabalho está organizado da seguinte forma: no Capítulo 1 está a introdução, que aborda o motivo da pesquisa, o estado da arte e os objetivos do trabalho; no Capítulo 2 é apresentada a fundamentação teórica abordando o aparelho fonológico, tecnologias assistivas, domótica, automação inclusiva, reconhecimento de fala, dispositivos eletrônicos e *softwares* voltados para o reconhecimento de fala, sobre a plataforma embarcada *Raspberry Pi B+ e Zero W*, ESP8266 12E, o *software Jasper* e o sistema operacional *Raspbian Wheezy*; no Capítulo 3 são apresentados os materiais e métodos empregados em duas abordagens para a proposta. No Capítulo 4 são apresentados os resultados obtidos e discussões para a primeira e segunda abordagem após os testes de validação do sistema, seguido do Capítulo 5 que apresenta as considerações finais acerca dos resultados obtidos, contribuições e sugestões para trabalhos futuros; e, por fim, no Capítulo 6, as referências bibliográficas utilizadas como fonte de pesquisa.

2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados alguns conceitos básicos necessários ao entendimento do sistema implementado. Para tanto, são abordados aspectos básicos do aparelho fonador e das tecnologias assistivas no âmbito da domótica e da automação inclusiva. São abordados, ainda, tópicos como sistemas biométricos, com foco nos sistemas que utilizam a fala como biometria e os dispositivos e *softwares* empregados na operacionalização do sistema implementado.

2.1. Tecnologias assistivas

A ausência de adaptações apropriadas, em residências, prejudica a acessibilidade e o desempenho de pessoas com deficiências. Estas pessoas, em seu cotidiano, se deparam com diversas dificuldades de locomoção e acesso, o que limita ou mesmo inviabiliza sua independência e autonomia. Assim, é necessário que estas pessoas utilizem materiais, equipamentos adaptados, adequação do mobiliário e estrutura arquitetônica, ou seja, recursos que lhes propiciem condições seguras de mobilidade e conforto.

O termo Tecnologia Assistiva é utilizado para identificar todo este arsenal de recursos e serviços que contribuem para propiciar ou ampliar habilidades funcionais de pessoas com deficiência e, conseqüentemente, promover vida independente e inclusão (SARTORETO e BERSCH, 2014). Exemplos de tecnologias assistivas podem ser vistos na Figura 1. Observa-se que a tecnologia assistiva está presente desde um instrumento adaptado, como em um aparato simples para facilitar a escrita (imagem esquerda), com artes curvadas e circulares que se adaptam à mão da pessoa com deficiência. Para acessibilidade em computadores pode ser um teclado adaptado, leitores de tela ou *softwares* especiais (de reconhecimento de voz, etc.). Equipamentos de entrada e saída (síntese de voz, Braille) de dados, ou auxílios alternativos de acesso como a ponteira de cabeça (imagem direita), que permitem as pessoas que não possuem o movimento das mãos usarem o computador através de movimentos da cabeça (SILVEIRA, 2016).

Os recursos de tecnologias assistivas são organizados ou classificados de acordo com objetivos funcionais a que se destinam tais como mobilidade, adequação postural, comunicação, recursos para cegos ou pessoas de baixa visão.

Figura 1 - Tecnologias assistivas.

Fonte: Psicologiaacessivel.net.

Para surdos ou pessoas com perdas auditivas, instrumentos que promovam independência em atividades da vida diária, recursos para educação, recreação, acessibilidade arquitetônica, adaptações de veículos, recursos para acesso ao computador, órteses, próteses e outros (BERSCH et al., 2007). Essas tecnologias podem ser aplicadas de diversas formas e para determinados tipos de deficiências, como apresentado na Tabela 2.

Tabela 2 - Recursos de tecnologias assistivas.

(continua)

Categorias	Características
Auxílios para vida diária	Materiais e produtos para auxílio em tarefas rotineiras tais como comer, cozinhar, vestir-se, tomar banho e executar necessidades pessoais, manutenção da casa, etc.
Recursos de comunicação	Recursos, eletrônicos ou não, que permitem a comunicação expressiva e receptiva das pessoas sem a fala ou com limitações da mesma.
Recursos de acessibilidade ao computador	Equipamentos de entrada e saída (síntese de voz, Braille), auxílios alternativos de acesso (ponteiras de cabeça, de luz), teclados modificados ou alternativos, acionadores, <i>softwares</i> especiais (de reconhecimento de voz, etc.), que permitem as pessoas com deficiência a usarem o computador.
Sistemas de controle de ambiente	Sistemas eletrônicos que permitem às pessoas com limitações moto-locomotoras, controlar remotamente aparelhos eletroeletrônicos, sistemas de segurança, entre outros, localizados em seu quarto, sala, escritório, casa e arredores.

(continuação)

Categorias	Características
Auxílios de mobilidade	Qualquer equipamento que ajude a pessoa com restrição física ou mobilidade reduzida a se locomover com mais facilidade, tais como: cadeira de rodas motorizada, cadeira de rodas elevador, ou escada elevador.
Órteses e próteses	Troca ou ajuste de partes do corpo, faltantes ou de funcionamento comprometido, por membros artificiais ou outros recursos ortopédicos (talas, apoios etc.). Incluem-se os protéticos para auxiliar nos déficits ou limitações cognitivas, como os gravadores de fita magnética ou digital que funcionam como lembretes instantâneos.
Adequação postural	Adaptações para cadeira de rodas ou outro sistema de sentar visando o conforto e distribuição adequada da pressão na superfície da pele (almofadas especiais, assentos e encostos anatômicos), bem como posicionadores e contentores que propiciam maior estabilidade e postura adequada do corpo através do suporte e posicionamento de tronco/cabeça/membros.
Melhorias sensoriais	Qualquer recurso que possa tornar mais fácil à interação com mundo ao seu redor daqueles que estão parcialmente ou totalmente cegos ou surdos. Por exemplo, um decodificador <i>telecaption</i> para um aparelho de TV seria um dispositivo auxiliar para uma pessoa que possui deficiência auditiva.
Recursos terapêuticos	Equipamentos ou processos que ajudem a recuperar o máximo possível de uma doença ou lesão. Terapia pode envolver uma combinação de serviços e tecnologia, como um fisioterapeuta ter que usar uma unidade de massagem especial para restaurar um mais amplo movimento para os músculos rígidos.
Projetos arquitetônicos para acessibilidade	Adaptações estruturais e reformas na casa e/ou ambiente de trabalho, através de rampas, elevadores, adaptações em banheiros entre outras, que retiram ou reduzem as barreiras físicas, facilitando a locomoção da pessoa com deficiência.

2.2. Domótica

Esse termo domótica é a junção da palavra latina “Domus” (casa) com “Robótica”, que vem sendo utilizada para trazer segurança ou mesmo conforto. A domótica pode ter foco em duas arquiteturas, na Arquitetura Baseada em Automação (ABA), onde o usuário se adapta à automação, ou na Automação Baseada em Comportamento (ABC), onde o sistema é que se adapta e aprende com o usuário (LINS e MOURA, 2010).

De acordo com Roque (2002), domótica é a utilização simultânea da eletricidade, da eletrônica e das tecnologias da informação no ambiente residencial, permitindo realizar a sua gestão, local ou remota, e oferecer uma vasta gama de aplicações integradas nas áreas da segurança, conforto, comunicação e gestão de energia.

Para a implementação de um sistema domótico é executada uma série de funções que podem ou não estar integrado a outros dispositivos envolvidos. A execução dessas funções determina a classificação, segundo Muratori (2004) em:

- Autônomos: Esse tipo possui funcionalidades somente de ligar ou desligar dispositivos ou subsistemas, onde nenhum tem relação com o outro (TERUEL, 2008, p.28);
- Integrados: Nos integrados existem múltiplos subsistemas integrados a um único controlador. Esse tipo de sistema não permite ser personalizado, o funcionamento dos subsistemas é unicamente na forma a qual seu fabricante pretendia. Basicamente trata-se apenas de controle remoto estendido a diferentes ambientes (TEZA, 2002, p.32). O processamento pode ser centralizado na central de automação ou distribuído pela rede (TERUEL, 2008, p.28);
- Complexos: possuem como grande diferencial a possibilidade da personalização dos produtos manufaturados de modo que atenda às necessidades do usuário. É dependente de comunicação de mão dupla e realimentação de *status* entre todos os subsistemas (TEZA, 2002, p.33).

As formas de controle desses sistemas podem ser centralizadas, onde a base controladora envia as informações para os controladores e interfaces. O recebimento de informações se dá através de sensores, sistemas interconectados e ação do usuário (CASADOMO, 2008).

Sua vantagem é que são sistemas de baixo custo de implementação comparado a descentralizada. A desvantagem fica por conta do grande número de cabeamento, da dificuldade da interação homem-máquina de modo que não corresponde com a filosofia dos sistemas domóticos, muitos sistemas desse tipo são executados com CLPs (Controlador Lógico Programável), outros não possuem barramento de comunicação e inexistência de comunicação entre equipamentos (DÍAZ; PARDO; PULIDO, 2001).

Na placa de controle dos dispositivos centralizada, um barramento interconecta vários controladores que possibilitam a troca de informações entre eles. Já os atuadores, as interfaces e os sensores não necessariamente comunicam com mais de um controlador diretamente, ou seja, a proposta é dividir o sistema para suprir necessidades complexas (CASADOMO, 2008).

Neste modelo de arquitetura a principal ideia é tornar o sistema mais robusto a falhas, fácil desenho de instalação e fácil uso, cumprindo com os requisitos que um sistema domótico deve ter. Porém, o custo de integração é mais elevado em função das tecnologias empregadas (DÍAZ; PARDO; PULIDO, 2001).

2.3. Automação inclusiva

A automação residencial por reconhecimento de fala se torna uma das mais adequadas a pessoas com algum tipo de mobilidade reduzida, seja por uma deficiência ou idade avançada. Ela possibilita o usuário ter o controle dos dispositivos eletroeletrônicos como lâmpadas, ventiladores, portas, TVs, entre outros, se utilizando apenas da fala.

Muratori e Bó (2011) definem automação residencial como um conjunto de serviços proporcionados por sistemas tecnológicos integrados como o melhor meio de satisfazer as necessidades básicas de segurança, comunicação, gestão energética e conforto de uma habitação. Porém, neste contexto preferem o termo “domótica”, que é largamente empregado na Europa, por ser mais abrangente. A Asociación Española de Domótica (Cedom), tem uma definição mais completa sobre o termo:

Domótica é a automatização e o controle aplicados à residência. Esta automatização e controle se realizam mediante o uso de equipamentos que dispõem de capacidade para se comunicar interativamente entre eles e com capacidade de seguir as instruções de um programa previamente estabelecido pelo usuário da residência e com possibilidades de alterações conforme seus interesses. Em consequência, a domótica permite maior qualidade de vida, reduz o trabalho doméstico, aumenta o bem-estar e a segurança, racionaliza o consumo de energia e, além disso, sua evolução permite oferecer continuamente novas aplicações (MURATORI e BÓ, 2011, p. 70).

A domótica tomou várias vertentes nos últimos anos, uma delas dando bastante ênfase à automação inclusiva, voltada para pessoas com deficiência ou mobilidade reduzida, que busca soluções que garantam a acessibilidade, segurança, conforto e saúde de seu usuário (GUEDES et al., 2012).

Segundo Rockenbach (2005), a automação inclusiva possibilita o acionamento remoto ou automático de sistemas automatizados, proporcionando conveniência ao indivíduo, de forma que este possa ter maior autonomia e independência em suas atividades diárias. Tais ajudas tecnológicas propiciam inúmeras formas de adaptação e controle, sejam recursos instalados à cadeira de rodas do indivíduo, oferecendo a este controle total dos itens ligados a este sistema, ou na própria residência.

Segundo afirma o secretário Nacional para a Reabilitação e Integração das Pessoas com Deficiência (SNRIPD) de Portugal:

Entende-se por ajudas técnicas qualquer produto, instrumento, estratégia, serviço e prática utilizada por pessoas com deficiência e pessoas idosas, especialmente produzido ou geralmente disponível para prevenir, compensar, aliviar ou neutralizar uma deficiência, incapacidade ou desvantagem e melhorar a autonomia e a qualidade de vida dos indivíduos (PORTUGAL, 2007).

Prado, Lopes e Ornstein (2010) definem acessibilidade remetendo ao princípio do direito universal de ir e vir, como sendo a capacidade de o ambiente construído oferecer segurança e autonomia a qualquer pessoa que o utilize, independentemente de suas limitações.

2.4. Sistemas biométricos

Existem atualmente dispositivos que se utilizam de método automático de reconhecimento individual baseado em medidas biológicas, principalmente para o controle de segurança. O primeiro passo utilizado por esses processos são a captura em si do material a ser analisado, em seguida o sistema faz a extração onde o material capturado é convertido em informações identificáveis pelo sistema. Após traduzido para linguagem de máquina, o sistema cria um padrão único para esse cadastro que é salvo em um banco de dados. Por fim, acontece a comparação com cadastros já salvos no sistema (GREGORIN, 2017).

Dentre essas medidas biométricas estão como as mais utilizadas:

- O reconhecimento da impressão digital: o sistema captura a imagem da impressão digital com um leitor biométrico óptico e compara com um banco de dados com imagens pré-gravadas. A aplicação é bastante confiável e de baixo custo;
- Reconhecimento de face: Apresenta menor confiabilidade por ser sensível às mudanças de expressão, luz ambiente e distância entre a face e a câmera, porém de fácil implementação e baixo custo;
- Reconhecimento pela íris: Bastante confiável por ser imutável ao passar dos anos, apresenta um alto custo de implementação;
- Reconhecimento pela retina: Também bastante confiável, porém de difícil leitura e alto custo de implementação;
- Reconhecimento da assinatura: As características únicas de cada indivíduo como pressão, movimentos aéreos entre outras, tornam difícil a sua falsificação, sendo um método prático de se aplicar e de médio custo;
- Reconhecimento de locutor: De confiança média devido a ruídos no ambiente ou por mudança na voz do usuário devido a alguma patologia, apresenta uma demora no processo de cadastramento, porém tem baixo custo de implementação.
- Reconhecimento de fala: De fácil implementação e baixo custo, porém ruídos ambientes, mudanças na voz por alguma patologia ou estado emocional podem dificultar o reconhecimento.

2.5. Reconhecimento de fala

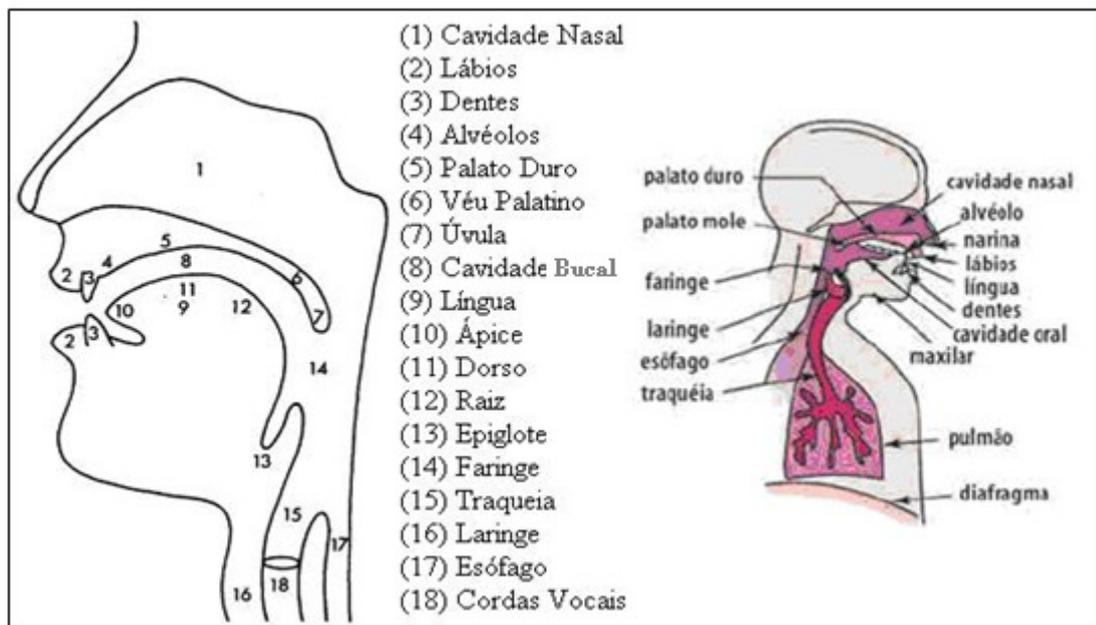
A voz é uma ferramenta de grande complexidade e possui inúmeros fatores envolvidos na sua produção. As estruturas anatômicas e os processos fisiológicos envolvidos no sistema de produção da fala são responsáveis pelos diversos fonemas emitidos e, conseqüentemente, carregam informações sobre suas formas, fisiologia e dinâmica (BEHLAU, 2001).

Segundo Behlau (2004), a voz é uma das ferramentas primárias e mais imediata que o ser humano dispõe para interagir com a sociedade.

O aparelho fonador, representado na Figura 2, é o conjunto de órgãos responsáveis pela fonação humana, ou seja, pela geração da voz. A voz é produzida na laringe, onde se localiza as cordas vocais ou pregas vocais.

A produção da voz está relacionada a fatores biológicos e genéticos, mas também culturais e psicossociais. Também tem relação com a personalidade, o estado emocional e a forma de expressar emoções (GOULART, 2002).

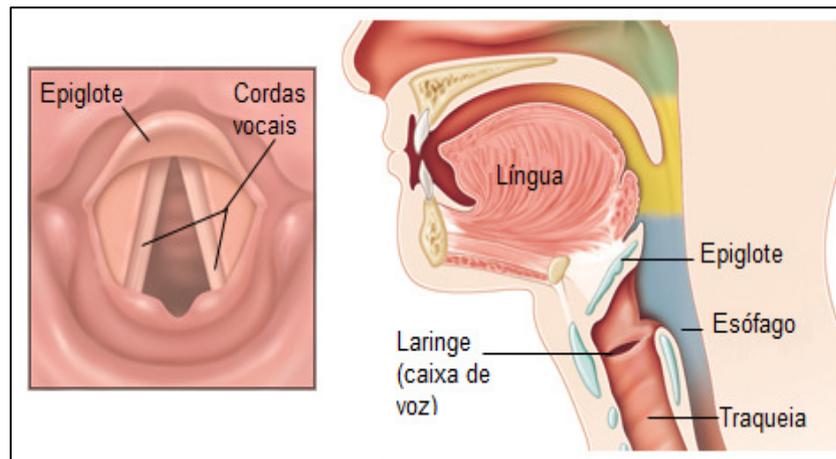
Figura 2 - Aparelho fonador.



Fonte: Fonticaarticulatria.com.

A voz é produzida na laringe, onde se encontram as pregas vocais, que no ato da fala se aproximam suavemente e realizam um movimento de vibração, por conta da passagem do ar que vem dos pulmões durante o ato de expiração, sendo produzido o som. Este som é ampliado e modificado pelas cavidades ressonantes (faringe, cavidade bucal, cavidade nasal e seios perinasais) e pelos órgãos de articulação (lábios, dentes, maxilares, língua, palato duro e mole) que cooperam na articulação de consoantes e vogais.

As pregas vocais (Figura 3) são duas faixas de tecido elástico, localizadas lado a lado na laringe, um pouco acima da traquéia. Quando se permanece em silêncio, as pregas mantêm-se abertas, criando assim uma via através da qual se respira. Quando se emite som, o ar que sai dos pulmões é forçado através das pregas vocais, fazendo com que elas vibrem. Quando vibram mais rápido, produzem sons de intensidade inferior. Ao vibrarem, o som é emitido (MONTEIRO e FERREIRA, 2011).

Figura 3 - Pregas vocais.

Fonte: hmsportugal.wordpress.com.

As pregas vocais em mulheres tendem a ser menores do que a dos homens. Sendo assim, as mulheres tendem a ter uma voz mais aguda. Também com relação ao tamanho das mesmas, quanto menor mais vibram. Normalmente as pregas vocais das mulheres vibram 220 vezes por segundo (220 Hz) enquanto as pregas vocais dos homens vibram 110 vezes por segundo (110 Hz) (GOULART e COOPER, 2002).

Cada indivíduo, assim, possui uma voz única como uma impressão digital, tornando essa característica genética, que pode ser empregada para a automatização de sistemas controlados por comandos de voz. Este recurso está sendo utilizado inclusive em controle de acesso para segurança, controle de ambientes, ou mesmo itens pessoais como uma cadeira de rodas, como na proposta de Chiele e Zerbetto (2010) e Suresh e Rao (2015).

