

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CÂMPUS CORNÉLIO PROCÓPIO
CURSO SUPERIOR DE ENGENHARIA ELETRÔNICA
ENGENHARIA ELETRÔNICA**

PEDRO HENRIQUE GARCIA DE SOUZA

**DESENVOLVIMENTO DE UMA APLICAÇÃO DE AQUISIÇÃO E
PROCESSAMENTO DE VOZ**

TRABALHO DE CONCLUSÃO DE CURSO

CORNÉLIO PROCÓPIO

2019

PEDRO HENRIQUE GARCIA DE SOUZA

**DESENVOLVIMENTO DE UMA APLICAÇÃO DE AQUISIÇÃO E
PROCESSAMENTO DE VOZ**

Trabalho de Conclusão de Curso apresentado ao Curso Superior de Engenharia Eletrônica da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Bacharel em Engenharia Eletrônica”.

Orientador: Prof. Dr. Paulo Rogério Scalassara

CORNÉLIO PROCÓPIO

2019

Sobrenome, Nome
S661a A análise sensorial como ferramenta para otimização do
processamento de alimentos. Nome e Sobrenome. – São
Paulo: [s.n.], 2009.
54f.: il.

Monografia apresentada à “Nome da Instituição de Ensino
Superior” como parte dos requisitos exigidos para a conclusão
Do Curso de ...

Orientador: Prof. Nome e Sobrenome

1. Processamento de alimentos. 2. Controle da produção.
I. Título. II. Sobrenome, Nome. (Orientador)

641.3 CDD



TERMO DE APROVAÇÃO

Desenvolvimento de uma aplicação de aquisição e processamento de voz

por

Pedro Henrique Garcia de Souza

Esta Trabalho de Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharel em Engenharia Eletrônica” e aprovado em sua forma final pelo Curso Superior de Engenharia Eletrônica da Universidade Tecnológica Federal do Paraná.
Cornélio Procópio, 29/07/2019.

Banca Examinadora:

Luis Fernando Caparroz Duarte, Prof. Dr.
Coordenador do Curso

Paulo Rogério Scalassara, Prof. Dr.
Orientador

Maria Eugenia Dajer, Profa. Dra.
Universidade Tecnológica Federal do Paraná

Cristiano Marcos Agulhari, Prof. Dr.
Universidade Tecnológica Federal do Paraná

À minha família.

AGRADECIMENTOS

A Deus e aos meus mentores espirituais por terem me dado saúde e força para superar as dificuldades.

Agradeço ao meu orientador Prof. Dr. Paulo Rogério Scalassara por todo compartilhamento do conhecimento sobre o assunto para me guiar até o final do trabalho.

Agradeço também a minha família que sempre esteve ao meu lado me incentivando e apoiando nos momentos difíceis enfrentados nesta trajetória. Agradeço a minha namorada Izabela Milo por também estar ao meu lado em todos os momentos.

E a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

"O Princípio é o que importa"

Frank Sherman Land

RESUMO

SOUZA, Pedro Henrique G. **Desenvolvimento de uma aplicação de aquisição e processamento de voz**. 2019. 37 f. Trabalho de Conclusão de Curso – Engenharia Eletrônica, Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2019.

O foco deste trabalho é desenvolver um aplicativo móvel que opere no sistema operacional Android. Este aplicativo coleta dados de um sinal vocal e informações do paciente, utilizando somente os recursos de um aparelho celular. O aplicativo se comunica com um servidor através de uma nuvem, assim todo o processamento é realizado no servidor. O projeto é realizado em 4 etapas: 1) Escolha do sistema operacional e do servidor; 2) Desenvolvimento da aplicação (servidor e aplicativo); 3) Aquisição e processamento dos dados utilizando correlação; 4) Apresentação dos dados na interface gráfica do servidor e no aplicativo para o usuário.

Palavras-chave: Aplicativo Android. Detecção frequência fundamental. Sinal de Voz

ABSTRACT

SOUZA, Pedro Henrique G. **Development of a voice acquisition and processing application.** 2019. 37 f. Master Thesis – Electrical Engineering Graduate Program, Federal University of Technology - Paraná. Cornélio Procópio, 2019.

The focus of this work is to develop a mobile application that runs on Android operating system. This application should collect, data of a vocal signal and patient information, using only the resources of a mobile device. This app will communicate with a server through a cloud, thus all processing will be performed on the server. The project will be carried out in 4 steps: 1) Choice of operating system and server, 2) Application development(server and application), 3) Data acquisition and processing using correlation, 4) Apresentation of data in the server's graphical user interface and mobile application.

Keywords: Android application. Fundamental frequency detection. Voice Signal

LISTA DE FIGURAS

FIGURA 1 – Exemplo de uma figura	17
FIGURA 2 – Aparato Fonador.(a: cavidade nasal, b: cavidade oral, c: laringe, d: pulmões, e: faringe, f: traqueia, g: diafragma.)	18
FIGURA 3 – Caixa torácica. (1: Esterno, 2: Porção óssea, 3: Porção cartilaginosa, I-VII: Costelas verdadeiras, VIII-X: Costelas falsas e XI-XIII: Costelas flutuantes)	19
FIGURA 4 – Pulmões	20
FIGURA 5 – Laringe.	21
FIGURA 6 – Cartilagens da laringe.	22
FIGURA 7 – Músculos da laringe.	23
FIGURA 8 – Histologia das pregas vocais.	24
FIGURA 9 – Pregas vocais.	25
FIGURA 10 – Boca. (a: arco palatoglosso, b: arco palatofaríngeo, c: amígdala palatina, d: língua, e: frênulo do lábio inferior, f: istmo dos maxilares, g: úvula, h: frênulo do lábio superior.)	26
FIGURA 11 – Amostragem no domínio do tempo.	28
FIGURA 12 – Versões do Android.	31
FIGURA 13 – Raspberry Pi 3 B+.	31
FIGURA 14 – Famílias de processadores ARM.	32
FIGURA 15 – Diagrama funcional do processador ARM1176JZF-S.	33

LISTA DE TABELAS

TABELA 1 – Porções do trato vocal.	18
--------------------------------------------	----