Os sistemas de reconhecimento automático de fala (RAF), implementado para automação, passam por alguns processos necessários envolvendo dispositivos de *hardware* (Figura 4) e *software* (Figura 5), para o reconhecimento da voz e posteriormente a realização de ações pré-determinadas.

Para que um sistema computacional reconheça a fala e juntamente a fonética da palavra dita, são seguidos alguns passos. Primeiramente, o som da voz precisa passar por um processo de filtragem para eliminar a maior parte do ruído e interferências possível, proveniente do ambiente. Depois essas ondas sonoras são transformadas em dados digitais através de um conversor analógico-digital. Esses filtros e conversores podem estar já inseridos em um único chip no próprio microfone. Após, é aplicado uma medida para cada uma dessas ondas capturadas, então são computadas as características espectrais contidas na voz e sincronizadas com modelos de som já armazenadas na memória do classificador. Essa

Por fim, o resultado é analisado e comparado com palavras ou frases conhecidas. O resultado é convertido para uma funcionalidade escolhida pelo usuário, como converter para texto, para reconhecimento do locutor ou acionar dispositivos eletroeletrônicos.

O maior problema do reconhecimento de fala não é a transcrição do sinal analógico para o sinal digital, mas sim o do reconhecimento deste sinal digital em modelos conhecidos de palavra. Para lidar com reconhecimento de fala é necessário interpretar a fala através da manipulação da representação do conhecimento fonético-fonológico (VIEIRA; LIMA, 2001).

Segundo Silva (2010, p. 9), mesmo com tantos avanços na área da computação e do reconhecimento de palavras, o entendimento completo da fala humana por um computador é um processo complexo. Alguns fatores relacionados são:

- a) A mesma palavra pronunciada várias vezes pode apresentar características diferentes devido à articulação dos órgãos do aparelho fonador (SILVA, 2010, p. 9);
- b) O sotaque do locutor tem grande influência na forma como a palavra é pronunciada (LOUZADA, 2010, p. 13), o que dificulta a implementação do reconhecimento do português do Brasil que, por ser um país de tamanho continental e ter sofrido influência na colonização por vários povos, tem sotaques muito variados;
- c) Em idiomas que possuem um vocabulário muito extenso, como o português do Brasil, existem várias palavras com a mesma pronúncia e significados diferentes, como por exemplo: sessão e cessão, concerto e conserto, mas e mais, entre outras (SILVA, 2010, p. 9);
- d) A pronúncia de uma palavra pode variar de locutor para locutor, sendo que alguns podem falar mais rápido ou “engolir” sílabas e letras (LOUZADA, 2010, p. 13);
- e) A relação sinal-ruído influencia o reconhecimento de voz, pois quanto menos ruído perto da fonte de entrada de áudio, melhor será o reconhecimento, enquanto ambientes com muito ruído prejudicam o desempenho do reconhecedor de voz (SILVA, 2010, p. 9).

2.6. Dispositivos eletrônicos e *Software* de assistência social voltados para o reconhecimento de fala

Com o grande crescimento do setor de automação e IoT (*Internet of Things*), dispositivos dos mais diversos tipos e finalidades estão sendo inseridos no mercado. Dispositivos como sensores, atuadores e plataformas embarcadas com microcontroladores ou microprocessadores já desenvolvidos para a finalidade. Com o invento de tais plataformas também surgiu a necessidade de *softwares* para programar ou interfacear essas plataformas. *Softwares* que podem ser desenvolvidos especificamente para a aplicação ou adaptados à mesma.

A plataforma mais conhecida e utilizada atualmente é o Arduino UNO, com baixo custo, linguagem de fácil programação e uma grande comunidade. Tornou-se uma das plataformas de aprendizado mais utilizadas em sistemas de automação. Com isso, fabricantes passaram a desenvolver dispositivos de fácil comunicação baseado na plataforma, como *Shields* e módulos prontos para as mais diversas finalidades, como reconhecimento de cor, comunicação com rede Wi-Fi, medição de temperatura e humidade, reconhecimento digital e reconhecimento de voz.

A *ELECHOUSE* desenvolveu um módulo de reconhecimento de voz V2 (Figura 6), para plataformas embarcadas, compacto e de fácil controle dependente de locutor, capaz de gravar até 15 comandos de voz divididos em três grupos, possuindo duas formas de controle: pela porta serial ou pinos de saída. O módulo de reconhecimento V3, posterior ao V2, conta com armazenamento e reconhecimento de até 80 comandos, com precisão de até 99% em ambientes adequados.

Esses módulos podem ser facilmente adaptados às mais diversas plataformas, com tensão de trabalho entre 4,5 a 5,5 Volts e corrente de consumo máximo de 40 mA, pode ser alimentada direto da plataforma e programada através de sua IDE.

O *Shield* EasyVR3 (Figura 7) é mais robusto do que o módulo V2, podendo ser encaixado direto no Arduino UNO sem a necessidade de fios ou conectores. O *Shield* aceita 32 comandos principais definidos pelo usuário (os chamados *Speaker Dependent* - SD) e outros 26 comandos secundários (os chamados *Speaker Independent* - SI), possui conectores adicionais para entrada de microfone, uma saída para alto-falantes, saída de áudio em P2 fêmea, e acesso aos pinos de entrada e saída do módulo, além de LEDs na placa para mostrar *feedback* durante o uso (USINAINFO, 2015).

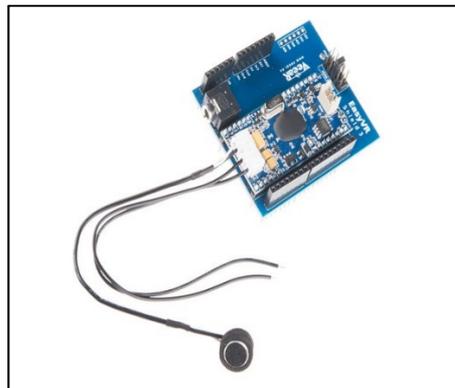
Trabalha com nível de tensão TTL de 3,3 a 5 Volts e comandos básicos no idioma inglês, italiano, japonês, alemão, espanhol, francês e português.

Figura 6 - Módulo V2.



Fonte: Smartkits.com.br.

Figura 7 - Shield EasyVR3.



Fonte: Multilogicashop.com.br.

Algumas empresas também disponibilizam serviços de reconhecimento de fala automático a partir de *softwares* que podem ser instalados em qualquer PC, *smartphone notebook* ou plataformas embarcadas que aceitem um sistema operacional, seja para obter informações de forma prática ou mesmo automação.

Algumas alternativas para aparelhos celulares disponíveis no mercado é o *Google Now* (TECMUNDO, 2016), *Cortana* (MICROSOFT, 2017), *Siri* (IPODSCHOOL, 2015) e *Tina* (DESIGNSOFT, 2016). As aplicações são disponíveis para sistemas Android, IOS, e *Windows Phone*, tanto para *smartphone* ou *Tablet*, e podem oferecer o controle do dispositivo

e acesso a serviços com notificações de redes sociais, e-mail, hora, temperatura, efetuar chamadas etc.

A *Alexa* (AWS, 2017) desenvolvida pela *Amazon* é, atualmente, o assistente doméstico portátil mais promissor. Diferente das propostas dos demais, lançou o alto-falante inteligente *Echo* (Figura 8), que dispensa o uso de aparelhos como celulares ou *Tablets* e que, além dos serviços básicos dos demais, oferece serviços de pedido em redes *delivery* associados. Por outro lado, o dispositivo não tem suporte para automação residencial.

Figura 8 - Alto-falante *Echo*.



Fonte: Uol.

2.7. *Raspberry Pi B+*, *Raspberry Pi Zero W* e ESP8266 12E

2.7.1. *Raspberry Pi B+*

O *Raspberry Pi B+* (Figura 9) é um pequeno computador embarcado de desenvolvimento, barato e educacional. Os sistemas operacionais suportados utilizam o núcleo Linux, em sua maioria gratuitos e de código aberto. Há uma imensa liberdade de personalização da máquina e de suas funções por intermédio da linguagem de programação *Python*.

O *Raspberry Pi* não possui disco rígido e tanto o sistema operacional como todos os arquivos, são armazenados em um cartão SD.

Característica da placa *Raspberry Pi B+*:

- Processador de um chip Broadcom BCM2835 de 700 MHz, 32 bits, construído sobre a arquitetura ARM11;

- Possui 512 MB de memória RAM;
- *Slot* para cartão de memória micro SD (*Secure Digital*);
- Quatro portas USBs 2.0 podendo fornecer até 500 mA de corrente, possibilitando assim a conexão de mais periféricos como *mouse*, teclado, *webcam*, adaptador *Wi-Fi*, *Bluetooth*, etc;
- Uma porta *Ethernet* padrão RJ45;
- Conector HDMI (1920 × 1200 pixels) que oferece saída de áudio e vídeo digital de alta definição;
- Um conector de interface serial para câmera, que permite que um módulo de câmera seja conectado diretamente à placa;
- Um conector de interface serial do display que pode ser usado para se comunicar com uma tela de exibição LCD ou OLED;
- Saída de vídeo compostos em um Jack de 3,5mm de quatro polos, junto com o áudio;
- LEDs de *status*;
- Quarenta pinos I/Os, para controle ou monitoramento de equipamentos, sensores ou atuadores externos;
- Conector micro USB para alimentação externa de 5 V máximo e 700 mA mínimo.

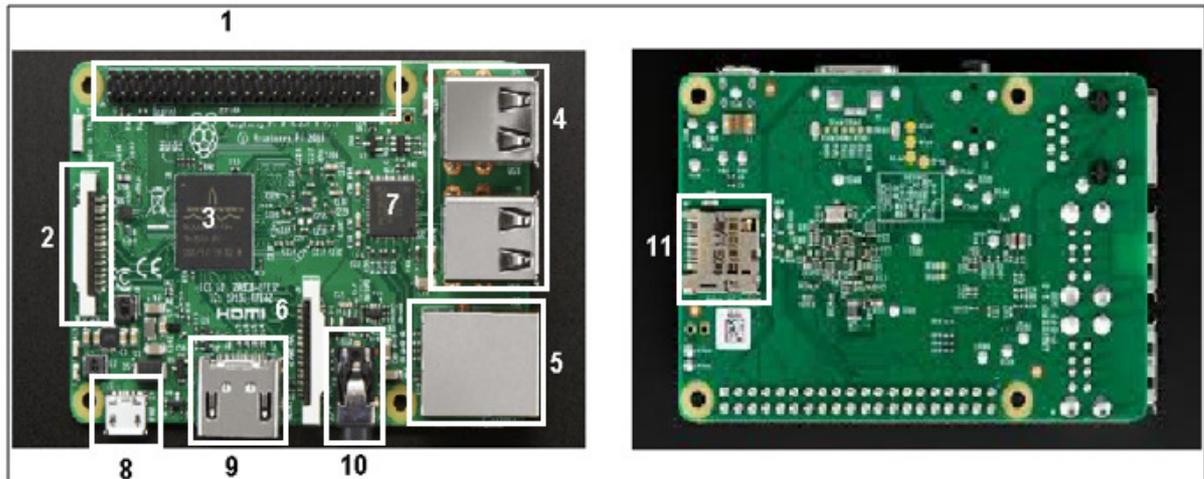
Na Tabela 3 estão enumerados e descritos cada componente da placa de acordo com a Figura 9.

Tabela 3 - Especificações da placa *Raspberry Pi B+*.

Número	Especificação
1	Pinos GPIO
2	Conector de interface serial do display
3	Processador Broadcom BCM2837, 700 MHz, 32 bits
4	4 x USB 2.0
5	Conector LAN RJ45
6	Conector para interface serial para câmera
7	Chip LAN (<i>Local Area Network</i>) 10/100 Mbps
8	Entrada Micro USB de energia
9	Saída de vídeo HDMI
10	Saída de vídeo compostos e áudio
11	<i>Slot</i> para cartão <i>SD Card</i>

Fonte: Autoria Própria.

Figura 9 - *Raspberry Pi*, (a) vista de cima, (b) vista de baixo.



Fonte: Autoria Própria.

O barramento de pinos, ilustrado na Figura 10, dispõe de:

- Dois pinos (1 e 17) de alimentação de 3.3 Volts e dois (2 e 4) de 5 Volts;
- Oito pinos (6, 9, 14, 20, 25, 30, 34 e 39) ligados ao GND (*Ground*);
- Dois pinos (3 e 5) de comunicação I2C (Circuito Inter-integrado) protocolo para fazer conexões entre periféricos de baixa velocidade, onde utiliza-se um barramento entre dois fios, sendo um de dados e outro de *clock*, para comunicação serial entre circuitos integrados montados em uma mesma placa;
- Duas portas (8 e 10) seriais RX/RT, que utilizam o protocolo RS-232 para o envio e recebimento de sinal digital;
- Dezesete pinos (7, 11, 12, 13, 15, 16, 18, 22, 29, 31, 32, 33, 35, 36, 37, 38 e 40) de GPIOs que servem de envio e recebimentos de dados digitais;
- Cinco pinos (19, 21, 23, 24 e 26) de entrada e saída digitais, também podem fazer comunicação serial *Full Duplex* síncrono, que permite o processador do *Raspberry* comunicar com algum periférico externo de forma bidirecional;

- Dois pinos (27 e 28) ID EEPROM (*Electrically-Erasable Programmable Read-Only Memory*). Este é um tipo de memória que pode ser programado e apagado várias vezes, através de uma tensão elétrica interna ou externa.

Figura 10 - Pinagem GPIO's *Raspberry Pi* B+ e Zero W.

Raspberry Pi GPIO Header A+, B+, Zero, Pi2				
Pin#	NAME		NAME	Pin#
01	3.3v DC Power		DC Power 5v	02
03	GPIO02 (SDA1 , I ² C)		DC Power 5v	04
05	GPIO03 (SCL1 , I ² C)		Ground	06
07	GPIO04 (GPIO_GCLK)		(TXD0) GPIO14	08
09	Ground		(RXD0) GPIO15	10
11	GPIO17 (GPIO_GEN0)		(GPIO_GEN1) GPIO18	12
13	GPIO27 (GPIO_GEN2)		Ground	14
15	GPIO22 (GPIO_GEN3)		(GPIO_GEN4) GPIO23	16
17	3.3v DC Power		(GPIO_GEN5) GPIO24	18
19	GPIO10 (SPI_MOSI)		Ground	20
21	GPIO09 (SPI_MISO)		(GPIO_GEN6) GPIO25	22
23	GPIO11 (SPI_CLK)		(SPI_CE0_N) GPIO08	24
25	Ground		(SPI_CE1_N) GPIO07	26
27	ID_SD (I ² C ID EEPROM)		(I ² C ID EEPROM) ID_SC	28
29	GPIO05		Ground	30
31	GPIO06		GPIO12	32
33	GPIO13		Ground	34
35	GPIO19		GPIO16	36
37	GPIO26		GPIO20	38
39	Ground		GPIO21	40

Fonte: Element14.com.

Os pinos definidos como saída apresentam uma tensão de 0 V no caso do nível lógico 0, e 3.3V para o nível lógico 1. A corrente para esses pinos não deve superar 16 mA por pino, e 50 mA na totalidade de todos os pinos.

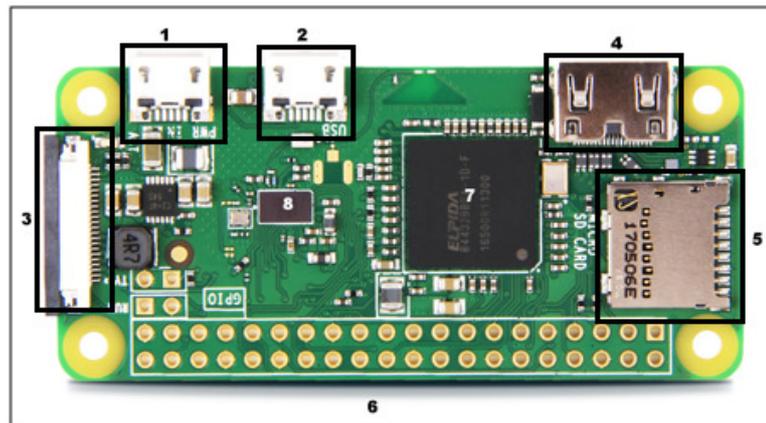
2.7.2. *Raspberry Pi* Zero W

O *Raspberry Pi* Zero W, representado na Figura 11, é uma versão reduzida das demais placas *Raspberry Pi*, porém possui uma maior capacidade de processamento e valor de custo reduzido em relação as outras placas da mesma família. O diferencial do *Raspberry Pi* Zero W são os recursos de Wi-Fi e *Bluetooth* já embarcado na placa o que torna fácil a implementação da plataforma em diversos projetos, incluindo automação residencial e IoT. O chip Cypress CYW3438 sem fio incorpora um layout de antena PCB (*Printed circuit board*) licenciado pela empresa sueca ProAnt. As especificações da placa podem ser vistas na Tabela 4.

Configurações da placa *Raspberry Pi Zero W*:

- CPU BCM2835 1 GHz;
- 512 MB de memória RAM;
- Mini HDMI e USB;
- Entrada de energia micro USB;
- 40 pinos entre GPIOs e de alimentação;
- Conector para câmera e display;
- *Wireless* LAN 802.11 b/g/n;
- *Bluetooth* 4.1 BLE (*Bluetooth Low Energy*);

Figura 11 - *Raspberry Pi Zero W*.



Fonte: Seedstudio.com.

Tabela 4 - Especificações da placa *Raspberry Pi Zero W*.

Número	Especificações
1	Entrada de força
2	Entrada Micro USB
3	Conector para câmera
4	Saída micro HDMI
5	Slot para cartão <i>SD Card</i>
6	Pinos de GPIO's
7	CPU
8	Chip LAN e <i>Bluetooth</i>

Fonte: Autoria própria.

10. A pinagem das GPIO's do *Raspberry Pi Zero W*, correspondem às mesmas da Figura

2.7.3. ESP 8266 12E

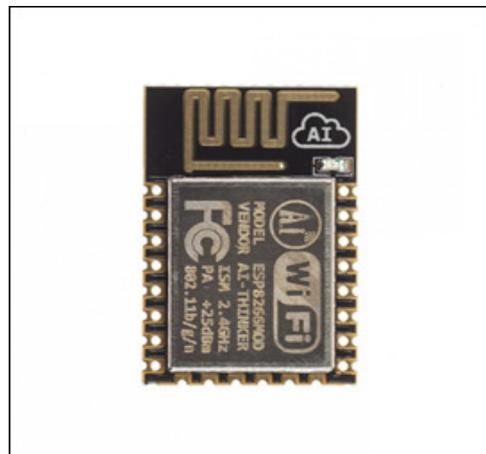
A família das plataformas embarcadas ESP8266 (Figura 12) foram desenvolvidas para práticas de IoT (*Internet of Things*), possuem um *chip* de arquitetura 32 bits com Wi-Fi integrado com a antena já embutida na placa, interface USB-Serial e trabalha com uma tensão de 3,3V.

Conecta-se ao computador por meio de um conversor USB-Serial e pode ser programado em LUA ou *Wiring* através da IDE do Arduino, conta com 11 pinos I/O (Entrada e Saída) e conversor analógico-digital.

Configurações da placa ESP8266 -12E:

- *Wireless* padrão 802.11 b/g/n;
- Modos de operação: STA/AP/STA+AP;
- Suporta 5 conexões TCP/IP;
- GPIO com funções de PWM, I2C, SPI, etc;
- Taxa de transferência: 110-460800bps;
- Suporta *Upgrade* remoto de *firmware*;
- CPU 32 bits RISC Tensilica Xtensa LX106 rodando à 80/160 MHz;
- RAM: 64 KB;
- *FLASH*: QSPI Externo – de 512 KB até 4 MB;
- WiFi: IEEE 802.11 – b/g/n.

Figura 12 - ESP8266 12E.