LISTA DE SIGLAS

CFM	Conselho Federal de Medicina
DMA	<i>Direct Memory Access</i>
FPGAs	<i>Field-programmable Gate Array</i>
GPIO pins	<i>General Purpose Input/Output pins</i>
GPU	<i>Graphics Processing Unit</i>
HDMI	<i>High-Definition Multimedia Interface</i>
IoT	<i>Internet of things</i>
ONU	Conselho Econômico e Social das Nações Unidas
PEP	Protuário Eletrônico do Paciente
RISC	<i>Reduced Instructions Set Computing</i>
SBIS	Sociedade Brasileira de Informática em Saúde
TIC	Tecnologias da informação e da comunicação
USB	<i>Universal Serial Bus</i>

LISTA DE SÍMBOLOS

T_a	Período do sinal
$\frac{1}{T_a}$	Frequência fundamental de amostragem
τ	Atraso

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVOS	15
1.1.1	Objetivo Geral	15
1.1.2	Objetivos Específicos	15
1.1.3	Estrutura	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	A VOZ HUMANA	17
2.2	ANATOMIA DO APARATO FONADOR	18
2.2.1	Combustível do aparato fonador.	18
2.2.2	Porção vibradora do aparato fonador.	21
2.2.3	Porção ressonadora do aparato fonador.	25
2.3	PROCESSAMENTO DE SINAIS DIGITAIS	27
2.3.1	Estimação da frequência fundamental da voz	28
2.4	SISTEMA OPERACIONAL ANDROID	29
2.4.1	Plataforma de código aberto	30
2.4.2	Versões do Android	30
2.5	RASPBERRY PI 3	31
2.6	INFORMAÇÕES ELETRÔNICAS DO USUÁRIO	33
2.7	SERVIDOR E CLIENTE	34
	REFERÊNCIAS	35

1 INTRODUÇÃO

A voz humana é um dos instrumentos de comunicação mais importantes utilizados no intuito de expressar e organizar os sentimentos e ideias. Fisicamente, os sons da fala são classificados como ondas de pressão, resultantes de vibrações longitudinais das moléculas do ar. Vale ressaltar que essas vibrações são oriundas de órgãos do corpo humano que exercem as funções vitais de alimentação e respiração (ALCAIM; SANTOS, 2011).

A glote, espaço entre as pregas vocais, se destaca por ser onde o fluxo de ar contínuo dos pulmões é transformado em vibrações rápidas durante a produção da fala. O fechamento dessas pregas vocais causa um aumento gradativo da pressão exercida sobre elas, que ocasiona a abertura repentina das mesmas, assim liberando a pressão, para então se fecharem novamente. A frequência com que essas pregas vocais se abrem é chamada de frequência fundamental de fonação, F_0 , enquanto o intervalo de tempo entre sucessivas aberturas das pregas vocais é denominado período fundamental. Os sons produzidos dessa forma são classificados como sonoros, ou vozeados, onde normalmente se encontram as vogais; caso contrário, são classificados como surdos, ou não-vozeados (RABINER; SCHAFER, 1978).

Analisando esse sistema dinâmico diversas áreas vêm estudando maneiras para conscientizar e diagnosticar de forma eficaz os distúrbios, que normalmente são: rouquidão, tosse, dor de garganta/queimação, fadiga vocal, garganta seca e perda de voz (BEHLAU et al., 2012). Estudos apontam que aproximadamente 70% da população economicamente ativa precisa da voz para realizar seu trabalho, o que pode indicar maiores incidências de distúrbios de voz. Alterações na voz podem, portanto, impactar de forma negativa as relações profissionais e pessoais, o que demonstra que a qualidade vocal influencia a qualidade de vida do indivíduo (CATANI; CARVALHO, 2012).

Atualmente, a maioria dos diagnósticos de distúrbio na região laríngea para doenças como a Distonia Laríngea, é realizado com a coleta de imagens da região da laringe. Estes exames são de alto custo (BENAVIDES et al., 2014), além de serem invasivos e de difícil realização em pacientes dentro da faixa etária pediátrica (ROTHFIELD; RUSSO, 2012). Exames como a videolaringoestroboscopia, no caso da Distonia Laríngea, em que uma fonte de luz com uma câmera acoplada e conectada a um aparelho de monitoramento é inserida na laringe através da narina ou garganta, leva em alguns casos a intenso reflexo nauseante, torna-se difícil a realização do exame (GÓMEZ-PUERTA et al., 2014).

Desta forma, uma consulta clínica pode não ser eficiente para analisar com precisão um caso de distúrbio vocal, pois é analisada somente uma situação ideal do uso da voz naquele instante (MEHTA et al., 2013). Considera-se recomendável realizar um diagnóstico do comportamento vocal em uma amostra do sinal de voz coletada durante um período prolongado de

tempo. O monitoramento de longo prazo da utilização de voz do paciente facilita aos médicos otorrinolaringologistas e/ou fonoaudiólogos quantificarem o tempo, os ciclos, a dosagem e a energia da vibração das pregas vocais e, assim, mensurar a distância que elas percorrem durante um dia de trabalho (HILLMAN et al., 2006).

Contado com o crescimento do interesse e do potencial das tecnologias móveis em auxiliar e aperfeiçoar o atendimento na área da saúde, leva à gradual propagação e avanço do campo de m-health, que é uma abreviação do termo *Mobile Health* (saúde movel), que liga os aparelhos móveis com seus sensores e sistemas em prol da saúde (AIKENS et al., 2014). O m-health pode ser de grande contribuição para os atendimentos médicos em uma diversidade de formas, por exemplo, todo atendimento na área da saúde requer e gera informações, como um prontuário médico, para que um profissional possa atender adequadamente o paciente. Além desses registros das atividades médicas atuais é necessário o histórico do paciente, inclusive de diferentes áreas da saúde, que auxilia na decisão do diagnóstico (ALSOS; DAS; SVANÆS, 2012).

O objetivo desse trabalho é desenvolver um aplicativo responsável pela aquisição de informações dos usuários e aquisição de um sinal de voz que é coletado em diferentes horários predeterminados do dia. O aplicativo tem a opção de reproduzir o áudio gravado ou enviar para nuvem, onde é monitorada por um servidor, que no caso é um Raspberry pi. Esse servidor realiza o download desse sinal da nuvem, processa e analisa a frequência fundamental da voz, em seguida essa informação retorna para nuvem, onde por fim retorna para o aplicativo para visualização do usuário.

A escolha do aparelho celular é reflexo de um constante crescimento da comunicação móvel que penetra em diferentes culturas e classes sociais (PELLANDA, 2009). E nesse sentido, a necessidade de um meio ambulatorio e acessível, economicamente capaz de coletar dados do sinal da voz em diferentes períodos do dia, utilizando uma forma não evasiva, classificou-se como uma opção viável para esse estudo.