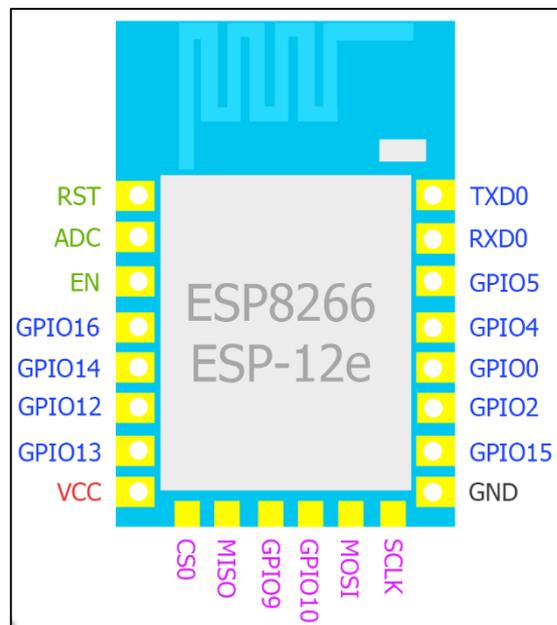


Fonte: Jaycon Sistemas.com.

O barramento de pinagem (Figura 13) conta com:

- Dois pinos para comunicação UART (TX/RX);
- Onze pinos de entrada e saída digitais (GPIOs);
- Um pino ADC (conversor analógico digital) de 10 bits;
- Um pino de Reset;
- Um pino de habilitação Wi-Fi;
- Um pino de alimentação 3.3V;
- Um pino de GND (Ground);
- Três pinos para comunicação SPI (*Serial Peripheral Interface*);
- Dois pinos para comunicação I2C.

Figura 13 - Pinagem ESP8266 12E.



Fonte: Espruino.com.

2.8. Jasper

O *Jasper* criado em 2014 por Charles Marsh e Shubhro Sahaé, é uma plataforma de código aberto instituído especificamente para permitir a interação de usuários com o *Raspberry Pi* através de comandos de voz, além de comandos de ativação de periféricos que podem ser implementados através de uma interface simples em *Python*. Ele disponibiliza informações e notificações das redes sociais através de autofalantes integrados à placa, como temperatura e hora atual da localidade, notificações de mensagens recebidas tanto do *E-mail* ou do *Facebook* do usuário cadastrado nas configurações.

O *script* de instalação *Jasper* deve ser baixado no diretório principal do *Raspbian*, diretamente de uma plataforma de hospedagem de código-fonte *Open Source*, o *GitHub*. Para que um comando de voz seja entendido, por *Jasper* se tratar de um assistente, é necessário a instalação de um mecanismo voz para texto. Dentre os motores STT (*Speech-To-Text*) sugeridos estão:

- ***PocketSphinx STT engine*** - é uma versão do *CMU Sphinx*, podendo ser usada em sistemas embarcados (por exemplo, com base em um processador ARM). É um decodificador de fala de código aberto rápido e projetado para funcionar bem em sistemas como o *Raspberry Pi*. Entretanto, a taxa de reconhecimento não é a melhor.
- ***Julius STT*** - é um *software* de decodificação de reconhecimento de fala contínuo (LVCSR) de alto desempenho. Por outro lado, o modelo acústico exige treinamento.
- ***AT & T STT*** - é um decodificador de voz feito pela empresa de telecomunicações AT & T.
- ***Wit.ai STT*** - conta com os serviços da nuvem *wit.ai* e usa *Crowd Sourcing* para treinar algoritmos de reconhecimento de fala. Assim como o *Google Speech* e o *AT&T* ele também realiza decodificação online e, portanto, precisa de uma conexão ativa com a internet.
- ***Google STT*** - é o sistema de fala para texto do Google. O *API Google Speech* conta com muitos padrões de comparação e com muitos computadores que agem com uma rede de neurônios que processa a informação e devolve um resultado comum, o que torna seu serviço de conversão um dos melhores existentes. O *speech* pode apresentar diferentes resultados em função do utilizador baseando-se no seu histórico de procura e utilizando-se de probabilidade para reduzir os erros de diálogo. Quanto mais dados forem trabalhados, mais preciso se tornará. Quando o *STT* recebe o arquivo de áudio

para o reconhecimento, o mesmo passa para o domínio da frequência, divide o espectrograma em partes e envia para diferentes computadores. Esses computadores processam o som baseando-se no modelo de redes neurais e tentam encontrar os elementos individuais que o compõem como as vogais e consoantes numa das camadas. Outra camada tenta identificar os grupos dos quais fazem parte esses sons fundamentais, e assim sucessivamente, até encontrar uma estimativa final da palavra, retornando ao usuário.

2.9. Sistema Operacional *Raspbian Wheezy*

O sistema operacional a ser utilizado neste trabalho no *Raspberry Pi* será o *Raspbian*, que é uma variante do Debian baseada no ARM *hard-float*, sendo um suporte da arquitetura *Wheezy*, otimizada para o conjunto de instruções ARMv6 do *hardware* do *Raspberry Pi*.

A imagem de disco deverá ser gravada em um cartão *SD Card* de no mínimo 8 GB com um *software* específico *Win32Diskimager*, onde o sistema de arquivos deverá ser expandido através das configurações ao iniciar o *Raspberry Pi*.

Logo após a gravação da imagem e inicialização do sistema, será feita a instalação manualmente de algumas bibliotecas necessárias para o funcionamento do *Jasper*, diretamente de um repositório para as duas abordagens. São elas:

- *Python*, que é uma ferramenta que serve basicamente para desenvolver e interpretar módulos em *Python*.
- *Libasound*, que contém os arquivos necessários para desenvolver *softwares*, que usa a biblioteca ALSA (*Advanced Linux Sound Architecture*).
- *LibaPortAudio*, é uma biblioteca de entrada e saída de áudio, que utiliza um mecanismo de retorno de chamada para solicitar processamento de áudio.
- *Python-pyaudio*, que tem a função de fornecer ligações *Python* para o *PortAudio* facilitando, assim, a utilização do *Python* para reproduzir e gravar áudio.
- *Pip*, gerenciador de pacotes usados para instalar e gerenciar pacotes de *software* escritos na linguagem de programação *Python*.

As configurações de entrada de áudio devem ser alteradas no ALSA (sistema de controle de áudio do sistema operacional), selecionando o *driver* externo como dispositivo de entrada e saída de áudio principal.

3. MATERIAIS E MÉTODOS

Neste trabalho, duas abordagens foram realizadas para a automação residencial por comandos de voz. Neste capítulo, serão apresentados os materiais e métodos utilizados em ambas as abordagens, primeiramente explicando sobre os *softwares* conversores de fala para texto e a metodologia empregada e, em seguida, os dispositivos de *hardware* para o controle de equipamentos eletroeletrônicos.

Na primeira abordagem foi utilizado o *Pocketsphinx* na língua inglesa e uma placa de circuito que é acoplada ao *Raspberry Pi B+*, podendo controlar até seis cargas individuais. Na segunda abordagem é apresentado o sistema na língua portuguesa com o STT Google *speech* e uma nova metodologia para o controle das cargas via Wi-Fi através do ESP8266 12E, que além do sistema apresentar melhores resultados em termos de reconhecimento de fala em relação a primeira abordagem e benefício, teve um custo de *hardware* também menor em relação a primeira abordagem. Esses custos são apresentados ao final da descrição de cada abordagem, realizada a seguir.

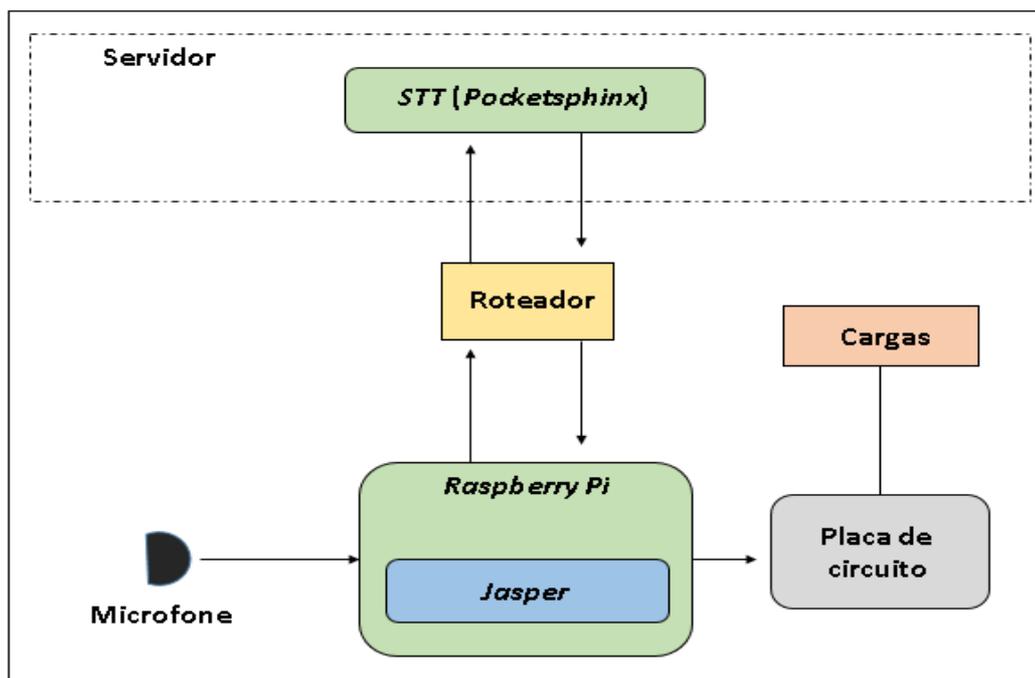
3.1. Primeira abordagem – Reconhecimento na língua inglesa.

3.1.1. *Software*

Para a primeira abordagem foi utilizado o modelo apresentado na Figura 14. Como conversor de voz para texto foi escolhido o *Pocketsphinx* na língua inglesa. Escolheu-se usar o *Pocketsphinx* por ser um conversor bastante conhecido em sistemas automáticos de voz na língua inglesa, por fornecer taxas de acurácia acima de 90% (SURESH E RAO, 2015).

O *Pocketsphinx* é um decodificador de fala independente de falantes, de código aberto baseado no projeto CMU *Sphinx*, que utiliza modelos acústicos baseados em HMMs (*Hidden Markov Models*) e um modelo de linguagem estatística de *n*-grama. O *PocketSphinx* foi especialmente adaptado para dispositivos portáteis e móveis que tenham em sua estrutura um processador ARM como o *Raspberry Pi*.

Figura 14 - Modelo Geral da Metodologia Aplicada com o *Pocketsphinx*.



Fonte: Autoria própria.

O *Raspberry Pi B+* embarca o *software Jasper*, que após a captura da fala através de um microfone, processa e o converte para texto a partir de um motor STT. O texto, então, é validado de acordo com as palavras chaves dos módulos, que são *scripts* criados para o controle das GPIO's (*General Purpose Input/Output*) do *Raspberry Pi B+*, que, por sua vez, ativa e desativa as cargas (Lâmpadas, ventiladores, TVs etc) conectados à placa de circuito.

3.1.2. Hardware

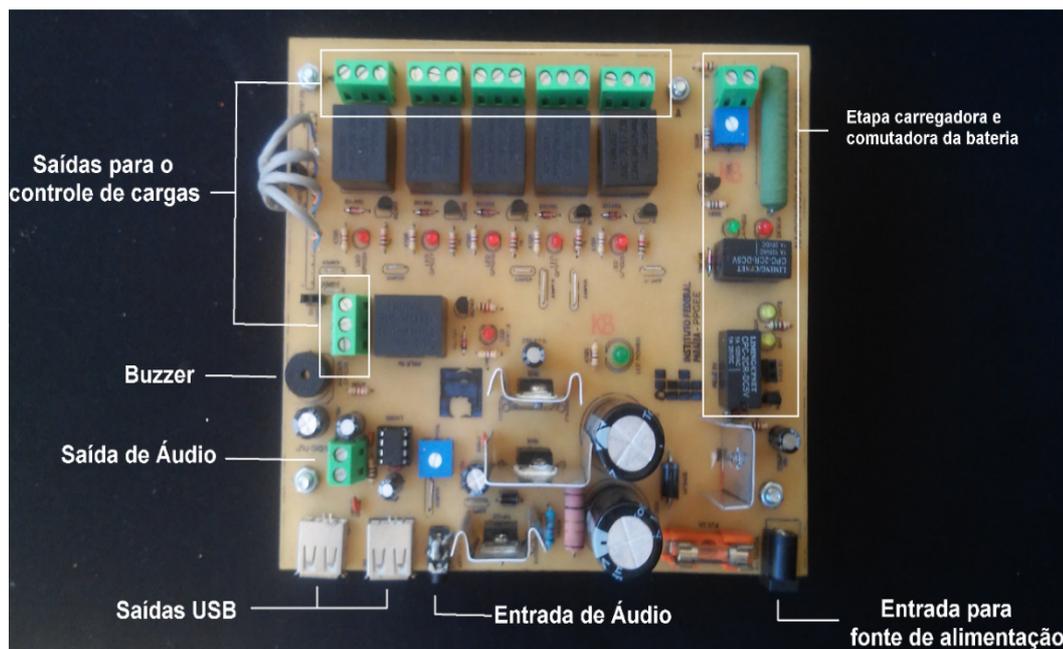
A plataforma *Raspberry Pi* foi escolhida por ser um microcomputador portátil de tamanho reduzido, com alta taxa de processamento, configurações de *hardware* e um sistema operacional *Open Source* que atenderia à proposta. E em contrapartida à proposta de Pimentel (2014) que se utilizou da plataforma *BeagleBone Black*, desenvolver um sistema de reconhecimento automático de fala com igual ou superior eficiência, porém com o valor de custo mais reduzido utilizando-se plataformas embarcadas de fácil aquisição no mercado. Pesquisas de valores feitas no Brasil em janeiro de 2018 entre as duas, mostraram que o *Raspberry Pi Zero W* apresenta o menor valor entre a categoria, com uma diferença de custo 66% mais barato que a *BeagleBone Black*.

O modelo de placa utilizado para a primeira proposta foi o *Raspberry Pi B+* (Figura 9) pela disponibilidade, e um cartão *SD Card* de 8 GB com uma versão do *Raspbian* pré-gravada.

Para a prototipagem foi feita uma placa de circuito impresso (Figura 15), que é acoplada ao *Raspberry Pi B+* através de um suporte (Figura 16) e conectores, facilitando, assim, o acionamento dos atuadores, que serão controlados. A mesma consta com:

- Seis relés para o controle de cargas, com tensão nominal de até 220 V 10 Ampères cada, suportando uma potência máxima de 2200 W.
- Um *Buzzer* para indicação sonora;
- Fusível para proteção contra sobrecarga;
- Amplificador de áudio de 3 W;
- Duas saídas USB's para alimentar dispositivos que venham a ser necessários para o funcionamento do protótipo;
- Circuito carregador de bateria selada 6 V, que tem a função de alimentar a placa na falta de energia da fonte externa fazendo a comutação automática entre as duas formas de fornecimento, mantendo assim o SO (Sistema Operacional) ativo.

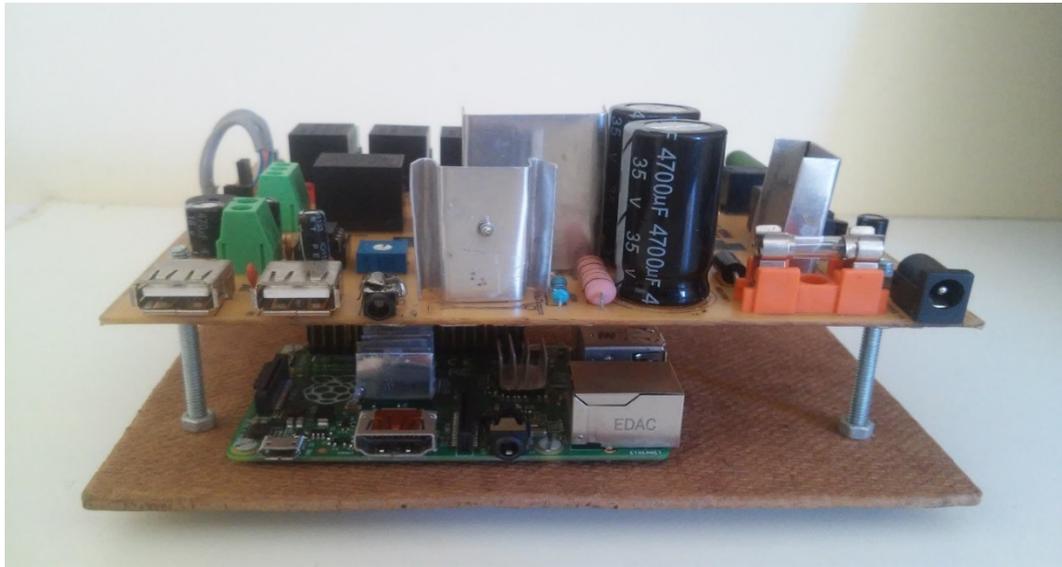
Figura 15 - Placa de circuito.



Fonte: Autoria Própria.

A placa é alimentada com uma fonte externa chaveada de 9 V por 2 A. Na placa de circuito, reguladores de tensão fazem a conversão para os 5 V necessários para o funcionamento do *Raspberry Pi B+* e dos outros dispositivos.

Figura 16 - Estrutura para acoplamento com o *Raspberry Pi*.



Fonte: Aatoria Própria.

Outros dispositivos de *hardware* foram utilizados para a programação, configuração e testes no *Raspberry Pi*, tais como teclado, mouse, adaptador *Wi-Fi*, cabo HDMI para conexão com o monitor de vídeo e um microfone com filtro “POP” (Figura 18), acoplado a um drive de som externo USB (Figura 17), para a captura da voz.

Figura 17 - Adaptador de Áudio.



Fonte: Lazada.com.my.

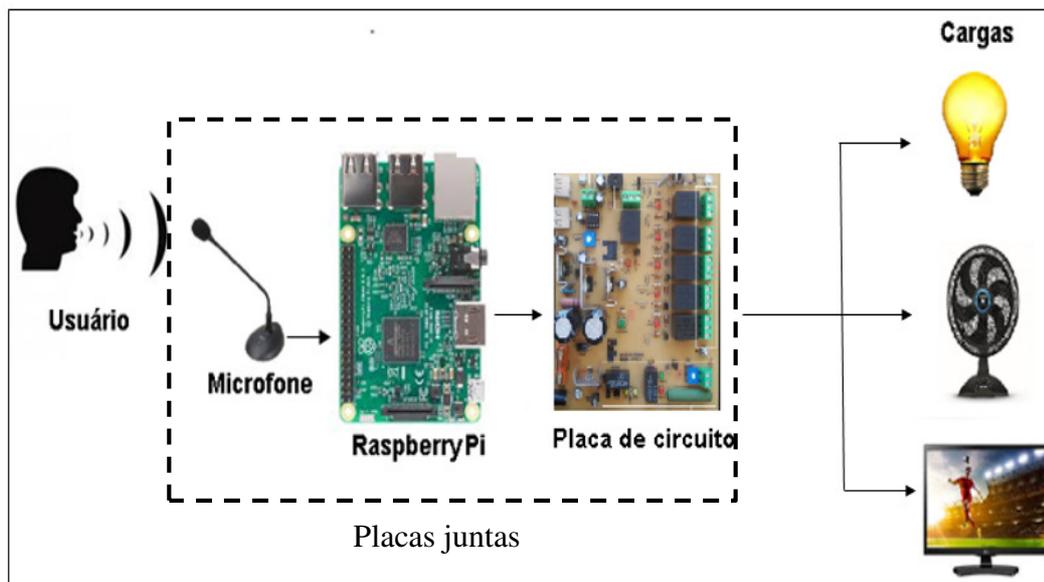
Figura 18 - Microfone plugue P2 com filtro “POP”.



Fonte: Filipeflop.com.

Para essa abordagem, o dispositivo teve a placa de controle das cargas centralizada, ou seja, a placa fica em um determinado ponto fixo onde o usuário deve-se locomover até sua proximidade para ditar o comando. A placa também é dedicada a um número máximo de 6 aparelhos ou dispositivos a serem controlados, conforme a Figura 15.

Figura 19 - Placa de controle dos dispositivos centralizada.



Fonte: Autoria própria.

O motor TTS (*Text-to-speech*) que converte texto para voz, é uma ferramenta indispensável para o *Jasper* padrão. Foi dispensado para este projeto, por aumentar o custo do sistema, sendo substituídos os alertas e notificações internas do *Jasper*, por sinais sonoros provenientes de um *Buzzer*, adaptado na placa de circuito, obtendo um ganho em custo de processamento.

O *software* disponibiliza dois modos de operação: a forma passiva, onde deve ser pronunciada uma palavra de ativação antes do comando desejado, palavra esta que pode ser definida pelo usuário, e a forma ativa, que fica “escutando” constantemente a espera de alguma palavra definida em um de seus módulos.