1.1 OBJETIVOS

1.1.1 Objetivo Geral

Desenvolvimento de um aplicativo móvel para aquisição do sinal de voz e informações do usuário. Desenvolvimento de um servidor para o processamento dos dados para obtenção da frequência fundamental e geração de um banco de dados.

1.1.2 Objetivos Específicos

- Estudar a voz humana e suas características acústicas.

- Desenvolver um aplicativo para o sistema operacional Android para aquisição de áudio.
- Analisar formatos de compactação de áudio.
- Coletar informação do usuário.
- Aprofundar o conhecimento em processamento digital de sinais, para detecção da frequência fundamental.
- Desenvolver um servidor contendo uma interface gráfica com o raspberry pi.

1.1.3 Estrutura

No próximo capítulo deste trabalho, Fundamentação Teórica, será apresentado a voz humana nos aspectos da anatomia e fisiologia. Já no terceiro capítulo, Metodologia, serão apresentados os métodos adotados para a realização da captação e armazenamento do sinal de voz que posteriormente será processado. Para concluir, as considerações finais do projeto.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 A VOZ HUMANA

A voz existe desde o nascimento por vocalizações expressivas, como o riso e o choro. Ela transporta não só um conteúdo, mas também sentimentos. Desta forma ela se torna um dos meios de maior interação para provocar respostas no outro (BEHLAU; PONTES; MORETI, 2018).

A voz é produzida graças à ação coordenada de quase todo o nosso corpo. O aparelho fonador ou vocal é composto de estruturas musculares de diferentes regiões e elementos do sistema respiratório e sistema digestivo. É importante notar que nenhuma estrutura do nosso corpo tem como função única a produção da voz. A voz foi uma adaptação evolutiva muito mais tarde do que outras ações essenciais para a vida. Assim, a laringe, à qual nos relacionamos de uma maneira automática com a voz, sua principal função é proteger o trato respiratório (TORRES, 2007).

Em relação aos sons, a sonoridade é produzida na laringe, especificamente pela vibração das pregas vocais. Já o aparelho digestório auxilia na articulação dos ativos (lábios, língua, mandíbula e palato mole) e passivos (dentes, palato duro e parede posterior da faringe) e de ressoadores (cavidade oral, nasal, faríngea e labial) que alteram o som produzido na glote.

O ouvido humano sadio tem capacidade de perceber sons no intervalo da frequência entre 20Hz até 20kHz, e a voz humana emite sons na frequência que varia de 80Hz até 12Khz. Analisando a energia do sinal, constata-se que a maior parte dela se encontra na faixa de 200Hz a 3.400Hz (PINHEIRO, 2004), intervalo esse que é adotado pelas companhias telefônicas para telecomunicações, como ilustrado na Figura 1.

Figura 1 – Energia do sinal de voz concentrada na frequência.



Fonte: Adaptado de (PINHEIRO, 2004).

2.2 ANATOMIA DO APARATO FONADOR

O aparato fonador se divide em três porções diferentes, apresentado da Tabela 1.

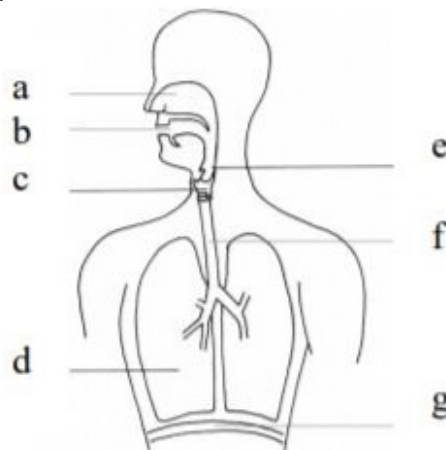
Tabela 1 – Porções do trato vocal.

Combustível	Formada por estruturas infraglólicas que determinam a pressão de ar exalado.
Vibrador	Consiste nas pregas vocais (cordas vocais) da laringe.
Ressonadores	Integrado por cavidades supraglólicas onde o som produzido nas pregas vocais é amplificado e modificado.

Fonte: Adaptado de (TORRES, 2007)

Mesmo com essa divisão, o aparato fonador é um todo homogêneo e inseparável, onde qualquer alteração em qualquer de suas partes acarretará em uma modificação no sistema em consequência na emissão da voz. Na Figura 2 é ilustrado o sistema fonador, os músculos do abdômen estão juntamente com o diafragma responsável por controlar nossa voz.

Figura 2 – Aparato Fonador.(a: cavidade nasal, b: cavidade oral, c: laringe, d: pulmões, e: faringe, f: traqueia, g: diafragma.)



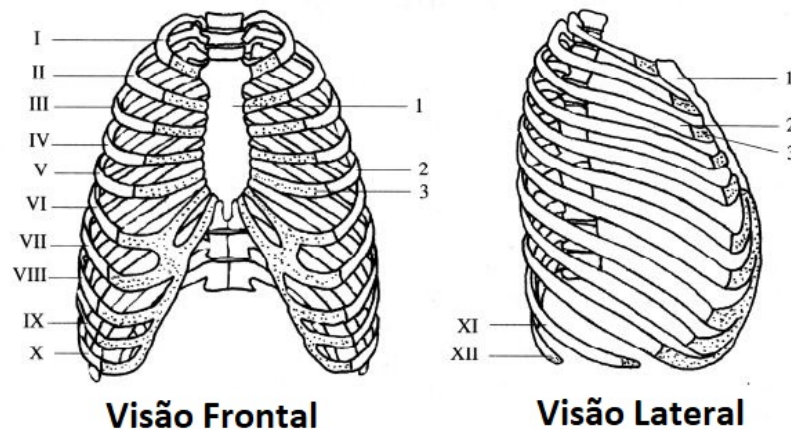
Fonte: Adaptado de (TORRES, 2007).

2.2.1 Combustível do aparato fonador.

O conjunto formado pela caixa torácica, pulmões, diafragma, músculos (músculos inspiradores) e abdômen (exaladores) fornecem a energia necessário para o ar ser exalado e dessa forma esse sistema foi chamado de combustível da voz (TORRES, 2007).

A caixa torácica, Figura 3 é constituída pela junção entre as costelas, o esterno e a porção torácica (dorsal) da coluna vertebral. Esta união é feita por várias articulações que conferem mobilidade e elasticidade a todo o conjunto, o que permite, durante a respiração, que os diâmetros da caixa torácica variem e os pulmões encham e esvaziem de ar.

Figura 3 – Caixa torácica. (1: Esterno, 2: Porção óssea, 3: Porção cartilaginosa, I-VII: Costelas verdadeiras, VIII-X: Costelas falsas e XI-XIII: Costelas flutuantes)



Fonte: Adaptado de (TORRES, 2007).

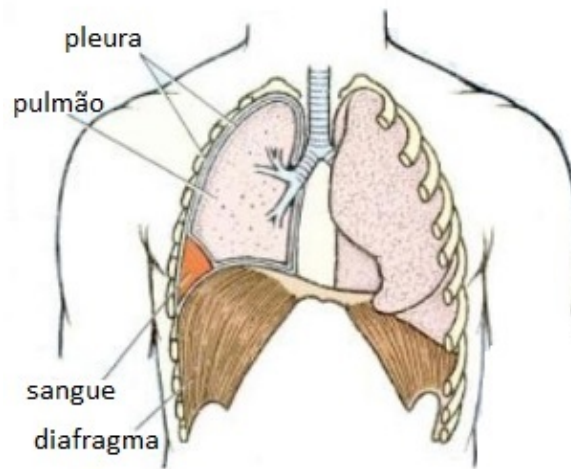
Os movimentos do tórax ósseo é a soma dos movimentos individuais de seus componentes durante a inspiração e exalação. Assim, qualquer músculo que produza a ascensão costal será inspirador, e qualquer um que determine sua descida é um músculo exalador. Durante a inspiração, as costelas superiores são direcionadas para a frente, produzindo aumento do diâmetro ântero-posterior do tórax. As costelas inferiores são direcionadas para os lados produzindo um aumento no diâmetro transversal da caixa torácica (MOORE; DALLEY, 2009).

A traquéia está localizada antes do esôfago. Estende-se entre a laringe e os brônquios principal, direita e esquerda, onde se bifurca. Sua função é dirigir o ar na direção dos pulmões ou fora deles. Os pulmões, apresentados na Figura 4, são os órgãos da respiração, sua função básica é oxigenar o sangue. Cada pulmão é envolvido em sua pleura ou saco pleural. O pulmão direito é formado por três lobos e o esquerdo por dois. A pleura é um saco de parede dupla, um interno intimamente ligado ao pulmão, um externo ligado à parede torácica e ao diafragma. A camada interna e externa da pleura e os pulmões seguem o diafragma e as costelas em seus movimentos respiratórios. Na inspiração, a capacidade da cavidade torácica aumenta nas três direções do espaço. À medida que os pulmões se alarga, há uma redução na pressão intra-alveolar e o ar é inspirado para dentro (MOORE; DALLEY, 2009).

Quando o ar entra nos pulmões as cúpulas direita e esquerda descem achatadas. Esta descida aumenta o diâmetro vertical da cavidade torácica; assim sua abertura inferior da caixa torácica irá atuar em suas articulações determinando o movimento das costelas que causa o aumento dos diâmetros anteroposterior e transversal. Com esse aumento dos três eixos do tórax, os pulmões se enchem de ar.

O diafragma é o principal músculo da respiração. Está situado como uma folha que separa a cavidade torácica do abdominal. Tem um formato de cúpula dupla e constitui o piso

Figura 4 – Pulmões



Fonte: Adaptado de (MOORE; DALLEY, 2009).

da cavidade torácica e o teto abdominal. Durante a inspiração ocorre a contração e durante a expiração o relaxamento.

O diafragma é mais alto na parte anterior (onde se junta ao esterno) do que na costas (anexadas à coluna lombar). Na respiração diafragmática, utilizada quando o diafragma se contrai, ele desce e empurra as vísceras para dentro da cavidade abdominal. Elas empurram a parede do abdome (formada pela musculatura abdominal) que é relaxada (SOUCHARD, 1989).

A expiração silenciosa é um processo passivo no qual a elevação do diafragma é produzido por sua elasticidade e a dos elementos da cavidade torácica. Esse retorno de diafragma determina a saída de ar intrapulmonar que passará pela laringe na qual as pregas vocais se encontram. Na fala ou no canto, a expiração é ativada e será controlada pelos músculos abdominais. Na fonação as cordas vocais recebem o ar expirado pelos pulmões com uma certa pressão para poder abrir a fenda glótica e produza a vibração das pregas vocais.

O abdômen é a porção do tronco entre o tórax e a pelve. Os músculos desta região formam a maioria das paredes abdominais eles funcionam como um cinto que segura fortemente as vísceras abdominais. Durante a inspiração, esses músculos relaxam e o diafragma se contrai. Na expiração, os músculos do abdome se contraem enquanto o diafragma relaxa. Essa ação coordenada constitui o chamado suporte da voz (SOUCHARD, 1987).

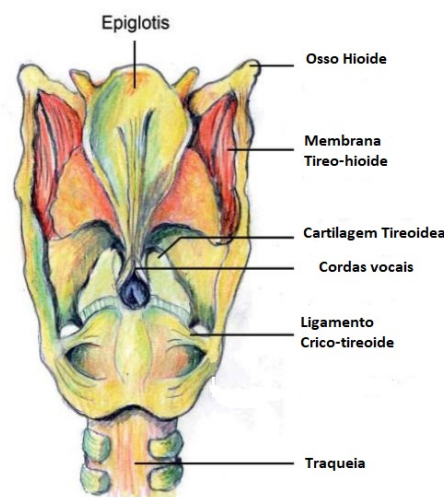
Na fonação, a contração da musculatura do abdome causa o aumento da pressão intra-abdominal. As vísceras que estão dentro do abdômen são comprimidas por esse aumento de pressão e assim empurram o diafragma, causando o seu aumento. A ascensão do diafragma empurra os pulmões e determina um aumento na pressão subglótica, a pressão é suficiente e o ar é exalado fortemente produzindo a abertura e vibração das pregas vocais.

2.2.2 Porção vibradora do aparato fonador.

A porção vibradora do aparato fonador é constituída pela laringe, apresentada na Figura 5, onde se localizam as pregas vocais, que está localizada na parte medial e anterior do pescoço, à frente da faringe. Ela está envolvida nos processos de respiração, deglutição e fonação.