Para os testes iniciais, o aparato funciona nos dois modos, ativo e passivo, onde para alterar de um modo para o outro, se faz necessário dizer a palavra “*mode*” (Modo), a qualquer momento. A palavra para a forma passiva escolhida foi “*home*” (Casa) que, após identificada, é emitido um sinal sonoro (dois bips) que indica o início da escuta, onde o usuário tem 7 segundos para falar um dos comandos (Tabela 5). Esse tempo foi estimado para que o usuário completasse os comandos sem que houvesse o corte da palavra. Em seguida, outro sinal (um bip) indicava o fim da escuta. O tempo em que o sistema fica em estado de escuta pode ser alterado, conforme a necessidade.

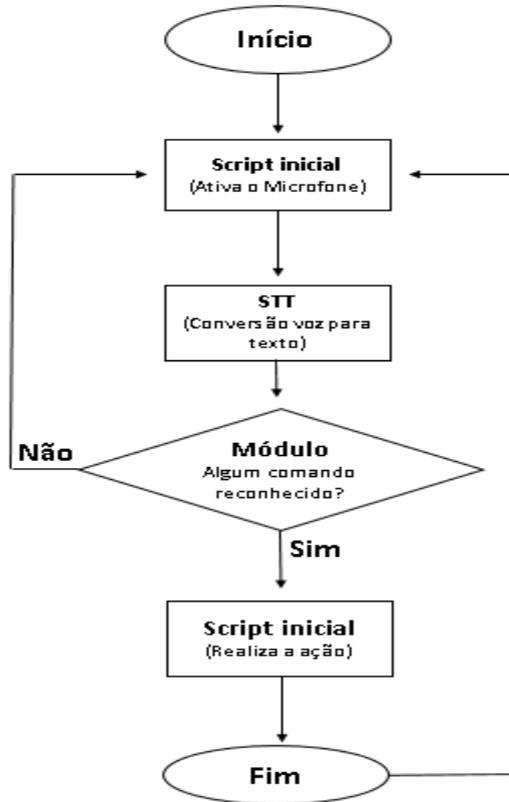
Tabela 5 - Comandos definidos para os módulos e suas respectivas ações.

Comandos		Ação
<i>Door</i>	<i>Open</i>	Ativa relé 1
	<i>Close</i>	Desativa relé 1
<i>Garden</i>	<i>Light on</i>	Ativa relé 2
	<i>Light off</i>	Desativa relé 2
<i>Kitchen</i>	<i>Light on</i>	Ativa relé 3
	<i>Light off</i>	Desativa relé 3
<i>Bathroom</i>	<i>Light on</i>	Ativa relé 4
	<i>Light off</i>	Desativa relé 4
<i>Hall</i>	<i>Light on</i>	Ativa relé 5
	<i>Light off</i>	Desativa relé 5

Fonte: Autoria Própria.

Para cada comando que deveria ser reconhecido pelo sistema, foi criado um módulo (*Script* em *Python*), esses módulos podem ser definidos em ordem de prioridade (hierarquia), o que faz com que o sistema busque primeiramente aqueles módulos com maior nível, seguindo uma ordem decrescente. Por exemplo, para os comandos de ativação referentes aos cômodos, o nível de prioridade era um. Para os comandos “*Open*”, “*Close*” e “*Light*” o nível de prioridade era 2 e para “*Off*” e “*On*”, o nível de prioridade era 3. Nestes módulos são inseridas as palavras chaves (comandos) que deverão ser reconhecidas. Segue, na Figura 20, o fluxograma do algoritmo implementado.

Figura 20 - Fluxograma do algoritmo desenvolvido em *Python*.



Fonte: Autoria própria.

Após a implementação, realizou-se um teste para validação do sistema. Os testes foram realizados com cinco voluntários adultos, alheios ao desenvolvimento do projeto, com idade entre 18 e 26 anos, sendo três homens e duas mulheres, todos com nível de inglês básico, não fluentes na língua. Segundo estimativas de Nielsen (1994) para uma avaliação de usabilidade, de três a cinco pessoas, são suficientes para identificar 95% dos problemas existentes. O microfone foi ajustado nas configurações do *Raspbian*, com sensibilidade igual a 50%, dessa forma capturando o sinal de voz com o menor ruído ambiente possível. Os comandos (Tabela 5) eram ditos a uma distância de aproximadamente 10 cm do microfone, que segundo Gustavo (2013), para uma gravação natural da fala, a distância entre 10 a 13 cm do microfone é suficiente. Os testes foram realizados em um ambiente domiciliar comum, sem as cargas ligadas ao circuito.

O valor individual de cada componente e a média geral de custo, para essa implementação, podem ser vistos na Tabela 6.

Tabela 6 - Custo para a implementação da primeira abordagem.

VALORES PARA A PRIMEIRA ABORDAGEM		
Nº	Descrição	Valor
1	<i>Raspberry Pi B+</i>	R\$ 152,00
2	Drive de Áudio	R\$ 25,00
3	Fonte 5V/2A	R\$ 15,00
4	Placa de circuito para 6 cargas	R\$ 160,00
5	Adaptador Wi-Fi	R\$ 20,00
Custo Total		R\$ 327,00

Fonte: Autoria própria.

3.2. Segunda abordagem – Reconhecimento na língua portuguesa

Visto que os sistemas automáticos que utilizam comandos de fala, os usuários têm que se deslocar até próximo do dispositivo para ditar o comando, uma alternativa pensada foi tornar o dispositivo de reconhecimento móvel, dentro do ambiente doméstico, ou seja, que possa ser facilmente deslocado para próximo do usuário, permitindo assim maior comodidade e segurança. Outro recurso implementado é o serviço de SMS (*Short Message Service*), onde o usuário através de um único comando de voz pode enviar uma mensagem de texto pré-definida ao aparelho celular de pessoas cadastradas no sistema como, por exemplo, um pedido de ajuda.

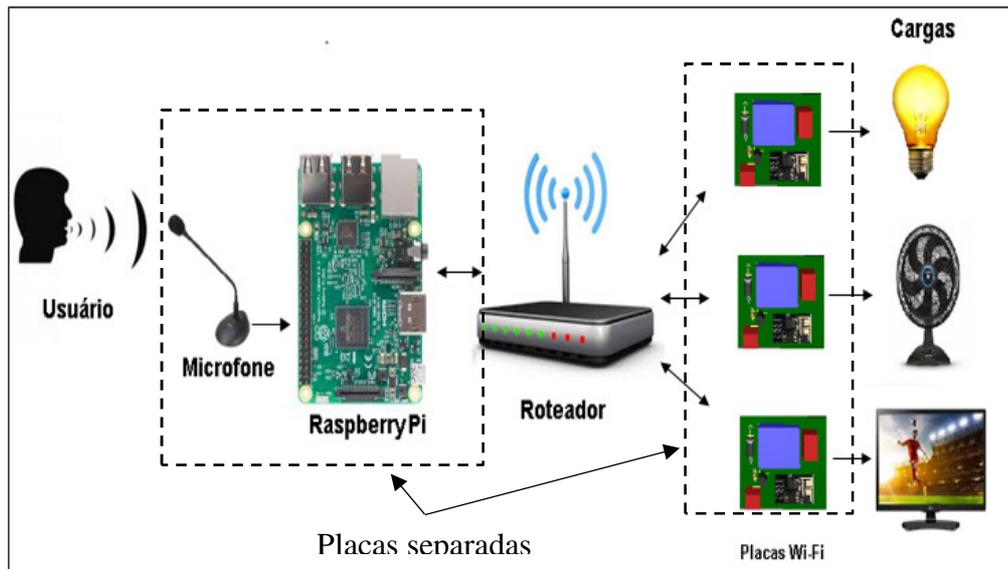
A ideia implementada para a segunda abordagem do projeto é a de que o *Raspberry Pi Zero W*, após fazer o reconhecimento do comando, envie um dado para o controle dos dispositivos a serem controlados por *Wireless* à placa de circuito. Além disso, que embarque o ESP8266 12E, que é um dispositivo embarcado de baixo custo, com um bom alcance determinado pelo roteador e eficiente por possui GPIO's configuráveis e trabalhar com um baixo nível de tensão.

O aparato trouxe maior comodidade pela portabilidade e segurança, por conta do dispositivo de reconhecimento de fala ser isolado da placa controladora dos aparelhos que ficarão ligados da rede elétrica (Figura 21). Outra vantagem é que podem ser adicionados mais dispositivos a serem controlados em outros ambientes sem a necessidade de alteração nas placas de circuitos já existentes.

Também foi substituído o *PocketSphinx STT engine*, que por uma API (*Application Programming Interface*) com suporte para o idioma português-Brasil. Por ser o idioma nativo,

facilita a interação do usuário, o que melhorou a eficácia consideravelmente. Além disso, foi feita a substituição da plataforma embarcada por uma da mesma família, porém com o tamanho reduzido, maior capacidade de processamento e financeiramente mais acessível.

Figura 21 - Placa de controle dos dispositivos descentralizada.



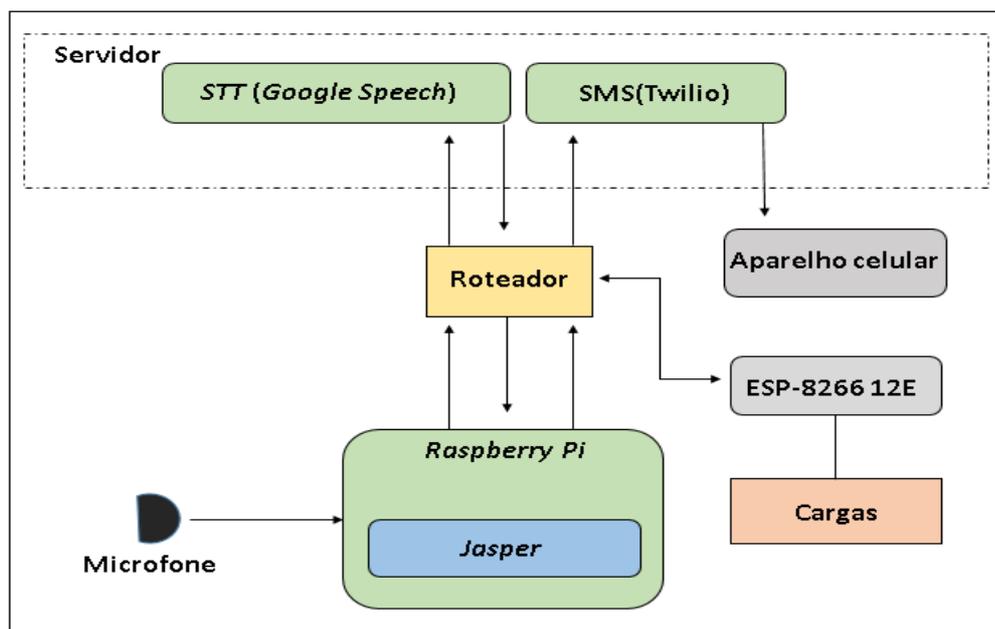
Fonte: Autoria Própria.

3.2.1. Software

Para a segunda abordagem o motor STT *Pocketsphinx* foi substituído pelo *Google speech API*. Uma das sugestões de conversores de voz para texto pelo *Jasper* é o *Google speech*, que reconhece mais de 110 idiomas entre eles o português do Brasil, o qual foi escolhido para a aplicação final deste projeto.

O *Google speech* converte arquivos de voz em texto, aplicando modelos de rede neurais avançadas, em que a precisão vai aumentando com o passar do tempo. Oferece *streaming* dos resultados de texto, onde o texto é reconhecido imediatamente após a fala, retornando os resultados de reconhecimento parcial, conforme ficam disponíveis. A API processa áudio mesmo em ambientes barulhentos, assim o sistema não precisa estar fazendo um processamento do sinal antes, como cancelamento de ruídos. O modelo metodológico para essa segunda implementação pode ser visto na Figura 22.

Figura 22 - Modelo Geral da Metodologia Aplicada com o Google *Speech API*.



Fonte: Autoria Própria.

Para a utilização do *Google speech API* se faz necessária a criação de uma conta no grupo *Chromium dev* e, em seguida, a criação de um projeto através do console do *Google Developers*. O console disponibiliza uma chave individual de acesso, onde a mesma será solicitada no *script* de configuração de dados pessoais do usuário *Jasper*. Quando o sistema passa a funcionar é feito um requerimento ao servidor onde se utiliza dessa chave de acesso. O idioma a ser reconhecido é selecionado no *script* STT *Jasper*, onde são colocadas a abreviação da língua a ser reconhecida e do país, como por exemplo “*pt-br*” para o português do Brasil, “*en-us*” para o inglês americano, o que torna o sistema flexível.

A proposta tem como objetivo descentralizar a placa que faz o reconhecimento da fala, da placa de controle dos dispositivos, que utilizará comunicação através do Wi-Fi. O roteador além de servir de intermédio entre o servidor do *Google speech* e *Twilio*, também é responsável por fazer a comunicação entre o *Raspberry Pi* e o ESP8266 12E.

Foi desenvolvido um algoritmo próprio para o ESP8266 12E em *Wiring* utilizando a IDE (*Integrated Development Environment*) utilizada para programação do Arduino. O ESP se comunica diretamente com o roteador através de um IP fixo determinado pelo usuário, e realiza a ação determinada em cada item condicional dependendo do caractere recebido do *Raspberry Pi*, após o comando de fala reconhecido e validado.

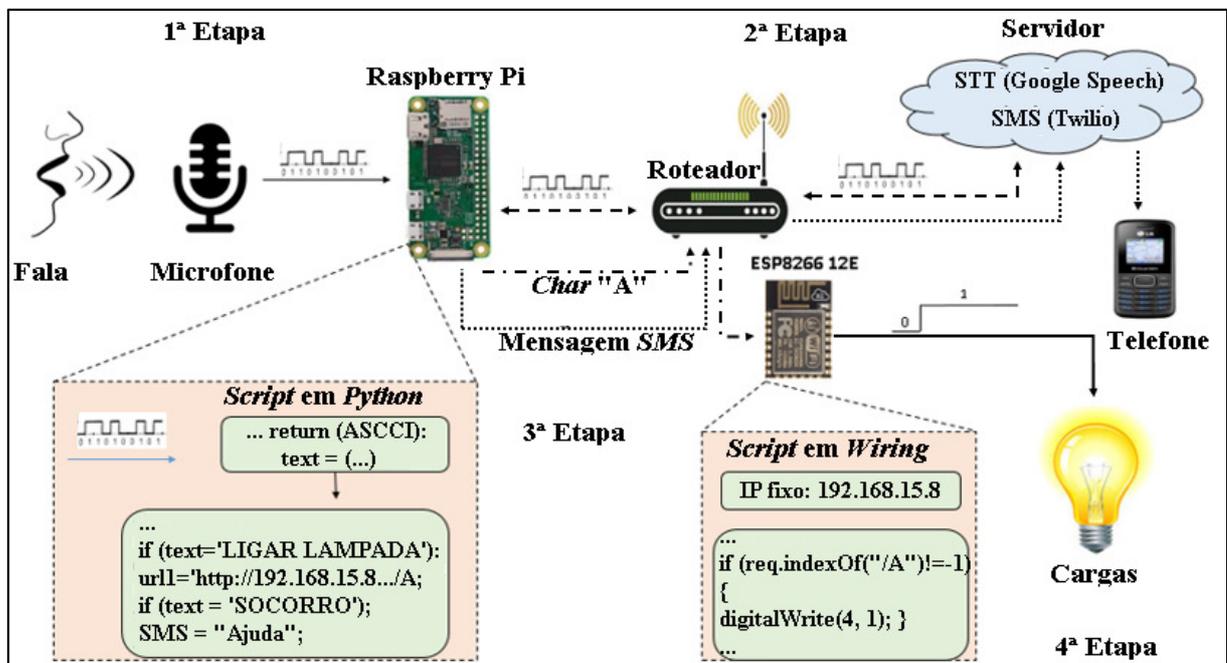
Foi necessário fazer modificações no *script* desenvolvido em *Python* da primeira abordagem para a implementação do *Google speech API*. Para comunicação com o ESP8266

12E foi acrescentada a biblioteca “*urllib2*” para o envio HTTP do caractere a ser reconhecido, a biblioteca “*Twilio*” para o envio de SMS e os comandos para português, visto que na primeira abordagem era na língua inglesa. Para essa abordagem não é necessário a criação de módulos como na abordagem anterior ou de dicionário.

Na pasta de controle do microfone “*mic.py*”, foi determinada a GPIO de saída para o *Buzzer*, onde foi determinado o som de dois bips para indicar o início da escuta e um bip para indicar o final em um período de 3 segundos, período estimado para o usuário falar o comando sem cortes. Para a taxa de amostragem de áudio foi setado um valor de 16 kHz, compatível com as taxas de amostragem de *stream* aceitáveis pelo *Google speech* que é entre 8 kHz e 48 kHz. Além disso, a sensibilidade de captura do microfone pode ser ajustada no diretório do ALSA instalado previamente no sistema operacional. Para essa proposta, o ajuste foi em 50%.

A Figura 23 mostra detalhadamente o modelo de funcionamento do sistema implementado e, na Figura 24, é ilustrado o fluxograma. As setas contínuas representam conexão por fio ou por conector, as setas tracejadas indicam a comunicação via *Wireless*. Um guia para o significado das setas pode ser visto na Tabela 7.

Figura 23 - Metodologia detalhada da segunda abordagem.



Fonte: Autoria própria.

Tabela 7 - Indicativo de comunicação entre as setas.

Seta	Indicativo
→	Comunicação USB (Microfone x <i>Raspberry Pi</i>)
- - - →	Comunicação Wi-Fi (<i>Raspberry Pi</i> x Roteador x Servidor STT)
..... →	Comunicação Wi-Fi (<i>Raspberry Pi</i> x Roteador x Servidor SMS x Telefone)
- . - - →	Comunicação Wi-Fi (<i>Raspberry Pi</i> x Roteador x ESP8266 12E)
→	Comunicação por fio (ESP8266 12E x Cargas)

Fonte: Autoria própria.

Tabela 8 - Comandos definidos a serem reconhecidos pelo sistema.

Comandos		Ação
Ligar	Lâmpada	Ativa o LED 1 (GPIO 17)
Desligar		Desativa o LED 1 (GPIO 17)
Ligar	Ar condicionado	Ativa o LED 2 (GPIO 23)
Desligar		Desativa o LED 2 (GPIO 23)
Ligar	TV	Ativa o LED 3 (GPIO 24)
Desligar		Desativa o LED 3 (GPIO 24)
Abrir	Persiana	Ativa o LED 4 (GPIO 27)
Fechar		Desativa o LED 4 (GPIO 27)
Ligar	Abajur	Ativa o LED 5 (GPIO 22)
Desligar		Desativa o LED 5 (GPIO 22)
Socorro		Ativa e desativa todos os LED's (Pisca)
Encerrar		Ativa todos os LED's

Fonte: Autoria própria.

Na primeira etapa o usuário deve falar uma palavra de ativação do sistema de automação. As palavras a serem reconhecidas só serão validadas se a palavra de ativação “Quarto” for dita antes. O início da escuta será indicado por dois bips. O usuário terá três segundos para ditar o comando desejado. Em seguida, um bip indicará o final da escuta. O microfone faz um pré-processamento antirruídos e digitalização da fala antes de enviar ao *Raspberry Pi*.

Na segunda etapa após receber esses dados digitalizados o *Raspberry Pi* faz um requerimento ao servidor STT por meio do roteador para o reconhecimento da palavra, que retorna outro arquivo digitalizado referente a palavra reconhecida.

Na terceira etapa o *script* principal recebe esses dados e o converte em texto através da tabela ASCII que retorna a palavra que foi reconhecida pelo servidor sem o uso de acentos. Se a palavra for um dos comandos da Tabela 8, que está na rotina condicional, então é realizada a ação definida, que corresponde à etapa quatro.

Por exemplo, se a palavra reconhecida for “*LIGAR LAMPADA*” é enviado ao roteador o IP referente ao ESP8266 12E e o caractere referente à ativação da GPIO setada para acionar o relé que está ligado à lâmpada. O roteador, por fim, envia o comando ao ESP que realiza a ação determinada em outra rotina condicional, etapa quatro. No caso, se a palavra reconhecida for “*SOCORRO*”, é feito um requerimento ao “*Twilio*” que envia a mensagem (Figura 25) definida no *script* para o número de telefone salvo. Para a utilização desse servidor, é necessária a criação de uma conta na página, onde serão geradas uma identidade e uma senha temporária de proteção (*token*) para o usuário.