A função da laringe é bloquear corpos estranhos de adentrarem aos pulmões, função principalmente realizada pela epiglote, e auxiliar na deglutição através de um movimento vertical de toda a estrutura laríngea exercido pelos músculos extrínsecos, músculos que conectam-na a outras partes do corpo humano. A ação de produzir o sinal glotal é uma função secundária da laringe. Isso ocorre por causa da morfologia específica de alguns de seus músculos, particularmente das pregas vocais verdadeiras. Simplificadamente podemos dizer que a laringe é uma válvula de controle de vazão de ar (ROSA, 2002).

Figura 5 – Laringe.

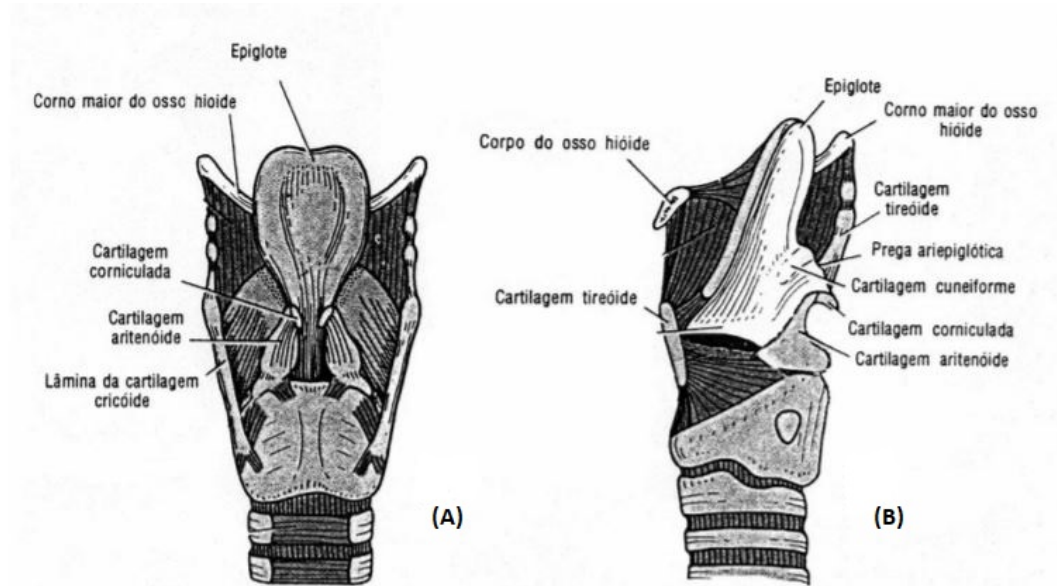


Fonte: Adaptado de (PINHO; KORN; PONTES, 2019).

Anexado a Laringe pela membrana tireo-hioide está o osso híóide, que é o único do corpo que não se articula com outro osso, mas é suspenso do crânio por pequenos ligamentos e estruturas musculares. Nesse osso, por um lado, a língua é inserida e, por outro, o chamado músculos extrínsecos da laringe. As três estruturas constituem, portanto, um bloco funcional inseparável tanto durante a fonação quanto durante deglutição. Dentro do grupo de ligamentos da laringe destacamos os ligamentos vocais e vestibular para constituir o esqueleto das pregas vocais (TORRES, 2007).

As cartilagens da laringe, apresentadas na Figura 6, são sustentadas principalmente pelo osso híóide, elas são nove no total: três ímpares (tireóide, cricóide e epiglote) e três pares (aritenóides, as corniculadas e as cuneiformes). Entretanto, as cartilagens corniculadas e as cuneiformes, tem pouca importância para o processo de fonação (DAJER, 2010).

Figura 6 – Cartilagens da laringe.



Fonte: Adaptado de (ROSA, 2002).

A cartilagem tireóide é composta por duas placas fundidas na posição anterior em ângulo próximo a 90° e abertas posteriormente, adquirindo assim um formato em "V". Os cornos superiores ligam-se às pontas dos cornos do osso hióide enquanto que os cornos inferiores conectam-se às faces da cartilagem cricóide.

Uma das funções da cartilagem tireóide é proteger a estrutura da laringe. É essa a cartilagem, cuja proeminência conhecida por "Pomo de Adão" corresponde a uma elevação onde as lâminas se fundem, que é palpável e vista no ser humano.

Possuindo a forma de um anel, a cartilagem cricóide serve de plataforma para a tireóide. Ela possui uma placa posterior e uma parte anterior estreita. Em cada lado da borda superior da lâmina há uma faceta para junção com uma das cartilagens aritenóides. Sua borda inferior define o término da faringe e laringe e o início do esôfago e traqueia.

Já a epiglote tem a função de evitar que corpos estranhos cheguem aos pulmões. Ela fecha a laringe durante a deglutição e a abre na fonação. Ela está presa ao osso hióide e tem forma de folha.

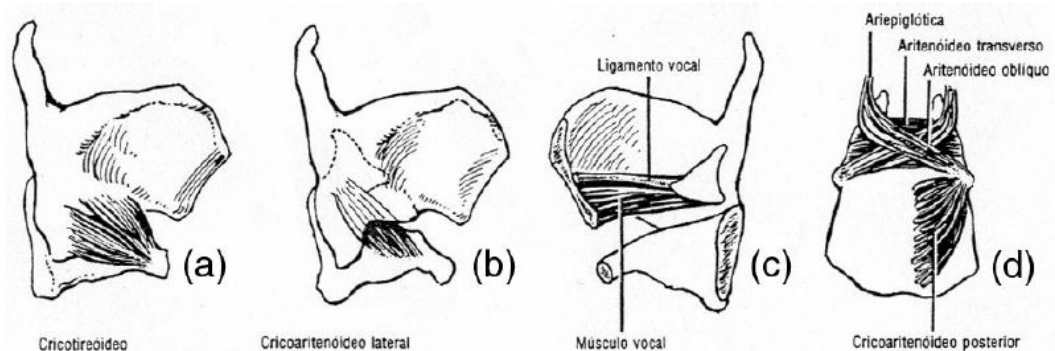
As cartilagens aritenóides estão apoiadas sobre a cartilagem cricóide. Esse acoplamento permite que elas tenham dois movimentos importantes: o "*rocking*", que corresponde a uma rotação, e o "*gliding*", que é um pequeno deslizamento dessas cartilagens sobre a cricóide. Ela possui um formato piramidal e sua ápice dá apoio às cartilagens corniculadas. Sobre duas de suas faces definem-se os processos ditos vocal e muscular. O primeiro corresponde à face em cada uma das cartilagens de onde o músculo vocal parte enquanto que o último à face onde os demais músculos se conectam às aritenóides. Essa diferença de posição de contato muscular

influencia na direção de rotação dessas cartilagens e, por conseguinte, na ação de abertura e fechamento glotal (ROSA, 2002).

A laringe é constituída de um conjunto de músculos intrínsecos e extrínsecos que movimentam ou sustentam as cartilagens na garganta. Os músculos extrínsecos movimentam a laringe como um todo, dando sustentação à mesma, enquanto que os músculos intrínsecos são responsáveis por movimentos que existem dentro da própria laringe. A fonação é particularmente dependente da ação desses músculos intrínsecos. Eles são responsáveis pela abertura e fechamento do espaço glotal, atividades que efetivamente controlam a vibração das pregas vocais.

Os Músculos intrínsecos da laringe, apresentado na Figura 7, são em pares, exceto o aritenóide transversal que é único. Todos eles são adutores das pregas vocais, com exceção do cricoaritenóideo posterior. Esse músculo é responsável por separar as pregas vocais durante a respiração e permite a entrada e saída de ar. O resto dos músculo próximo às pregas vocais fecham a entrada das vias aéreas. Durante a fonação (no período prefonatório) esses músculos são aqueles que trazem as pregas vocais e impedem o ar expirado de sair, assim fazendo as cordas vocais vibrarem.

Figura 7 – Músculos da laringe.



Fonte: Adaptado de (ROSA, 2002).

Além da ação de aproximar as pregas vocais, existem dois músculos intrínsecos que são responsáveis por mudanças no tom fundamental da voz: o cricotireóideo e o vocal.

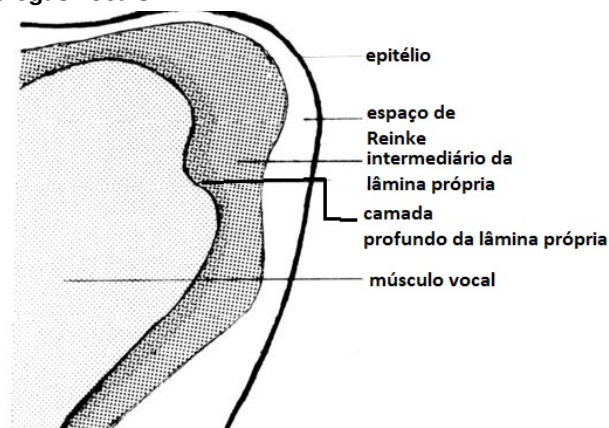
O músculo cricotireóideo faz com que a cartilagem tireoidiana se incline sobre o cricóide (na articulação cricotireóidea) alongando e enrijecendo as pregas vocais. As variações finas no tom de voz são devidas ao músculo vocal. Esse músculo, que é encontrado dentro da prega vocal ao contrair causa um aumento no volume das pregas vocais e, assim, modifica suas vibrações. Essa ação é adicionada ao músculo cricotireóideo que, agindo em conjunto, prolonga as pregas vocais tensas. Assim, cada dobra, graças à ação combinada desses dois músculos se comporta como se fosse um conjunto de estruturas diferentes, com espessuras e tensões

diferentes. Como esse processo ocorre nas duas dobras, o resultado é que temos dois corpos vibrantes de características variáveis que permitirão emitir todos os ricos sons desde a voz humana até a passagem do ar expirado.

Em cada lado da superfície interna da laringe encontram-se duas dobras de sua mucosa, as pregas vestibulares ou cordas vocais falsas, e as dobras vocais ou cordas vocais verdadeiras. A prega vestibular cobre o ligamento vestibular e a prega vocal cobre o ligamento vocal (TORRES, 2007).

As cordas vocais verdadeiras são compostas por um tecido superficial e um tecido muscular interno. Histologicamente, elas são divididas em epitélio, lamina própria e corpo muscular, divisão está apresentado na Figura 8.

Figura 8 – Histologia das pregas vocais.



Fonte: Adaptado de (TORRES, 2007).

O epitélio é constituído por células escamosas estratificadas. Essas células possuem uma maior adesão com as suas vizinhas pela presença de um componente orgânico, a desmossoma. É possível verificar que na proximidade da superfície, a quantidade de desmossoma é diretamente proporcional à profundidade em que estão os agregados celulares do epitélio, tornando-se praticamente nula na superfície do epitélio. Isso permite que as células mortas sejam facilmente eliminadas e substituídas, pois devido as colisões de corpos que ocorre durante a fonação, as células do epitélio exigem uma reposição maior (ROSA, 2002).

A lâmina própria pode ser subdividida em três camadas: superficial, intermediária e profunda. A diferença entre elas se dá pela constituição das fibras (elásticas e colágenas).

A camada superficial, também conhecida por espaço de Reinke, é flácida. É essa a camada que tem a maior vibração durante a fonação. Assim, qualquer modificação que a torne mais rígida resultará em uma menor vibração da prega vocal, dificultando a fonação.

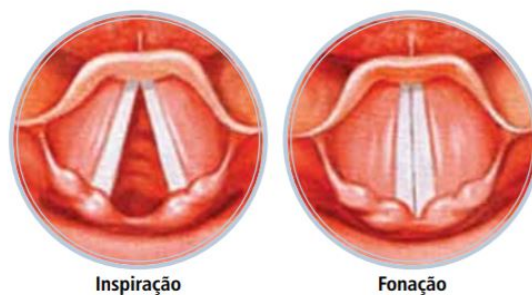
A camada intermediária é constituída, principalmente, de fibras elásticas enquanto que a camada profunda, de fibra colágenas. Quanto mais nos aproximamos do músculo vocal, o

número de fibras colágenas aumenta enquanto que o de fibras elásticas diminui. Assim, a rigidez das camadas intermediária, profunda e músculo vocal aumentam com a profundidade (VIEIRA, 2003).

O músculo vocal é constituído basicamente das mesmas estruturas dos músculos do esqueleto do corpo humano, sendo longo e cilíndrico. Danos nesse músculo acarretam a incapacidade de aproximação na redução do espaço glotal e consequentemente na incapacidade de fonação. Esse músculo é responsável por definir o volume das pregas vocais verdadeiras. Aumentando-se sua tensão, as fibras se reorganizam, aumentando a densidade fibrilar e reduzindo o volume de massa das pregas vocais.