O *Twilio* permite aos desenvolvedores a incorporação VoIP e mensagens em seus aplicativos. *Twilio* conta com um ambiente baseado em nuvem, onde o usuário pode acompanhar toda a infraestrutura. Dos serviços que dispõe, o *Twilio Voice* permite que a aplicação faça ligações telefônicas. *Twilio SMS* permite o envio e o recebimento de SMS e o Cliente de *Twilio* permite chamadas de voz sobre IP (VoIP) de qualquer telefone, *tablet* ou navegador.

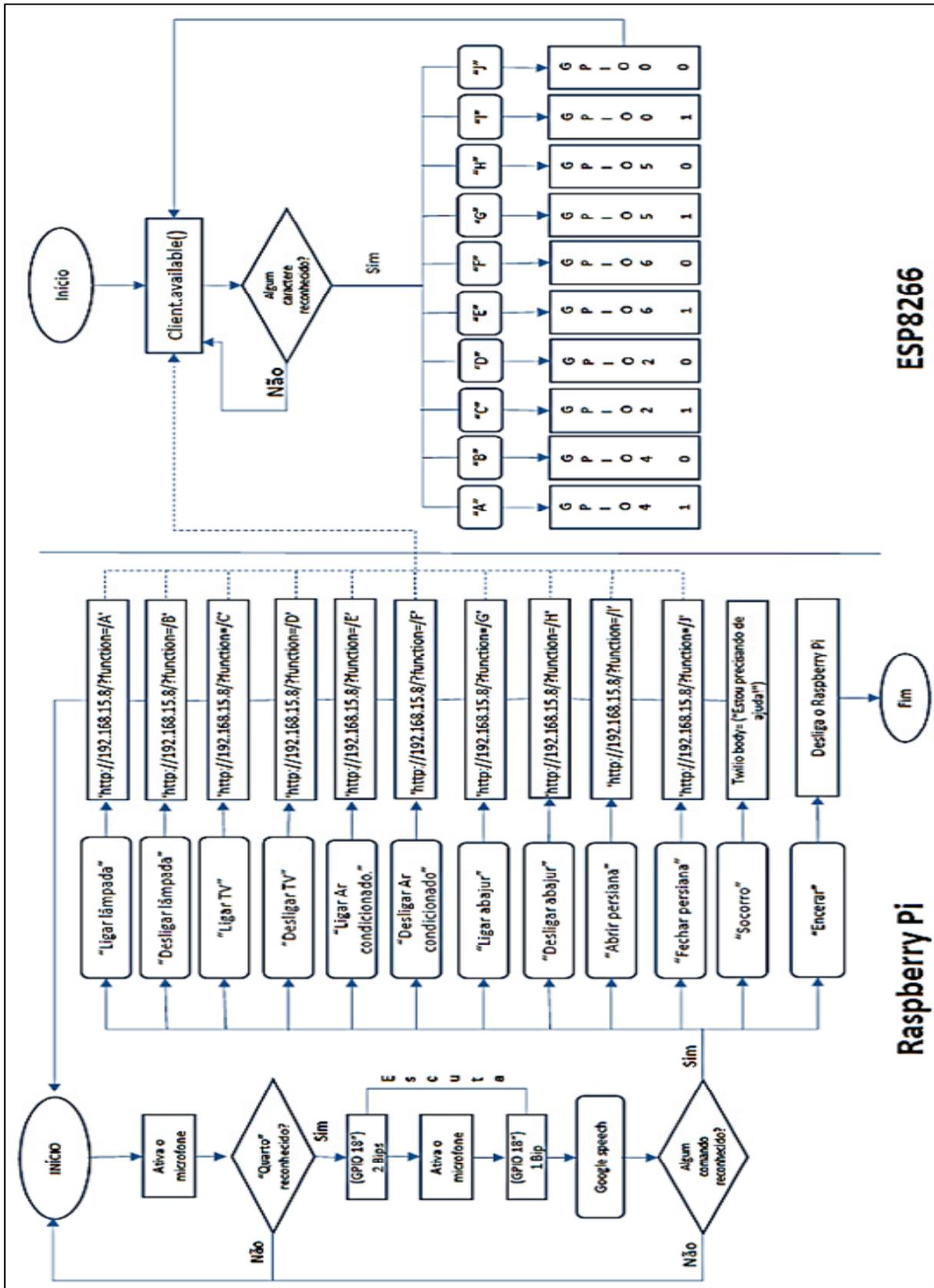
O comando “*ENCERRAR*” foi inserido visando desligar o *Raspberry Pi* no caso de uma eventualidade ou necessidade de deslocamento do sistema para outro ambiente. Esse comando encerra o sistema operacional via *software* evitando, assim, que o sistema operacional não seja corrompido ou o cartão *SD Card* danificado.

Figura 24 - Mensagem enviada através do *Twilio* API.



Fonte: Autoria própria.

Figura 25 - Fluxograma geral segunda abordagem.



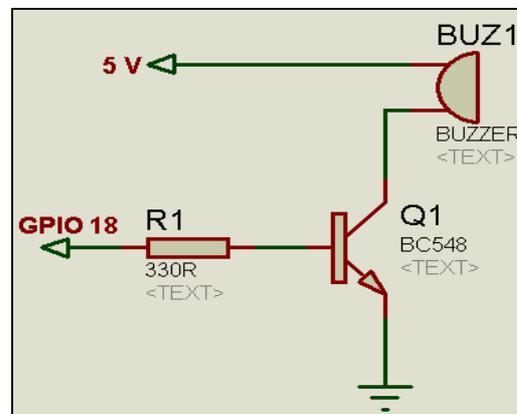
Fonte: Autoria própria.

3.2.2. Hardware

O modelo de placa utilizado foi o *Raspberry Pi Zero W* (Figura 11), um cartão *SD Card* de 8 GB com uma versão do *Raspbian Wheezy* pré-gravada e a plataforma *ESP 8266 12E* (Figura 12) juntamente com dispositivos eletrônicos para o controle das cargas.

Um sinal sonoro de alerta de início e final da escuta é reproduzido por um *Buzzer* conectado ao *Raspberry Pi*, seguindo o esquema da Figura 26. O tempo de escuta é setado no *script mic.py*, responsável também para ativar o microfone nas solicitações do sistema.

Figura 26 - Esquema eletrônico de conexão entre *Buzzer* e o *Raspberry Pi*.



Fonte: Autoria própria.

O *Driver* de áudio foi substituído por um com o microfone já embutido (Figura 27), também com conexão USB, menor tamanho e custo, com a única intenção de reduzir o valor do projeto, não diferenciando em nada a qualidade de captura da fala em relação ao *Driver* da primeira abordagem, reconhecendo comandos a uma distância de aproximadamente 50 cm.

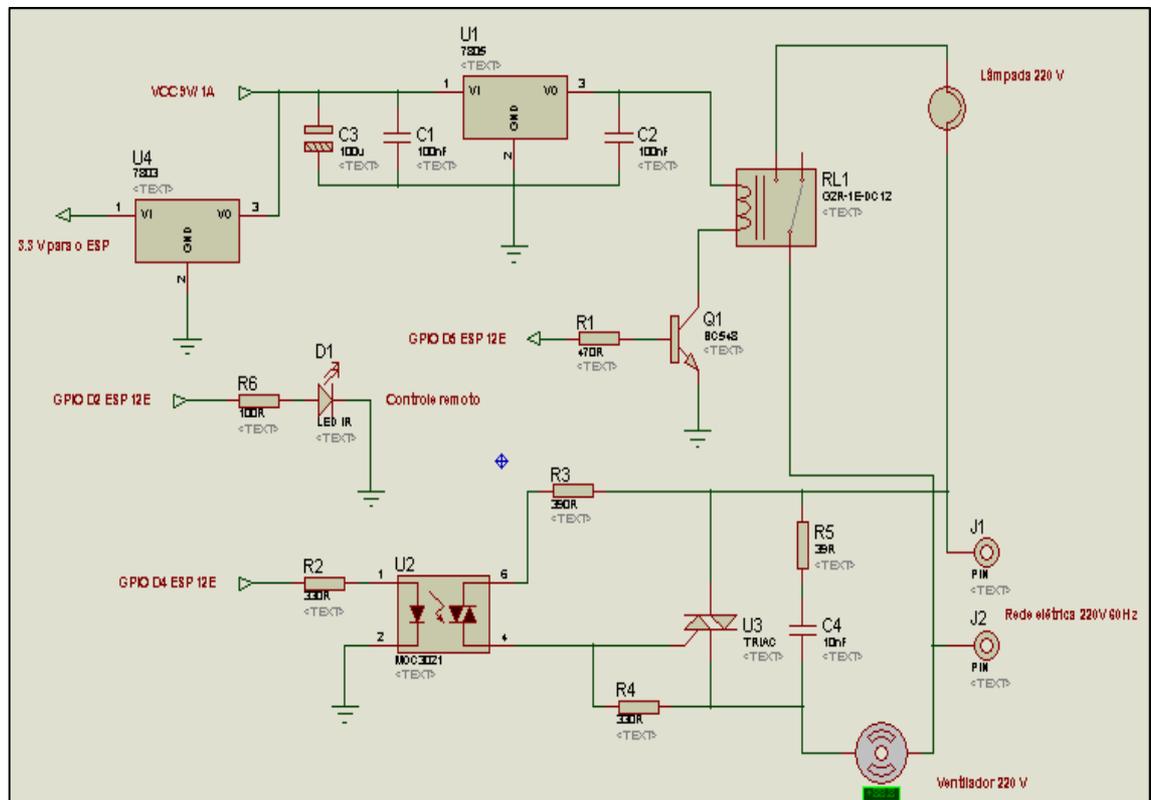
Figura 27 - Microfone utilizado na segunda abordagem.



Fonte: Retirado da Internet.

Para a segunda abordagem os testes com os voluntários foram realizados na Protoboard (Figura 30), utilizando somente LED's, e em uma aplicação real, ou seja, o sistema de automação foi instalado em um quarto para a validação, utilizando o ESP8266 12E. O esquema eletrônico está ilustrado na Figura 28 e a montagem final na Figura 29.

Figura 28 - Esquema eletrônico acionamento de cargas.



Fonte: Autoria própria.

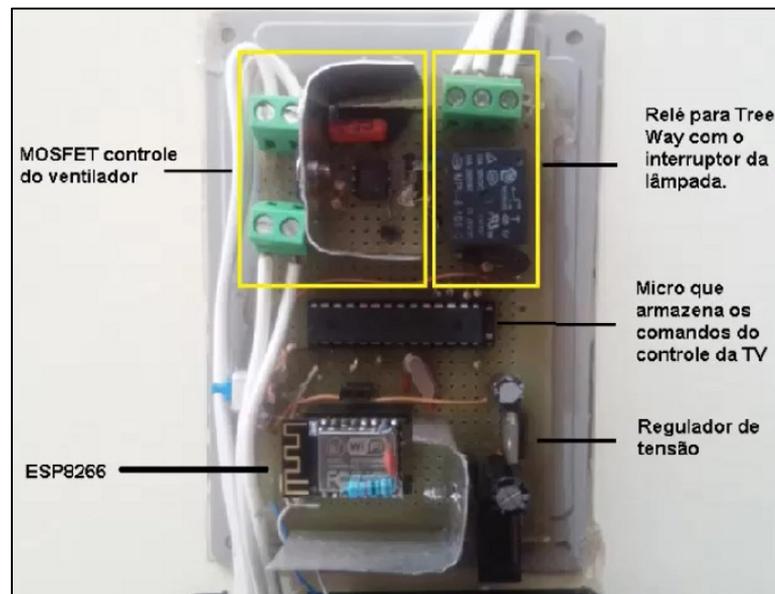
Figura 29 - (a) Caixa patota onde o circuito foi montado e fixado na parede (b) Montagem eletrônica.

(a)



Fonte: Autoria própria.

(b)



Fonte: Autoria própria.

Para a alimentação da placa foi utilizado uma fonte de 9 V/1A. A placa conta com dois reguladores de tensão um de 5 V para alimentação do relé e outro 3,3 V para a alimentação do ESP8266 12E.

O chaveamento da lâmpada feito por relé possibilita a ligação em paralelo (*Three-way*) com o interruptor onde, no caso de falha do circuito eletrônico, o usuário pode controlar a lâmpada manualmente. Para o ventilador, deu-se preferência a um relé de estado sólido com um Triac TIC226, comutado por optoacoplador MOC3021, que pode ser acionado com baixa tensão e tem maior vida útil que um relé convencional de contatos mecânicos.

A clonagem dos códigos do controle remoto da TV foi feita utilizando um receptor infravermelho IR TSOP4838 acoplado a um Arduino UNO para a captura. Após a captura, foi inserida a codificação no ESP8266 12E, e utilizado um LED emissor IR acoplado ao ESP para envio do código ao aparelho de TV quando requisitado.

Foi gravado um teste real do sistema em português, onde após o reconhecimento da fala o *Raspberry Pi* envia o comando ao ESP8266 12E, que por fim controla as cargas (Lâmpada, Ventilador e TV). O valor individual de cada componente e a média geral de custo para essa implementação pode ser vista na Tabela 9. Um vídeo está disponibilizado na Internet, através do seguinte link, no *YouTube*: <https://youtu.be/OEChlVZeiBc>.

Tabela 9 - Custo para a implementação da segunda abordagem.

VALORES PARA A SEGUNDA ABORDAGEM		
Nº	Descrição	Valor
1	<i>Raspberry Pi Zero W</i>	R\$ 85,00
2	Microfone USB	R\$ 17,00
3	Fonte 5V/2A	R\$ 15,00
4	Placa ESP8266 12E (controla até 11 cargas)	R\$ 20,10
	Custo Total	R\$ 137,10

Fonte: Autoria própria.

A segunda abordagem apresentou uma diferença de preço de 58,07% menor que na primeira abordagem.

4. RESULTADOS

Neste capítulo, será descrito como foram realizados os testes para as duas propostas implementadas, como o número de locutores participantes, *software* de gravação de voz e procedimentos adotados. Serão apresentados os resultados individuais para cada voluntário, links de vídeos disponibilizados na Internet para uma visualização demonstrativa do sistema nas duas línguas e, ao final, uma discussão sobre os resultados.

4.1. Resultados para a primeira abordagem

Os participantes fizeram a pronúncia de cada conjunto de palavras cinco vezes, obtendo um total de 50 comandos pronunciados por voluntário. Os resultados foram anotados na Tabela 10 onde, para cada conjunto de palavras pronunciadas, se obtivesse acerto era marcado um “V” e para não reconhecida um “X”. Ao final do teste, era somado o total de acertos de cada conjunto de palavras individualmente, feita a soma total de todos os comandos e calculada a média dos acertos em porcentagem. Os resultados com as essas taxas de acertos e média geral entre todos os locutores com o *Pocketsphinx* podem ser vistos na Tabela 11.

Tabela 10 - Tabela feita para anotação dos resultados.

Nome					Sexo		Idade		
					() M () F		() <15 () 15 -21 () 22 – 64 () > 64		
1° LED - OPEN DOOR					CLOSE DOOR				
2° LED - GARDEN LIGHT ON					GARDEN LIGHT OFF				
3° LED - KITCHEN LIGHT ON					KITCHEN LIGHT OFF				
4° LED - BETHROOM LIGHT ON					BETHROOM LIGHT OFF				
5° LED - HALL LIGHT ON					HALL LIGHT OFF				

Fonte: Autoria própria.

Tabela 11 - Resultados taxas de acertos e média geral com *Pocketsphinx*.

Usuários	Acertos
1° Locutor (M)	92 %
2° Locutor (M)	88 %
3° Locutor (M)	96 %
4° Locutor (F)	90 %
5° Locutor (F)	86 %
Média geral de acertos	90,4 %

Fonte: Autoria Própria.

4.1.1. Discussão

Os resultados foram satisfatórios, pois só em poucos casos os comandos pronunciados não foram reconhecidos, que podem ter sido ocasionados por alguns fatores:

- Pela pronúncia incorreta do idioma;
- Qualidade do microfone;
- Ruídos externos.

Foi observado que os comandos com o conjunto de três palavras eram mais difíceis de serem reconhecidos, como por exemplo “*Hall Light off* ” ou “*Hall Light on*”, onde foi percebido que as palavras “*off*” e “*on*” não eram reconhecidas de forma correta. A palavra “*off*” por exemplo era reconhecida como “*of*” e o conjunto de palavras referentes a “*Bathroom*”.

A eficiência, por outro lado, se mostrou favorável, sendo que para cada comando aceito a ação correspondente era realizada, ou seja, as GPIO’s eram ativadas ou desativadas instantaneamente.

Foi gravado um teste de validação do sistema em inglês, tendo sido o vídeo disponibilizado na Internet, através do seguinte link, no *YouTube*: <https://youtu.be/qvYYyk1J3XA>.

4.2. Resultados para a segunda abordagem

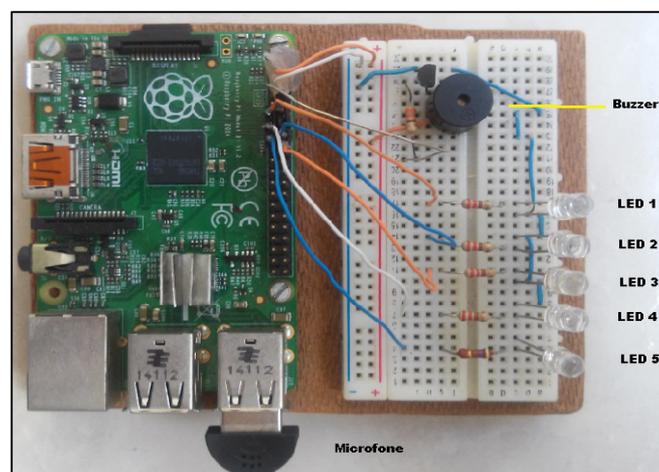
Para a validação da segunda abordagem foi coletada a voz com os respectivos comandos (Tabela 8) a serem reconhecidos pelo sistema com 10 voluntários, entre eles cinco homens e cinco mulheres. Os arquivos de áudio foram gravados através do *software Audacity*, com amostras de 16 kHz, canal mono, e salvos como arquivos .wav (Microsoft 16 bits PCM), que é recomendado pelo *Google speech*. Valores a baixo dessa taxa de amostragem pode prejudicar o reconhecimento e acima, o reconhecimento de fala não tem efeito significativos. O microfone utilizado foi o mesmo conjunto da primeira abordagem (Figuras 17 e 18), conectado ao *notebook*, onde os voluntários faziam as pronúncias dos comandos a uma distância de aproximadamente 10 cm, para evitar a captura da respiração.

Foi realizada uma montagem com uma Protoboard e LED's (Figura 30) para facilitar o transporte do sistema até os voluntários, assim logo após a captura da voz era possível fazer o teste diretamente no próprio sistema, empregando o microfone da placa.

A montagem para teste com os usuários logo após a gravação da voz foi feita utilizando o *Raspberry Pi B+* da primeira abordagem. Assim, o sistema poderia ser conectado de forma fácil à Internet através de um cabo de rede pelo conector RJ45 embarcado na placa, evitando assim ter que entrar no sistema operacional para se conectar a uma rede Wi-Fi.

O *Raspberry Pi* após conectado ao cabo de rede e à fonte de alimentação inicia o sistema operacional e em seguida o sistema de reconhecimento de fala, se não for detectado nenhum erro como falta de Internet ou microfone, o programa inicia de forma automática não precisando o usuário apertar qualquer botão.

Figura 30 - Montagem para testes da segunda abordagem.



Fonte: Autoria própria.

Para cada locutor foram realizadas cinco repetições de cada comando na entrada de áudio no sistema. Os resultados foram analisados e postos na Tabela 12, com a quantidade de acertos para cada comando por locutor. O sistema foi selecionado para funcionar de forma ativa, ou seja, sem a necessidade de falar a palavra de ativação “Quarto”, otimizando assim, os testes.