As falsas pregas vocais têm uma constituição flácida, semelhante à camada superficial da lâmina própria e contém poucas fibras musculares quando comparadas com as pregas vocais verdadeiras. A ação exercida pelo músculo vocal nas pregas vocais verdadeiras não ocorre aqui.

Figura 9 – Pregas vocais.



Fonte: Adaptado de (CORRÊA, 2016).

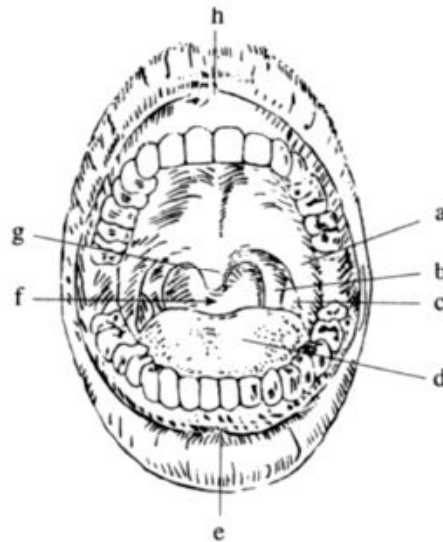
2.2.3 Porção ressonadora do aparato fonador.

Todas as cavidades localizadas acima das pregas vocais agem, ou podem agir, como caixas de ressonância sonora. São ressonadores ou cavidades supraglóticas a boca, faringe e narinas. Existem ressonadores móveis, como a boca, que podem modificar sua forma e volume adaptando-se ao som produzido e outros fixos, como as narinas, que não podem mudar sua forma ou volume. A boca será modificada dependendo da abertura mandibular e da posição da língua, dentes e lábios. A faringe muda sua morfologia principalmente com base nos deslocamentos da laringe, língua e véu do palato ou palato mole.

A boca, apresentado na Figura 10, é o principal ressonador da voz. É possível adaptar sua forma e volume do som emitido nas pregas vocais por meio de alterações na posição do língua, lábios, véu do palato e mandíbula. A intensidade ou volume final do som será diretamente proporcional à área de abertura dela. Isso explica a importância da abertura mandibular na

música. Além disso, a abertura da boca influenciará no timbre da voz. Para que a boca seja uma cavidade de ressonância eficiente, deve haver a coordenação de todas as suas estruturas.

Figura 10 – Boca. (a: arco palatoglosso, b: arco palatofaríngeo, c: amígdala palatina, d: língua, e: frênulo do lábio inferior, f: istmo dos maxilares, g: úvula, h: frênulo do lábio superior.)



Fonte: Adaptado de (CORRÊA, 2016).

Os lábios limitam a fenda oral. Lateralmente juntam-se formando o canto da boca e o ângulo da boca. Dentro dos lábios, entre outros tecidos, encontramos principalmente um músculo: o orbicular da boca. Os lábios são continuados com as bochechas que contêm o músculo bucinador; principal músculo desta região. Esses dois músculos fazem parte dos chamados músculos faciais; que são responsáveis pelas diferentes expressões. Dependendo da posição dos lábios, serão obtidas alterações no timbre da voz. De todos os músculos mencionados o orbicular da boca e o bucinador serão os mais importante para obter uma característica timbral.

A língua é um órgão muscular localizado no assoalho da boca, envolvido na sucção, mastigação, deglutição e fonação. Nela, distinguimos diferentes porções. A musculatura da língua é dividida em extrínseca e intrínseca. Os músculos intrínsecos estão dentro da língua e variam de forma, já os músculos extrínsecos são responsáveis por juntar a língua a diferentes porções ósseas.

A faringe é a parte do trato digestivo localizada atrás da cavidade nasal, vestibular e da laringe. Atua como um canal comum para engolir e respirar. Por sua parte anterior, comunica cranialmente com as narinas, medialmente com cavidade oral e caudalmente com a laringe.

A faringe pode atuar como um ressonador de voz. Dependendo do tamanho desta cavidade, o ar expirado ressoará nele com maior ou menor intensidade. Se durante a fonação, a

laringe é colocada muito baixa, devido à contração dos músculos infra-hióideos, ocorre principalmente o aumento do volume da porção laríngea consequentemente o da faringe com a qual a voz ressoa, adquirindo timbres diversificados (TORRES, 2007).

A cavidade nasal, pode ser total ou parcialmente separada da cavidade oral através do véu do palato, produzindo sons da fala oral ou nasal. O nariz é a porção externa que projeta a cavidade nasal para frente. A cavidade nasal é formada pelas narinas direita e esquerda separadas pelo septo nasal. As narinas são duas cavidades localizadas sob o crânio e sobre a cavidade bucal, dos quais são separados pelo palato. Eles têm um teto, formado por ossos do crânio, um piso formado pelo palato, uma parede medial o septo nasal e uma parede lateral. Eles subsequentemente se comunicam com a porção nasal da faringe através das coanas.

Durante a inspiração, o ar entra no salão, uma cavidade comum à região respiratória e olfativa, e é dividida em duas correntes. As passagens mais importantes pela região respiratório para a faringe. A outra corrente atinge a região olfativa carregando as partículas de odor para os receptores olfativos. A mucosa nasal é muito vascular e, principalmente na área das torres, esquenta e umidifica o ar, fato muito importante para olfação.

Como vimos anteriormente, a cavidade nasal atua como uma cavidade de ressonância então, é necessário que o véu do palato fique relaxado e o ar expirado saia através dessa região. Ambas as narinas podem atuar como caixas de ressonância independentes.

2.3 PROCESSAMENTO DE SINAIS DIGITAIS

O sinal de voz $x(t)$ é um sinal analógico, contínuo em amplitude e no tempo. Antes de qualquer operação de processamento digital, esse sinal deve ser amostrado no tempo. Se T_a é o período de amostragem, em segundos, então F_a , a frequência fundamental de amostragem, em Hertz, é igual a $\frac{1}{T_a}$. Assim, o sinal passa a ser representado por uma sequência de números $x(n)$, conforme ilustrado na Figura 11.