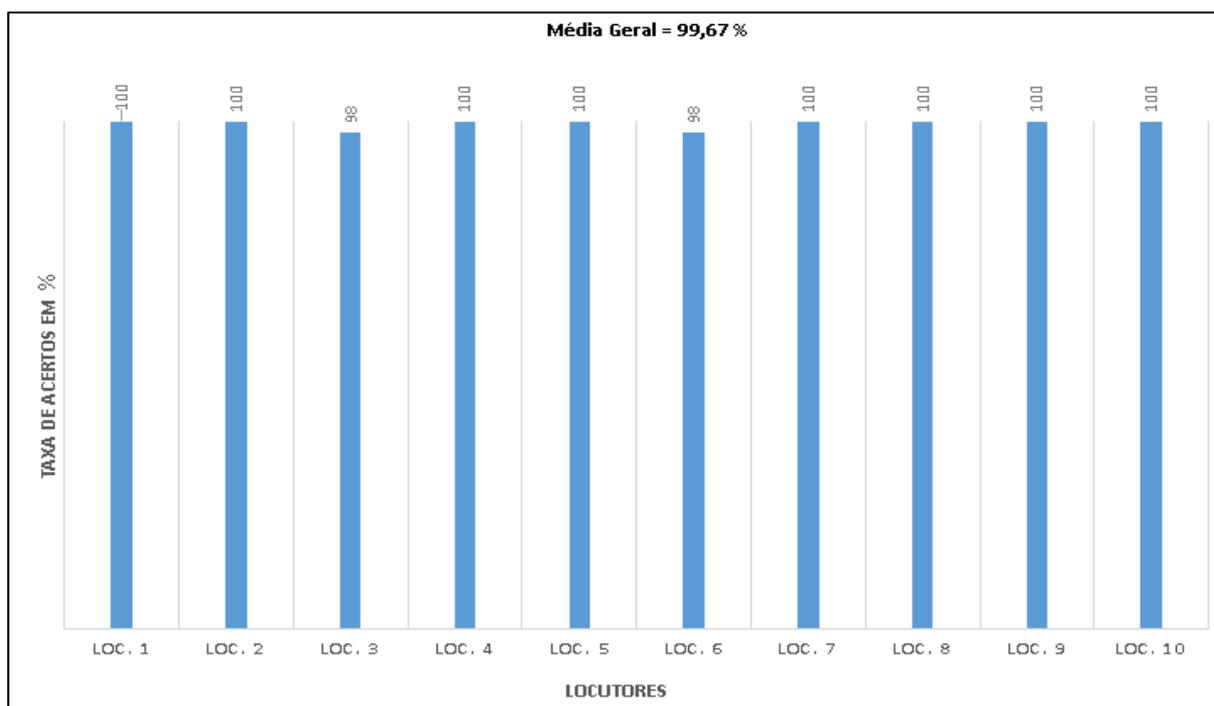
Tabela 12 - Número de acertos para cada locutor.

Comando		L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
Ligar	Lâmpada	5	5	5	5	5	5	5	5	5	5
Desligar	Lâmpada	5	5	5	5	5	5	5	5	5	5
Ligar	Ar condicionado	5	5	5	5	5	5	5	5	5	5
Desligar	Ar condicionado	5	5	4	5	5	5	5	5	5	5
Ligar	Abajur	5	5	5	5	5	5	5	5	5	5
Desligar	Abajur	5	5	5	5	5	5	5	5	5	5
Ligar	TV	5	5	5	5	5	5	5	5	5	5
Desligar	TV	5	5	5	5	5	5	5	5	5	5
Abrir	Persiana	5	5	5	5	5	4	5	5	5	5
Fechar	Persiana	5	5	5	5	5	5	5	5	5	5
Socorro		5	5	5	5	5	5	5	5	5	5
Encerrar		5	5	5	5	5	5	5	5	5	5

Fonte: Autoria própria.

Para o locutor 3 do sexo masculino, o sistema apresentou erro apenas em uma entrada de fala “ Desligar Ar condicionado”, e para a locutora 6 na entrada de fala “Abrir persiana”, que foi ocasionado pelo corte da palavra em ambos os casos. Ou seja, o comando não foi dito dentro do tempo estimado de escuta do sistema.

Os resultados com as taxas de acertos e média geral entre todos os locutores com o *Google speech* podem ser vistos na Figura 31.

Figura 31 - Taxa de acertos individual e média geral.

Fonte: Autoria própria

Foi gravado um teste de validação do sistema em português, tendo sido o vídeo disponibilizado na Internet, através do seguinte link no *YouTube*: https://youtu.be/Awd4_wdXRwo.

4.2.1 - Discussão

Para os testes com a segunda abordagem, o sistema se mostrou satisfatório ao desempenho esperado, obtendo uma média geral de acertos de 99,67%. Além dos testes inserindo as gravações das vozes no sistema, alguns testes foram realizados diretamente no sistema pelos voluntários. Foi observado que o sistema fazia o reconhecimento a uma distância superior a um metro, em um ambiente com baixo ruído.

Os comandos reconhecidos eram monitorados através do *notebook* que se comunicava com o *Raspberry Pi* por SSH (*Secure Shell*). Um exemplo pode ser visto na Figura 32.

O corte das palavras acontece esporadicamente, podendo possivelmente ser considerado erro no sistema e não no conversor STT, pois o mesmo problema foi observado na primeira abordagem com o uso do *Pocketsphinx*.

Para o tempo de escuta do sistema a duração foi de três segundos depois dos dois bips. Nos testes, por exemplo, o sistema finalizava a escuta com um bip antes dos três segundos. Acredita-se que pode ter sido ocasionado pela conexão com a Internet ou pelos dispositivos de *Hardware*, como o *SD Card* que apresenta baixa velocidade de tráfego de dados, já que o sistema trabalha de forma dinâmica trocando informações com ele, ou por conta de interferências da fonte de alimentação.

Figura 32 - Exemplo corte de palavra pelo sistema.

```
(u'LIGAR TV', u'LIGAR A TV', u'LIGAR TE VER')
(u'DESLIGAR TV', u'RESUME DA TV', u'DESLIGAR A TV', u'RESUMO
WARNING:client.stt:Resultado: Nenhuma palavra foi dita.
[]
WARNING:client.stt:Resultado: Nenhuma palavra foi dita.
[]
(u'DESLIGAR AR COND', u'DESLIGAR AR CONDI', u'DESLIGAR A COMI
WARNING:client.stt:Resultado: Nenhuma palavra foi dita.
[]
WARNING:client.stt:Resultado: Nenhuma palavra foi dita.
[]
^Z
[1]+  Stopped                               /home/pi/jasper/jasper3.py
pi@raspberrypi ~ $
```

Fonte: Autoria própria.

5. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

Tecnologias assistivas voltadas para pessoas com algum tipo de mobilidade reduzida ou deficiência, visam a minimização das dificuldades de acesso encontradas, seja facilitando ou até mesmo realizando alguma tarefa cotidiana, trazendo mais segurança e conforto por meio de recursos específicos.

O uso da automação via voz, proposta desse trabalho, se mostra válido, por geralmente se utilizar de dispositivos de baixo custo e que se mostram eficientes, o que os tornam mais acessíveis e seguras às pessoas e proporcionam uma interação de forma prática e rápida.

Com os testes preliminares com o sistema utilizando o *PocketSphinx STT engine* e a plataforma embarcada *Raspberry Pi B+*, comprovou-se a viabilidade de se implementar uma proposta de interface por interação via voz, de baixo custo e eficiente. A grande facilidade de manipulação da plataforma embarcada *Raspberry Pi* e do *software* torna fácil a adaptação e o controle dos mais diversos dispositivos externos, como eletrodomésticos e a parte elétrica do ambiente.

A partir disso, a implementação de uma segunda proposta para o mesmo sistema na língua portuguesa e controle dos dispositivos via Wi-Fi foi testada. A substituição da plataforma *Raspberry Pi B+* por outra da mesma família *Raspberry Pi Zero W* tornou o protótipo 58,10% mais barato. O reconhecimento, por ser na língua do usuário, melhorou consideravelmente a taxa de acertos para 99,67%, em relação aos 90,4% obtidos na primeira abordagem na língua inglesa.

Em termos de valores, a proposta apresentada nesse trabalho se tornou mais acessível em comparação com o sistema proposto por Pimentel (2014), onde houve uma diminuição em 55,62% no custo e, ainda, uma redução de 81,23% em relação à proposta de Hall (2003).

5.1. Contribuições do trabalho

O seguinte trabalho além de apresentar uma metodologia barata e simples de um dispositivo de reconhecimento automático de fala, teve como propósito:

- Estimular o estudo e aperfeiçoamento de sistemas e *software* de assistência social já existentes, acrescentando recursos que otimizem e sejam utilizados com outros propósitos.

- Apresentar outras alternativas em termos de *hardware* de baixo custo para o controle dos dispositivos por Wi-Fi, trazendo maior segurança ao usuário e flexibilidade em adaptar o sistema de controle em outros ambientes.
- A opção de se utilizar o API *speech* do Google para o reconhecimento automático de fala, por apresentar uma taxa de reconhecimento elevada, a não necessidade de um pré-processamento antes e a flexibilidade de troca de idioma.
- A possibilidade de comunicação com outras pessoas ou serviços de emergências através de mensagens SMS ou chamadas de voz.
- Análise dos motores STT.

5.2. Sugestões para trabalhos futuros

Além do sistema se mostrar acessível financeiramente, apresentar uma boa eficiência e eficácia, futuramente poderiam ser vistas outras possibilidades de melhoria como:

- O Reconhecimento de fala também no modo *off-line*, o que se tornaria mais acessível até mesmo para pessoas que não tem acesso a uma rede de Internet;
- A criação de uma interface com *display* com botões onde o idioma possa ser selecionado através de um menu, possibilitando o uso por pessoas de outras nacionalidades e acréscimo ou mudança de comandos do sistema;
- Interface de conexão com o roteador sem a necessidade de entrar no sistema;
- Ligações para os principais serviços de emergência.
- Criação de um aplicativo para *smartphone* como alternativa de controle das cargas, ampliando assim a funcionalidade para outras pessoas da casa.

6. REFERENCIAIS BIBLIOGRÁFICAS

ABREU, E.R.; VALIM P.R.O. **Domótica, controle de automação residencial utilizando celulares bluetooth**. In: VIII SEGet - Simpósio de excelência em gestão e tecnologia, 2011.

ANUSHA, G; Ballala, S. H; NARESH, S. **Low Cost Home Automation by Zigbee and Voice Commands using Raspberry pi2 (B+)**. Dept of ECE, Sri Indu Institute of Engineering & Technology, Hyderabad, TS, India, International Journal of Advanced Technology in and Innovative Research. ISSN 2348-2370, August-2016. Vol.08, Issue.10. Pages:1977-1980.

ÀVILA, C. **Como funciona o reconhecimento de voz do Google?**. Mybq. Disponível em: <http://www.mibqyyo.com/pt-artigos/2016/01/26/como-funciona-reconhecimento-voz-google>>. Acesso em: 01 jan. 2018.

ARAÚJO; T. **Raspberry Pi B+: Introdução a Porta GPIO**. Disponível em: <<http://blog.fazedores.com/raspberry-pi-b-introducao-porta-gpio/>>. Acesso em: 30 de jan. 2018.

AWS. **Documentação do Alexa for Business**.2017. Disponível em: <https://aws.amazon.com/pt/documentation/a4b/>. Acesso em: 6 de Mar. 2018.

BARROS, M.D. da S.; IANHAS, E.R.F. **Domótica-Automação Residencial por comando de Voz**. 2014. Disponível em: <<http://conic-semesp.org.br/anais/files/2014/trabalho-1000017298.pdf>>. Acesso em: 7 de Jun. 2017.

BEHLAU, M. **Voz, O Livro do Especialista**. E. Revinter. Ed., vol. I, 2001.

BERSCH, R. de C. R. **Design de um serviço de tecnologia assistiva em escolas públicas**. Porto Alegre, Universidade Federal do Rio Grande do Sul. Dissertação de mestrado, 2009, 231fl.

BERSCH, R.; MORAES, H.; PASSERINO, L. M.; BATISTA, V. **Tecnologia Assistiva e design na realidade Brasileira**. Anais do 3º Workshop Design & materiais da UFRGS. Porto Alegre, RS, 2007.

BOLZANI, C.A.M. **Desenvolvimento de simulador de controle de dispositivos residenciais inteligentes: uma introdução aos sistemas domóticos**. Ed. rev., São Paulo, 2004.

BORGES, A.K.; BURIOLA, J.F.L.; ELOI, J.; TELES, R.F. **SmartChair: Cadeira de rodas controlada por voz**. 2015. Disponível em: <paginapessoal.utfpr.edu.br/gustavobborba/if66j-s71.../IF66J-15a_RT_SmartChair.pdf>. Acesso em: 7 de Jun. de 2017.

BRASIL. O Acesso de Alunos com Deficiência às Escolas e Classes Comuns da Rede Regular. Ministério Público Federal: Fundação Procurador Pedro Jorge de Melo e Silva (organizadores) / 2ª ed. rev. e atualiz. Brasília: Procuradoria Federal dos Direitos do Cidadão, 2004, p. 13.

CHARLES, M.; SHUBHRO, S. **Jasper documentação**, 2014. Disponível em: <<http://jasperproject.github.io/documentation/>>. Acesso em: 15 de ago. 2017.

CHIELE, M. R.; ZERBETTO, A. **Desenvolvimento de uma Cadeira de Rodas Controlada por Voz.** ResearchGate, 2010. Disponível em: <https://www.researchgate.net/publication/266294187_DESENVOLVIMENTO_DE_UMA_CADEIRA_DE_RODAS_CONTROLADA_POR_VOZ>. Acesso em: 7 de Jun. de 2017.

DESIGNSOFT. Disponível em: <https://www.tina.com/Portuguese/tinacloud/document>. Acesso em: 6 de Mar. 2018.

DÍAZ, OSCAR MORENO; PARDO, Aller, J.L.; PULIDO, Ivan, M. **Domótica – Instalación y simulación.** Disponível em <http://bibliotecna.upc.es/bustia/arxiu/40441.pdf>. Acesso em: 03 de jan. 2018.

GOBINATHAN, P; NARAYANADASS, R. **Low Cost Home Automation Using Offline Speech Recognition**, International Journal of Signal Processing Systems, vol. 2, Dec, 2014.

GOOGLE CLOUD PLATFORM. **Cloud speech api.** Disponível em: <<https://cloud.google.com/speech/>>. Acesso em: 20 jan. 2018.

GOULART, D; COOPER, M. **Por todo o canto: Método vocal.** 1ª ed. São Paulo: G4, 2002. 85 p. ISBN: 8590170926.

GREGORIN, P. **Segurança: Biometrias e a Digitalização da Identidade.** Oz technology, 2017. Disponível em: <http://www.epoc.com.br/blog/tecnologia/seguranca-biometrias-e-a-digitalizacao-da-identidade/>. Acesso em: 19 jan. 2018.

GUCCIONE, A. A. **Avaliação Funcional do Idoso.** em: Guccione AA. Fisioterapia Geriátrica. 2ª ed. Rio de Janeiro: Guanabara Koogan, 2002. 480 p. ISBN: 9788527723442.

GUEDES, L.; ALVARENGA, L.; ROMANINI, A.; MARTINS, M. S. FOLLE, D. **O papel social da automação – automação inclusiva e mais sustentável.** Disponível em: <https://www.imed.edu.br/.../O%20papel%20social%20da%20automação_Automação...> Acesso em: 7 de Jun. 2017.

GUSTAVO, D. **Gravação de Voz – Posição e Distância do Microfone.** ArtSom Studio, 2013. Disponível em: <http://www.artsomstudio.com.br/blog/gravacao-de-voz-posicao-e-distancia-do-microfone>. Acesso em: 18 de jan. 2018

HALL B, MOLLOY J. **Installing A Voice Activated Environmental Control Unit For Under 500 Dollars**. RESNA 26th International Annual Conference [internet]. 2003. Disponível em: <http://web.resna.org/conference/proceedings/2003/Papers/EA/>. Acesso em: 23 de Jan. 2018.

IBGE. **Tábua completa de mortalidade para o Brasil**. 2015. Breve análise da evolução da mortalidade no Brasil, 2016. Disponível em: ftp://ftp.ibge.gov.br/Tabuas_Completas_de_Mortalidade/Tabuas_Completas_de_Mortalidade_2015/tabua_de_mortalidade_analise.pdf. Acesso em: 8 de Mar. 2018.

IPODSCHOOL. **Siri finalmente fala português**. 2015. Disponível em: <http://ipodschool.com/siri-finalmente-fala-portugues/>. Acesso em: 6 de Mar. 2018.

MUKESH, K; SHIMI S.L. **“Voice Recognition Based Home Automation System for Paralyzed People”**. International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4, Issue 10, October 2015.

KRONBAUER, A.H.; COSTA, V.F. **A New Approach to the Interpretation of Voice Commands**. Proceedings of the 21st Brazilian Symposium on Multimedia and the Web, Manaus, Brasil, Out. 2015.

LIMA, D.S; MENDES, J.C de F.; MANFIO, E.R. **Acionamento eletrônico de portão via comando por voz**. Disponível em: www.fatecgarca.edu.br/uploads/ict/artigos/edital02/danilojean.pdf >. Acesso em: 7 de Jun. 2017.

LIMA, T. **Raspberry Pi B+**. Embarcados, 2014. Disponível em: <https://www.embarcados.com.br/raspberry-pi-modelo-b-plus/>>. Acesso em: 20 ago. 2017

LINGARIA, D. M. **Assistive Voicer Recognition Device for GSM Calling Using Arduino UNO**. Rizvi College Of Engineering, Mumbai, India, 2015.

LINS, V; MOURA, W. **Domótica: Automação Residencial**. 2010. Disponível em: www.unibratec.edu.br/tecnologus/wp-content/uploads/2010/12/lins_moura.pdf. Acesso em: 02 jan. 2018.

MENDES, R. B.; TIBÚRCIO, T. M. de S. **Cadeira de rodas manual: o impacto dos problemas relacionados a tecnologia assistiva na melhoria da qualidade de vida de pessoas com mobilidade reduzida**. Artefactum – revista de estudos em linguagem e tecnologia ano VIII – nº 01/2016.

MICROSOFT. **Centro de desenvolvimento Cortana.2017**. Disponível em: <https://developer.microsoft.com/pt-br/cortana>. Acesso em: 06 de Mar. 2018.

MINATEL, P. **ESP8266: O guia básico de Hardware**. Disponível em: < <http://pedrominate.com.br/pt/esp8266/esp8266-o-guia-basico-de-hardware/>>. Acesso em: 30 de jan. 2018

MONTERIO, L; FERREIRA, N. **Perturbações das Cordas Vocais**. HARVARD MEDICAL SCHOOL – PORTUGAL PROGRAMA in translation research and information, 2011. Disponível em:< <https://hmsportugal.wordpress.com/2011/09/09/perturbacoes-das-cordas-vocais-2/>>. Acesso em: 31 de jan. 2018.

MORAES, C. e CASTRUCCI, P. **Engenharia de Automação Industrial**. 2ª ed. São Paulo: LTC, 2007. 360 p. ISBN: 9788521615323.

MURATORI, J. R; BÓ, P. H. D. **Automação residencial: histórico, definições e conceitos**. Disponível em:< http://www.instalacoeseltricas.com/download/Automacao_residencial1.pdf >. Acesso em: 17 de ago. de 2017.

NIELSEN, J. **Usability inspection methods**. In Conference companion on Human factors in computing systems. ACM: 413-414, 1994.

TERUEL, E. C. **Uma proposta de framework para sistemas de automação residencial com interface para WEB**. 2008. 158 f. Dissertação (Mestrado em Tecnologia: Gestão, Desenvolvimento e Formação) – Centro Estadual de Educação Tecnológica Paula Souza, São Paulo, 2008, 241 fl.

OBAID, T. “**Zigbee Based Voice Controlled Wireless Smart Home System**” International Journal of Wireless & Mobile Networks (IJWMN), Vol. 6, No. 1, February 2014.

OLIVEIRA, M.V.S.; OTTOBONI, R. **Projeto de automação residencial: Acionamento por comando de voz**. Centro universitário de Itajubá – FEPI, 2015.

PEREIRA, A. P. **Como funciona o reconhecimento de voz?** TECMUNDO. Disponível em: < <https://www.tecmundo.com.br/curiosidade/3144-como-funciona-o-reconhecimento-de-voz-.htm>>. Acesso em: 30 de jan. 2018.

PERICO, A.; SHINOHARA, C.S.; SARMENTO, C.D. **Sistema de reconhecimento de Voz para Automatização de uma Plataforma Elevatória**. Universidade Tecnologia Federal do Paraná-UTFPR, Trabalho de Conclusão de Curso, 2014, 96 fl.

PILOTI, J. S. **Sistema de automação residencial: acessibilidade no controle doméstico**. Universidade de Caxias do Sul, Trabalho de conclusão de curso, 2014, 96 fl.

PIMENTA, S. E. D. **Mobilidade reduzida dos idosos no domicílio – condições determinantes**. Escola Superior de Enfermagem do Porto, Dissertação de mestrado, 2016, 129 fl.

PIMENTEL, V.C. A.; BARBACENA, I.L.; CORREIA, S.E.N.; DIAS, M.C. **Uma plataforma de baixo custo comandada por voz para tecnologias assistivas com programação em python**. In: Congresso Brasileiro de Engenharia Biomédica, XXIV, MG. Anais..., Uberlândia: Universidade Federal de Uberlândia, 2014, pag. 1414-1417.

PORTUGAL. Secretariado Nacional de Reabilitação e Integração da Pessoa com Deficiência. Disponível em: <http://www.snripd.pt/default.aspx?IdLang=1> Acesso em: out. 2017.

PRADO, A. D. A.; LOPES, M. E.; ORNSTEIN, S. W. **Desenho universal - Caminhos da Acessibilidade no Brasil**. 1ª ed. São Paulo: Annablume, 2010. 306 p. ISBN: 978-8539100552

RAMLI, T.; DABIMEL, N. N.; MAMAT M.; PARIMON, N.; ROSALYN R. Porl,2016. **Simple speech controlled home automation system using android devices**. Journal of Scientific Research and Development 3 (1): 33-38 , ISSN 1115-7569.

PRASANNA, G.; RAMADASS, N.; "Low Cost Home Automation Using Offline Speech Recognition", International Journal of Signal Processing Systems Vol. 2, No. 2, December 2014.

RICHARDSON, M.; WALLACE, S. **Primeiros passos com o Raspberry Pi**. [tradução Patricia Tieme Maeda]. -- São Paulo: Novatec Editora, 2013.

ROCKENBACH, S. **Arquitetura, Automação e Sustentabilidade**. Cap. 5. Universidade Federal do Rio Grande do Sul – UFRGS. Programa de Pesquisa e Pós-Graduação em Arquitetura, Porto Alegre, 2005, 193 fl.

ROQUE, A. **Introdução a domótica**. Publicado na Revista O Electricista, nº1 Jul, Agos e Set de 2002. Disponível em: <<http://www.antonioroque.com/textos.asp?idCat=11&idArtigo=12>>. Acesso em: 03 fev. 2018.

SANT'ANA, D.G. **Controle por Voz utilizando a Engine Julius com Fala Contínua**. Universidade Regional de Blumenau-FURB, Trabalho de Conclusão de Curso, 2014, 64 fl.

SALES, R. IBGE: **Expectativa de vida dos brasileiros aumenta para 75,5 anos**. G1, Brasil, 2016. Disponível em: <<http://www.valor.com.br/brasil/4793049/expectativa-de-vida-dos-brasileiros-aumenta-para-755-anos>>. Acesso em: 25 ago. 2017.

SARTORETTO, M. L.; BERSCH, R. **Categorias de Tecnologia Assistiva**. Assistiva Tecnologia e Educação, 2014. Disponível em: <<http://www.assistiva.com.br/tassistiva.html#categorias>>. Acesso em: 14de Jun. 2017.

SASSAKI, R. K. **Como chamar as pessoas com deficiência. Vida independente: história, movimento, liderança, conceito, filosofia e fundamentos**. São Paulo: RNR, 2003. 12-16 p.

SILVA, C. P. A. **Um software de reconhecimento de voz para o português brasileiro.** Dissertação (Mestrado em Engenharia Elétrica) – Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2010.

SILVA, R. J.; SANTOS, L. F. L.; NASCIMENTO, D. M. M. L.; AMARAL, T. M.; NASCIMENTO, J. M. L. **Inclusão de estudantes com mobilidade reduzida na rede educacional do município de paulo afonso – ba.** REVASF, Petrolina, PE, vol. 6, n. 10, p. 27-38, jul. 2016.

SILVEIRA, D.P. **O que é tecnologia assistiva?.** Oficina da NET, 2016. Disponível em: <https://www.oficinadanet.com.br/post/17558-o-que-e-tecnologia-assistiva>. Acesso em: 6 de Mar. 2018.

SOUSA, M.S. de; GOMES ,R.M ; GUEDES, J. M. **Voz humana e comunicação.** Disponível em: < www.fumec.br/files/6814/3337/6578/Voz_humana-v7.pdf>. Acesso em: 24 de Jan. 2018.

SURESH, S.; RAO, Y. S. **Modelling Of Secured Voice Recognition Based Automatic Control System.** International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE) ISSN: 0976-1353 Volume 13 Issue 2 –MARCH 2015.

TECMUNDO. Google Now. 2016. Disponível em: <https://www.tecmundo.com.br/google-now>. Acesso em: 6 de Mar. 2018.

TEZA, V. R. **Alguns aspectos sobre automação residencial – Domótica.** Tese de Mestrado. UFSC, Santa Catarina, 2002, 108 fl.

USINAINFO. **Módulo de reconhecimento de voz e comando para Arduino.** Disponível em: <http://blog.usinainfo.com.br/modulo-de-reconhecimento-de-voz-para-arduino/>. Acesso em: 19 de jan. 2017.

VIEIRA, R.; LIMA, V. L. S. **Linguística computacional: princípios e aplicações.** In: IX Escola de Informática da SBC-Sul. Luciana Nedel (Ed.) Passo Fundo, Maringá, São José. SBC-Sul, 2001.

ZHANG, Y.; FANEURFF.J.J; HIDDEN, W.; HOTARY, J.T.; LEE, S.C.; IYENGAR, V. **Automobile speech-recognition interface.** US Pat. 7826945 B2, 02 nov. 2015.

APÊNDICES

APÊNDICE A – GUIA DE GRAVAÇÃO DO SISTEMA OPERACIONAL NO CARTÃO SD E CONFIGURAÇÕES DO *RASPBERRY PI*.

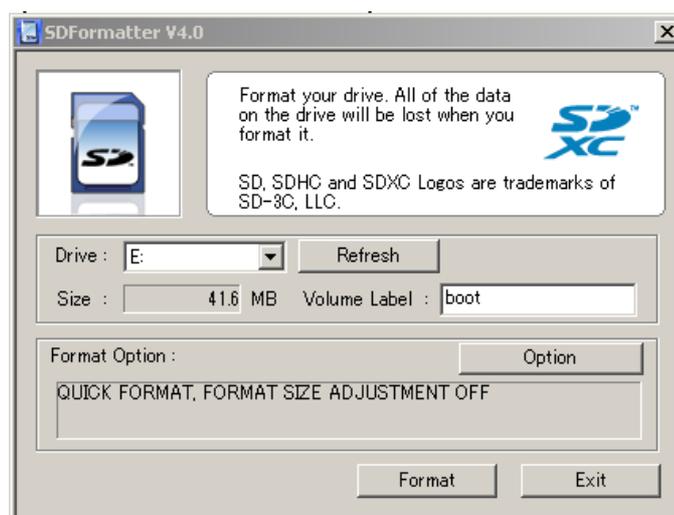
Neste guia será apresentada a forma adequada de formatação e gravação do sistema operacional *Raspbian Wheezy* em um cartão *SD Card*. Onde se faz necessário o uso de dois *softwares*. Observando-se que para esta aplicação o cartão SD deve possuir no mínimo 4 GB de memória. O sistema operacional utilizado pode ser baixado no seguinte *link*: <https://sourceforge.net/projects/raspbian-wheezy-pi/>.

A.1 – Formatação (*SD Formatter*)

O cartão precisa ser preparado para receber a imagem. ISO de forma correta. Deve ser formatado utilizando o *software SD Formatter* disponível no seguinte link <https://sd-card-formatter.br.uptodown.com/>.

Após a inicialização do *software* (Figura 33), o sistema deve reconhecer automaticamente o cartão SD, o mesmo já deve ter sido conectado ao PC através de um adaptador.

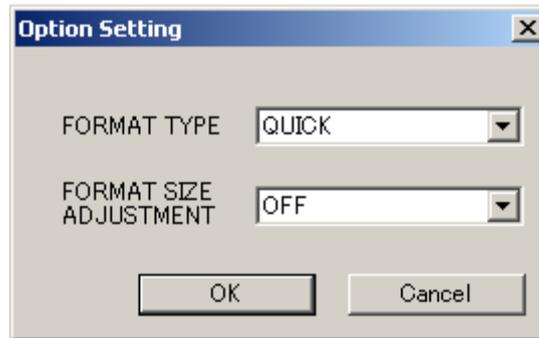
Figura 33 - Janela principal *SDFormatter*.



Fonte: Autoria própria.

Em “*Option*” devem ser setadas as seguintes configurações apresentadas na Figura 34. Após a confirmação “*OK*” e “*Format*” a formatação do cartão *SD Card* deve começar automaticamente.

Figura 34 - Configurações *SDFormatter*.

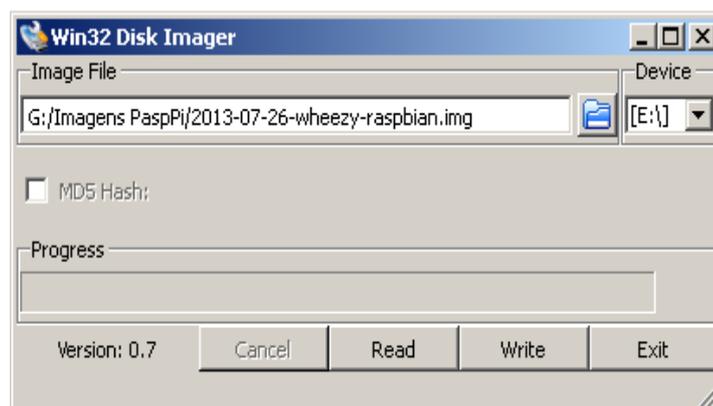


Fonte: Autoria própria.

A.2 - Gravação (*Win32DiskImager*)

Para a gravação da imagem *.ISO* se faz necessário a utilização de um *software* específico, que neste trabalho foi utilizado o “*Disk Imager*” disponível no seguinte endereço <https://sourceforge.net/projects/win32diskimager/>. Após inicialização do *software* (Figura 35), em “*Image File*”, deve ser posto o caminho do arquivo. Em “*device*” certificar-se que o *SD Card* está selecionado corretamente, em seguida, pressionando “*Write*” o arquivo será copiado para o cartão.

Figura 35 - Janela principal *Disk Imager*



Fonte: Autoria própria.

Seguido os passos acima, após todas as etapas concluídas com sucesso, o cartão SD pode ser ejetado do PC e inserido no *Raspberry Pi*. Prosseguindo com algumas configurações iniciais antes da instalação do *software Jasper*.

A.3 – Configurações iniciais *Raspbian Wheezy*

Nesta etapa serão apresentadas as linhas de comandos utilizadas no terminal do *Raspbian* para a expansão de arquivos, e a configuração da rede Wi-Fi, que apresenta uma certa particularidade entre os outros sistemas operacionais do *Raspberry Pi*.

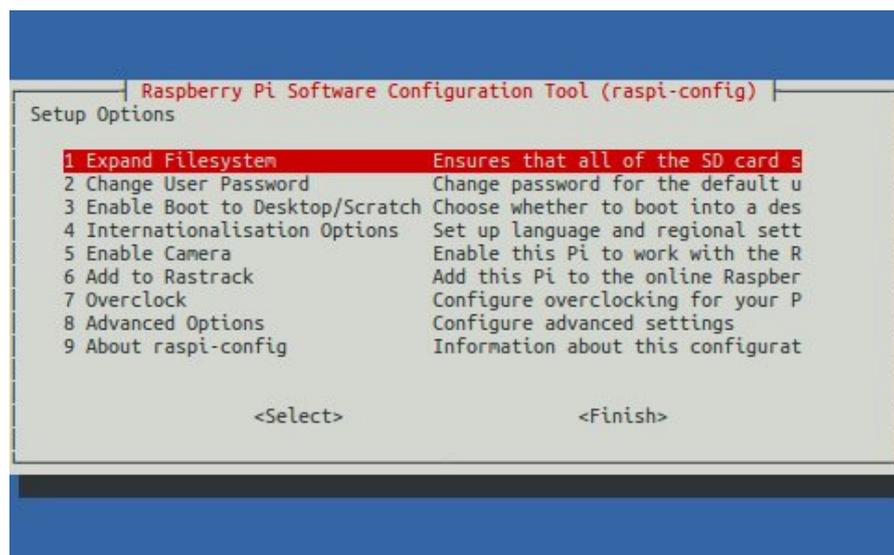
A.3.1 – Expandindo o sistema de arquivos

O sistema de arquivos deve ser expandido, pois a partição de dados inicialmente não ocupa todo o cartão de memória. Para fazer essa expansão é utilizado o seguinte comando no terminal:

```
$ sudo raspi-config
```

A seguinte janela (Figura 36) deve aparecer:

Figura 36 - Janela de configurações do *Raspbian*.

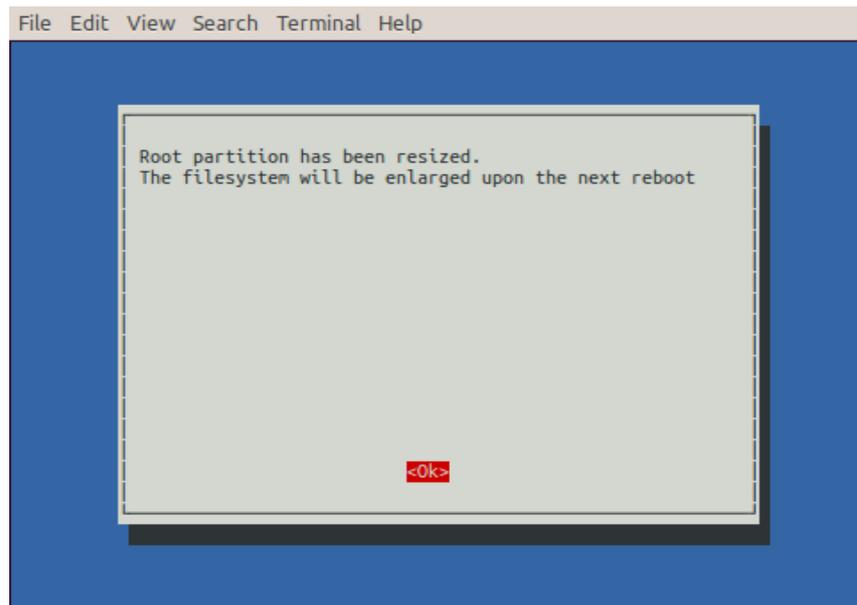


Fonte: Jabuti.edu

A primeira opção deve ser escolhida e confirmada. Uma segunda janela (Figura 37), após o “OK” o sistema será expandido. Ao término é necessário que o *Raspberry Pi* seja reinicializado para que sejam feitas as novas configurações, digitando o comando:

```
$ sudo reboot
```

Figura 37 - Janela de confirmação

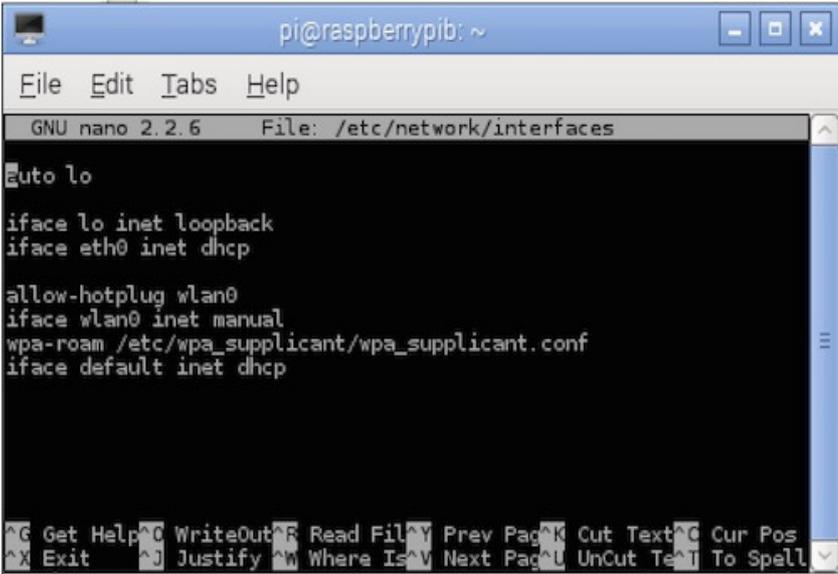


Fonte: Jabuti.edu

A.3.1 – Configurando a rede Wi-Fi

Para o uso de um adaptador Wi-Fi, é preciso habilitar o “eth0”, pois por padrão o *Raspberry Pi* usa DHCP para configurar suas interfaces de rede. Deve ser editado ou inserido no arquivo (Figura 38), após inserção da seguinte linha de comando no terminal:

```
$ sudo nano /etc/network/interfaces
```

Figura 38 - Interface *network Raspberry Pi*

```
pi@raspberrypi: ~  
File Edit Tabs Help  
GNU nano 2.2.6 File: /etc/network/interfaces  
auto lo  
  
iface lo inet loopback  
iface eth0 inet dhcp  
  
allow-hotplug wlan0  
iface wlan0 inet manual  
wpa-roam /etc/wpa_supplicant/wpa_supplicant.conf  
iface default inet dhcp  
  
^G Get Help ^O WriteOut ^R Read File ^Y Prev Page ^K Cut Text ^C Cur Pos  
^X Exit ^J Justify ^W Where Is ^V Next Page ^L UnCut Text ^T To Spell
```

Fonte: TechBlog.com

Em seguida as alterações devem ser salvas pressionando “ctrl + O” e enter, para sair da interface “ctrl + X” e enter. Para uso de dados de rede através do conector RJ45 não é necessário fazer essas configurações de Internet.

APÊNDICE B – GUIA DE INSTALAÇÃO DO SOFTWARE JASPER.

Inicialmente se faz necessário alguns documentos e configurações para o funcionamento correto do *software*:

B.1 – Atualização do sistema e configurações ALSA.

Para atualização do sistema operacional:

```
$ sudo apt-get update
$ sudo apt-get upgrade --yes
```

Documentos para dispositivos de áudio:

```
$ sudo apt-get install vim git-core python-dev python-pip bison libasound2-dev
libnortaudio-dev nvthon-nvaudio --yes
```

Instalação dos plug-ins de áudio ALSA (configuração do microfone):

```
$ sudo nano /etc/modprobe.d/alsa-base.conf
```

O arquivo deve ser modificado para a seguinte configuração:

```
# Load USB audio before the internal soundcard
options snd_usb_audio index=0
options snd_bcm2835 index=1

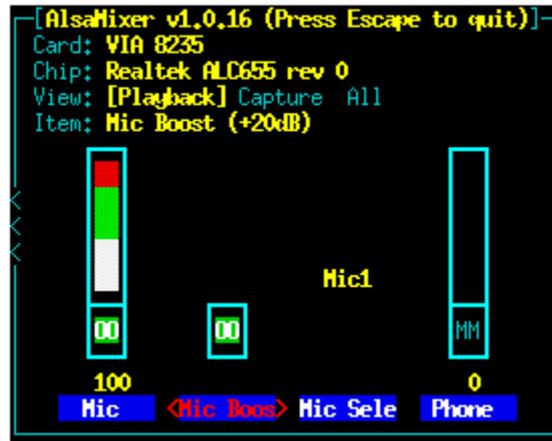
# Make sure the sound cards are ordered the correct way in ALSA
options snd slots=snd_usb_audio,snd_bcm2835

options snd-usb-audio index=-2
```

Com isso o sistema reconhece o *drive* de áudio externo USB (Microfone), como principal, a qualquer momento no terminal pode se ajustar a sensibilidade do microfone (Figura 39) utilizando o seguinte comando:

```
$ alsamixer
```

Figura 39 - Painel de controle ALSA



Fonte: Alsaproject

Na pasta:

```
$ sudo nano ~/.bash_profile
```

Deve ser acrescentado no final da linha:

```
export LD_LIBRARY_PATH="/usr/local/lib"
source .bashrc
```

Em:

```
$ sudo nano ~/.bashrc
```

Deve ser acrescentado:

```
LD_LIBRARY_PATH="/usr/local/lib"
export LD_LIBRARY_PATH
PATH=$PATH:/usr/local/lib/
export PATH
```

B.2 – Instalação Jasper e STT

Para esse projeto foi seguida a documentação do *software* para sua instalação que pode ser vista no seguinte link: <http://jasperproject.github.io/documentation/>. Só alguns arquivos foram necessários para seu pleno funcionamento, já que diversos recursos oferecidos não foram utilizados nessa proposta.

Comando para instalação do *software Jasper*:

```
$ git clone https://github.com/jasperproject/jasper-client.git jasper
$ sudo pip install --upgrade setuptools
$ sudo pip install -r jasper/cliente/requiriments.txt
```

Em seguida o executável:

```
$ chmod +x jasper/jasper.py
$ cd ~/jasper
$ sudo apt-get update
$ sudo apt-get install python-pymad
```

As configurações de definição do STT são editadas conforme desejado no seguinte arquivo:

```
$ sudo nano /home/pi/.jasper/profile.yml
```

Para o idioma inglês utilizando o *Pocketsphinx* segui as seguintes configurações:

```
stt engine: sphinx
```

Para o uso do *Pocketsphinx* se faz necessária a instalação de outros arquivos, que demora algum tempo por conta da compilação de alguns arquivos:

```
$ yaourt jasper -stt -pocketsphinx
```

Instalando o *Pocketsphinx*:

```
$ sudo apt-get update
$ sudo apt-get install pocketsphinx python-pocketsphinx
```

Fazendo a compilação e instalando a partir da fonte:

```
$ wget http://downloads.sourceforge.net/project/cmusphinx/sphinxbase/0.8/sphinxbase-0.8.tar.gz
tar -zxvf sphinxbase-0.8.tar.gz
cd ~/sphinxbase-0.8/
./configure --enable-fixed
make
```

```

$ sudo make install

wget
http://downloads.sourceforge.net/project/cmuspinx/pocketsphinx/0.8/pocketsphinx-
0.8.tar.gz

tar -zxvf pocketsphinx-0.8.tar.gz

cd ~/pocketsphinx-0.8/

./configure

make

$ sudo make install

cd ..

$ sudo easy_install pocketsphinx

```

Instalar algumas dependências:

```
$ sudo apt-get install subversion autoconf libtool automake gfortran g++ --yes
```

Instalando o CMUCLMTK:

```

$ svn co https://svn.code.sf.net/p/cmuspinx/code/trunk/cmuclmtk/

cd cmuclmtk/

./autogen.sh && make && sudo make install

cd ..

```

Instalando o MIT Language Modeling Toolkit, m2-aligner e *Phonetisaurus*:

```

$ wget http://distfiles.macports.org/openfst/openfst-1.3.4.tar.gz

wget https://github.com/mitlm/mitlm/releases/download/v0.4.1/mitlm_0.4.1.tar.gz

wget https://storage.googleapis.com/google-code-archive-
downloads/v2/code.google.com/m2m-aligner/m2m-aligner-1.2.tar.gz

wget https://storage.googleapis.com/google-code-archive-
downloads/v2/code.google.com/phonetisaurus/is2013-conversion.tgz

```

Após, se faz necessário desmarcar alguns downloads:

```
$ wget http:tar -xvf m2m-aligner-1.2.tar.gz
tar -xvf openfst-1.3.4.tar.gz
tar -xvf is2013-conversion.tgz
tar -xvf mitlm_0.4.1.tar.gz
```

Criando o OpenFST:

```
$ cd openfst-1.3.4/
sudo ./configure --enable-compact-fsts --enable-const-fsts --enable-far --enable-
lookahead-fsts --enable-pdt
sudo make install
```

Construindo o M2M e MITLMT:

```
$ cd m2m-aligner-1.2/
sudo make
cd mitlm-0.4.1/
sudo ./configure
sudo make install
```

Construindo o *Phonetisaurus*:

```
$ cd is2013-conversion/phonetisaurus/src
sudo make
```

Movendo os arquivos compilados:

```
$ sudo cp ~/m2m-aligner-1.2/m2m-aligner /usr/local/bin/m2m-aligner
sudo cp ~/is2013-conversion/bin/phonetisaurus-g2p /usr/local/bin/phonetisaurus-g2p
```

Construindo o modelo *Phonetisaurus* FST:

```
$ wget https://www.dropbox.com/s/kfht75czdwucni1/g014b2b.tgz
tar -xvf g014b2b.tgz
cd g014b2b/
./compile-fst.sh
cd ..
mv ~/g014b2b ~/phonetisaurus
```

Para o idioma português utilizando o *Google speech* segui as seguintes configurações:

```
stt_engine: google
keys:
  GOOGLE_SPEECH: "AlzaSyAHTEfKdg7IFNHaYC3Spw59tKszCKuYQLI"
```

Para o acesso a uma chave do *API Google Speech* devem ser seguidos os seguintes passos:

1º Criar uma conta com um e-mail válido no grupo Chromium dev no seguinte endereço: <https://groups.google.com/a/chromium.org/forum/?fromgroups#!forum/chromium-dev>.

2º Criar um projeto através do console do Google *Developers*: <https://console.developers.google.com/project>.

3º Em APIs e serviços > painel deve ser criado um projeto e ativado o API de reconhecimento de voz “*Google Cloud Speech API*”, a chave deve ser adquirida em credenciais e inseridas no arquivo *profile.yml*.

Após feita as configurações necessária para utilização do Google speech o idioma deve ser setado no arquivo *stt.py*:

```
$ sudo nano /home/pi/jasper/client/stt.py
```

Para a língua portuguesa, em “*language*” deve ser posto as iniciais do idioma e o país:

```
def __init__(self, api_key=None, language='pt-br'):
```

Se todos os procedimentos acima tiverem sido realizados com sucesso, o sistema de reconhecimento de fala pode ser iniciado com a seguinte linha de comando:

```
$ /home/pi/jasper/jasper.py
```

O sistema pode ser iniciado automaticamente ao ligar o *Raspberry Pi*, fazendo com que o usuário não precise entrar no sistema para ativar o programa de reconhecimento, através da seguinte linha de comando:

```
$ crontab -e
```

Acrescentando o seguinte comando:

```
@reboot python /home/pi/jasper/jasper.py;
```

Depois de salvo, o sistema passará a funcionar sem intermédio do usuário. Para isso é preciso estar conectado a uma fonte de Internet, seja por Wi-Fi ou cabo de rede através do conector RJ45 e o microfone deve estar conectado a uma entrada USB.

ANEXOS

ANEXO A – TERMO DE COMPROMISSO LIVRE E ESTABELECIDO (TCLE).

Termo de Compromisso Livre e Estabelecido (TCLE)

Prezado participante,

Você está sendo convidado a participar da pesquisa **Automação Residencial por Voz para Pessoas com Mobilidade Reduzida**, desenvolvida por **Náiron de Vasconcelos Sousa**, discente do Curso de Pós-Graduação *Strictu Sensu* de Mestrado em Engenharia Elétrica do Instituto federal de Educação, Ciência e Tecnologia da Paraíba – IFPB, sob orientação da Professora Dra. **Suzete Elida Nóbrega Correia**.

O objetivo central do estudo é: Desenvolver uma plataforma de reconhecimento da fala para o auxílio de pessoas com mobilidade reduzida para o controle de dispositivos do ambiente doméstico. Os objetivos específicos são: Realizar estudo sobre tecnologias eficientes e de baixo custo que se mostrem viáveis ao desenvolvimento de interfaces de comandos por voz para a interação humano-computador. Elaborar concepção de interface de baixo custo, portátil e de uso pessoal, para aplicações de comandos através do reconhecimento de voz em tecnologias assistivas que atuem no controle de dispositivos domésticos. Desenvolver recursos de base de texto (gramática) e base de áudio para a criação de modelo acústico específico para aplicações de comando e controle em domotica. Avaliar a interface concebida com fins de validar sua eficiência.

O motivo de sua participação se deve ao fato de você ser cidadão brasileiro, com preservadas boas capacidades da fala e idade igual ou superior a 18 anos. Sua participação é voluntária e você tem plena autonomia para decidir se quer ou não participar, bem como retificar sua participação a qualquer momento. Você não será penalizado de nenhuma maneira caso decida não consentir sua participação, ou desistir do seu consentimento. Contudo, ela é muito importante para a execução da pesquisa.

Sua participação consistirá em sua fala ser gravada, em ambiente silencioso, utilizando um microfone, pronunciando a leitura de uma lista de palavras no idioma inglês (fornecida pelo pesquisador), divididas em quarenta amostras, obtendo-se quarenta arquivos de áudio, cada um com a fala das palavras listadas na respectiva amostra. O processo de gravação deve durar de três a quatro horas, podendo ser realizado em até três sessões distintas. Essas gravações serão registradas e utilizadas para treinar um sistema de reconhecimento da fala, que se tornará, então, capaz de reconhecer comandos de voz com sua fala. Numa próxima etapa, realizaremos um procedimento para testar o sistema de comandos por voz, em que você pronunciará vinte repetições de cada um dos comandos definidos, no idioma inglês, para o cenário de testes do sistema. Esses testes serão utilizados para verificar se a plataforma está realizando os comandos de voz de forma satisfatória. O tempo de duração do teste é cerca de duas horas e meia, podendo ser realizado em até duas sessões distintas. Com o intuito de contribuir com a comunidade de desenvolvimento de tecnologias livres e de código aberto, de baixo custo, para o reconhecimento de fala, os arquivos gravados poderão ser disponibilizados no projeto Voxforge (<http://www.voxforge.org/>), conforme sua opção escolhida no devido campo desse termo, sem a necessidade de que você seja identificado. Ao final da pesquisa, todo material será mantido em arquivo, por pelo menos 5 anos, conforme resolução n °466/12 do Conselho Nacional de Saúde.

Com a sua colaboração nesta pesquisa, você não terá benefício pessoal direito, mas sua participação proporcionará a obtenção de dados que permitirão contribuir com a comunidade de desenvolvimento de tecnologias livre e de código aberto, de baixo custo, para

o reconhecimento de fala, que auxiliem no cotidiano de pessoas com mobilidade reduzida, proporcionando-lhes maior independência, autonomia e, conseqüentemente, mais qualidade de vida.

Em decorrência de sua participação na pesquisa, você poderá sentir cansaço devido ao esforço prolongado ao falar repetidamente para a gravação das amostras de treinamento, como também para a realização dos testes da plataforma. Para minimizar esse desconto, a realização das gravações de amostras de áudio de treinamento poderá ser realizada em até 3 sessões distintas, como também o teste do sistema ser realizado em até duas sessões distintas. Durante as sessões, serão realizadas pausas a cada 30 minutos de gravação, sendo fornecida água para sua hidratação. Além disso, você poderá sentir-se constrangido com a possibilidade de publicação no projeto Voxforge dos arquivos de áudio gravados com sua voz. Visando minimizar esse desconforto, a publicação do arquivo será realizada sem sua identificação, ou você pode ainda optar pela não publicação do arquivo gravado com sua voz, conforme a opção feita a seguir.

O participante permite a publicação dos arquivos de áudio gravados com sua voz no projeto Voxforge (<http://www.voxforge.org/>)? () Sim. () Não.

Os resultados desta pesquisa serão divulgados em artigo científico e em dissertação do Curso de Pós-graduação *Strictu Sensu* de mestrado em Engenharia Elétrica do IFPB. Seu nome não será identificado em nenhum desses documentos.

Este termo de Consentimento livre e Esclarecido é redigido em duas vias, sendo uma para o participante e outra para o pesquisador. Todas as páginas do documento serão rubricadas pelo participante da pesquisa e pelo pesquisador responsável, com exceção da última página, onde serão postas ambas as assinaturas.

Para qualquer outra informação sobre a pesquisa, você poderá entrar em contato com o pesquisador responsável pelo telefone (83) 98751-7898, e-mail naironsousa@yahoo.com.br. Em caso de dúvida quanto à condução ética do estudo, entre em contato com o Comitê de Ética em Pesquisa do IFPB através dos seguintes canais de comunicação: telefone (83) 3612-9725, e-mail eticaempesquisa@ifpb.edu.br, endereço Avenida João da Mata, 256, Jaguaribe, João Pessoa-PB.

Consentimento Pós-Informação

Eu, _____, abaixo assinado, fui devidamente esclarecido quanto aos objetivos da pesquisa, aos procedimentos aos quais serei submetido e os possíveis riscos decorrentes da minha participação. Diante do exposto, aceito livremente participar do estudo intitulado **Automação Residencial por Voz para Pessoas com Mobilidade Reduzida**, desenvolvido pelo pesquisador **Náiron de Vasconcelos Sousa**, sob a orientação da professora Dra. **Suzete Elida Nóbrega Correia**.

João pessoa, _____ de _____ de _____

Assinatura do Participante

Assinatura do pesquisador responsável

**ANEXO B – PARECER DO COMITÊ DE ÉTICA
APROVANDO O RELATÓRIO FINAL DE
PESQUISA.**



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E
TECNOLOGIA DA PARAÍBA -



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: AUTOMAÇÃO RESIDENCIAL POR VOZ PARA PESSOAS COM MOBILIDADE REDUZIDA EMPREGANDO JASPER E RASPBERRY PI

Pesquisador: NAIRON DE VASCONCELOS SOUSA

Área Temática:

Versão: 1

CAAE: 79535917.9.0000.5185

Instituição Proponente: INSTITUTO FEDERAL DE EDUCACAO, CIENCIA E TECNOLOGIA DA PARAIBA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.408.765

Apresentação do Projeto:

Trata-se de um projeto de dissertação de mestrado com abordagem metodológica qualitativa que envolve a avaliação do funcionamento de uma plataforma de reconhecimento da fala, dependente de locutor a ser empregada no controle de tecnologias assistivas a partir da interação humano-computador através da pronúncia de comandos pelo usuário. As amostras de áudio, são obtidas realizando-se a gravação da voz de 10 indivíduos que apresentem preservada boa capacidade das habilidades da fala, com idade maior ou igual a 18 anos, escolhidos por acessibilidade.

Objetivo da Pesquisa:

Primário

Desenvolver uma plataforma de reconhecimento da fala, empregando a placa de desenvolvimento Raspberry Pi e o software de assistência social Jasper, para o auxílio de pessoas com mobilidade reduzida.

Secundários

Realizar estudo sobre tecnologias eficientes, portáteis e de baixo custo que se mostrem viáveis ao desenvolvimento interfaces de comandos por voz para a interação humano-omputador;

Elaborar concepção de interface de baixo custo, portátil e de uso pessoal, para aplicações de comandos através do reconhecimento de voz em tecnologias assistivas que atuem no controle de

Endereço: Avenida João da Mata, 256 - Jaguaribe

Bairro: Jaguaribe

UF: PB

Município: JOAO PESSOA

CEP: 58.015-020

Telefone: (83)3612-9725

E-mail: eticaempesquisa@ifpb.edu.br



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E
TECNOLOGIA DA PARAÍBA -



Continuação do Parecer: 2.408.765

dispositivos domésticos;

Avaliar a interface concebida com fins de validar sua eficiência.

Avaliação dos Riscos e Benefícios:

Riscos

Segundo os pesquisadores, "Os participantes da pesquisa estarão sujeitos a riscos relacionados à realização de esforço prolongado ao falar repetidamente para a gravação das amostras de treinamento, como também para a realização das gravações de amostras de áudio de treinamento poderão ser realizadas em até três sessões distintas para cada indivíduo, como também o teste do sistema ser realizado em até duas sessões distintas. Durante as sessões, os participantes serão ainda orientados a realizar pausas a cada 30 minutos de gravação, sendo fornecida água para hidratação. Além disso, os participantes poderão optar por não ter seu arquivo gravado de voz, conforme marcação referente a essa opção que será devidamente formalizada pelo TCLE."

Benefícios

"A pesquisa disponibilizara recursos de base de texto e base de áudio para a criação de modelos acústicos e uma plataforma de reconhecimento de comandos de fala livres, de código aberto e de baixo custo, contribuindo para a popularização dessas tecnologias no âmbito das tecnologias assistivas que proporcionassem maior independência e autonomia para as pessoas com mobilidade reduzida, tendo em vista sua inclusão social e digital."

Diante do exposto, os benefícios se sobressaem aos riscos quanto à participação da amostra no estudo.

Comentários e Considerações sobre a Pesquisa:

Trata-se de um estudo com relevância acadêmica e social voltado para a área de Engenharia, que busca desenvolver uma plataforma de reconhecimento da fala, observando o comportamento do sistema desenvolvido com relação a sua resposta aos comandos de voz, de modo a caracterizar seu desempenho. O foco de atuação do projeto volta-se para o desenvolvimento de interface de comandos por voz que seja facilmente adaptada e incorporada aos já existentes sistemas e ferramentas de auxílio ao controle do ambiente doméstico (domótica). Apresenta justificativa, metodologia e hipótese compatíveis com os objetivos a serem

Endereço: Avenida João da Mata, 256 - Jaguaribe

Bairro: Jaguaribe

CEP: 58.015-020

UF: PB

Município: JOAO PESSOA

Telefone: (83)3612-9725

E-mail: eticaempesquisa@ifpb.edu.br



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E
TECNOLOGIA DA PARAÍBA -



Continuação do Parecer: 2.408.765

investigados, tornando o projeto em tela exequível metodologicamente.

Considerações sobre os Termos de apresentação obrigatória:

De acordo com a Resolução 466/2012 do CNS, os termos obrigatórios para pesquisas envolvendo seres humanos são:

- 1- Folha de rosto, está presente e devidamente preenchida e assinada pelo servidor responsável e dirigente máximo da instituição;
- 2- TCLE, está presente e atende a todas as exigências éticas contidas na Resolução 466/2012 da CNS, com linguagem clara para a população a ser investigada;
- 3- Orçamento, está presente e detalhado;
- 4- Recursos, está presente e compatível com a Resolução;
- 5- Informações básicas do projeto, estão presentes e devidamente preenchidas, incluindo toda equipe de trabalho.

Recomendações:

Sem recomendações.

Conclusões ou Pendências e Lista de Inadequações:

Após avaliação do parecer apresentado pelo relator, o Comitê de Ética em Pesquisa do IFPB discutiu sobre os diversos pontos da análise ética que preconiza a Resolução 466/2012 do Conselho Nacional de Saúde e deliberou o parecer de APROVADO para o referido protocolo de pesquisa.

Informamos ao pesquisador responsável que observe as seguintes orientações: 1- O participante da pesquisa tem a liberdade de recusar-se a participar ou retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 466/2012 - Item IV.3.d).

2- O Termo de Consentimento Livre e Esclarecido deve ser elaborado em duas vias, rubricadas em todas as suas páginas e assinadas, ao seu término, pelo convidado a participar da pesquisa, ou por seu representante legal, assim como pelo pesquisador responsável, ou pela(s) pessoa(s) por ele delegada(s), devendo as páginas de assinaturas estar na mesma folha. Em ambas as vias deverão constar o endereço e contato telefônico ou outro, dos responsáveis pela pesquisa e do CEP local e da CONEP, quando pertinente (Res. CNS 466/2012 - Item IV.5.d) e uma das vias entregue ao participante da pesquisa.

3- O pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e

Endereço: Avenida João da Mata, 256 - Jaguaribe

Bairro: Jaguaribe

CEP: 58.015-020

UF: PB

Município: JOAO PESSOA

Telefone: (83)3612-9725

E-mail: eticaempesquisa@ifpb.edu.br



**INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E
TECNOLOGIA DA PARAÍBA -**



Continuação do Parecer: 2.408.765

descontinuar o estudo somente após análise das razões da descontinuidade por parte do CEP que aprovou (Res. CNS 466/2012 - Item III.2.u), aguardando seu parecer, exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade de regime oferecido a um dos grupos da pesquisa (Item V.4) que requeiram ação imediata.

4- O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo (Res. CNS 466/2012 Item V.5).

5- Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas.

6- Deve ser apresentado ao CEP relatório final até 02/03/2018.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1006875.pdf	28/10/2017 08:43:49		Aceito
Projeto Detalhado / Brochura Investigador	Pesquisa.doc	28/10/2017 08:37:47	NAIRON DE VASCONCELOS SOUSA	Aceito
Outros	Oficio.docx	26/10/2017 20:50:01	NAIRON DE VASCONCELOS SOUSA	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx	26/10/2017 20:40:42	NAIRON DE VASCONCELOS SOUSA	Aceito
Folha de Rosto	FolhadeRosto.pdf	26/10/2017 20:39:32	NAIRON DE VASCONCELOS SOUSA	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Endereço: Avenida João da Mata, 256 - Jaguaribe

Bairro: Jaguaribe

CEP: 58.015-020

UF: PB

Município: JOAO PESSOA

Telefone: (83)3612-9725

E-mail: eticaempesquisa@ifpb.edu.br



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E
TECNOLOGIA DA PARAÍBA -



Continuação do Parecer: 2.408.765

JOAO PESSOA, 30 de Novembro de 2017

Assinado por:
Aleksandro Guedes de Lima
(Coordenador)

Endereço: Avenida João da Mata, 256 - Jaguaribe

Bairro: Jaguaribe

CEP: 58.015-020

UF: PB

Município: JOAO PESSOA

Telefone: (83)3612-9725

E-mail: eticaempesquisa@ifpb.edu.br