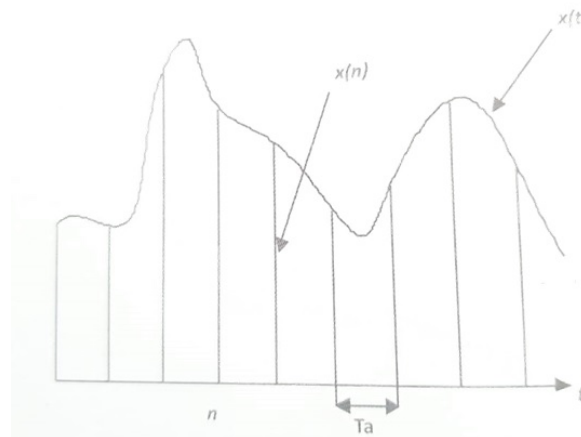
Um sinal no tempo discreto é aquele que pode ser representado por uma sequência de números. Por exemplo a sequência

$$x(n), n \in \mathbb{Z}. \quad (1)$$

onde \mathbb{Z} é o conjunto dos números inteiros, pode representar um sinal no tempo discreto onde cada número $x(n)$ corresponde à amplitude do sinal em cada instante (nT) . Se $x_a(t)$ é um sinal analógico, tempo que

$$x(n) = x_a(nT), n \in \mathbb{Z}. \quad (2)$$

Figura 11 – Amostragem no domínio do tempo.



Fonte: Adaptado de (ALCAIM; SANTOS, 2011).

Como n é um inteiro, T representa o intervalo entre dois pontos sucessivos nos quais o sinal é definido. É importante notar que T não é necessariamente uma unidade de tempo. Por exemplo, se $x_a(t)$ é a temperatura ao longo de uma barra de metal, então T pode ser uma unidade de comprimento, e nesse caso $x(n) = x_a(nT)$ pode representar a temperatura em sensores uniformemente posicionados ao longo da barra.

2.3.1 Estimação da frequência fundamental da voz

A detecção da frequência fundamental do sinal de voz é fundamental no processamento de sinais de voz, pois indica a presença de vocalização, ou seja, que as cordas vocais estão vibrando. A maioria das medidas acústicas depende da detecção da frequência fundamental e a diferença entre os valores das medidas acústicas nos programas existentes é devida, em grande parte, à diferença entre os diversos algoritmos usados (RABINER et al., 1976). Um algoritmo bastante utilizado é o de Boersma (BOERSMA, 1993), que utiliza o método de autocorrelação.

Em um processo de correlação, as formas de onda são comparadas em diferentes intervalos de tempo, sendo então calculada suas similaridades, podendo então se definir a correlação entre duas formas de onda como uma medida de quão similares elas são entre si aplicando-se um valor de atraso (τ). Quando falamos de autocorrelação, aplicamos o processo de correlação entre uma onda e ela mesma. Sendo representada matematicamente na equação 3, para uma função discreta finita, $x[n]$, de tamanho N . Na autocorrelação se é esperado total similaridade quando o valor de τ é igual a zero, e uma maior dissimilaridade conforme seu incremento (GERHARD, 2003). No método de autocorrelação, após uma busca dentre uma variação de τ , se é escolhido o maior pico para uma variação de τ , diferente de zero. Um dos problemas da autocorrelação é que também são identificados alguns picos que ocorrem

em sub-harmônicas, dificultando então a identificação do real pico que representa a frequência fundamental (CHEVEIGNÉ; KAWAHARA, 2002).

$$R_x(\tau) = \sum_{n=0}^{N-1-\tau} x[n]x[n+\tau] \quad (3)$$

2.4 SISTEMA OPERACIONAL ANDROID

De acordo com Tenenbaum (TANENBAUM, 2009), o sistema operacional pode ser representado por dois modos distintos. O primeiro é uma visão de baixo para cima, onde o sistema operacional é um gerente que controla quais processos podem ser executados quando e quais recursos(memória, disco, periféricos) podem ser utilizados. Já o segundo modo é uma visão de cima para baixo que consiste em uma abstração do hardware, onde sistema operacional tem a função de apresentar ao usuário o equivalente a uma máquina estendida que é mais fácil de programar e manipular, ou seja, o mediador entre programas e componentes de hardware.

O Android é uma plataforma para dispositivos móveis como *smartphones* e *tablets*, como também possui versões com interface específicas para TVs (Android TV), carros (Android Auto) e *smartwatch*(Android Wear).E, de acordo com a StatCounter (STARCOUNTER... , 2018), empresa que monitora o tráfego na internet, o Android é atualmente o sistema operacional líder absoluto, superando até mesmo o sistema operacional *Windows*.

A plataforma além de possuir a pilha de software denominado sistema operacional, ela também têm um middleware e aplicações base. Sendo baseado no kernel do Linux, o qual é responsável pelo gerenciamento de processos, drivers, memória, sistema de arquivos, E/S de rede e energia. O middleware, por sua vez, é o responsável por controlar a interação entre os aplicativos instalados no aparelho, facilitando a comunicação entre eles. Por fim, as aplicações base são os aplicativos como discador, agenda de contatos, correio de mensagens, navegador, relógio e entre outros (GLAUBER, 2014). Sendo assim, aplicações base estendem as funcionalidades do aparelho e são integradas ao sistema operacional.

A plataforma possui o Android SDK que provê as ferramentas e APIs necessárias para o programador desenvolver aplicativos para o sistema operacional. Atualmente é possível criar aplicações não só para o Android como para outras plataformas como Windows Phone e iOS utilizando a linguagem JavaScript. Entretanto, Java e Kotlin são as linguagens oficiais suportadas para o desenvolvimento nativo de aplicativos para o Android. Sendo o software Android Studio o principal ambiente para desenvolvimento por conter todas as ferramentas necessárias em um único lugar.

2.4.1 Plataforma de código aberto

O projeto de código do Android é *open source* liderado pela empresa Google e conforme estar licenciado sob licenças comerciais Apache/MIT. Isso permite que fabricantes de aparelhos faça modificações, utilize em seus produtos e isso não os obriga a compartilhar as modificações com a comunidade de desenvolvimento, nem tão pouco com seus concorrentes. Permitindo assim, que tenham versões do Android personalizadas (GARGENTA, 2011).

Uma vez que a plataforma é *open source*, isso possibilita que qualquer pessoa possa contribuir com o desenvolvimento, fazendo com que a plataforma seja atualizada com mais frequência e que seja amadurecida mais rápida, ou seja:

"O fato de o Android ser de código aberto contribui muito para seu aperfeiçoamento, uma vez que desenvolvedores de todos os lugares do mundo podem contribuir para seu código-fonte, adicionando novas funcionalidades ou simplesmente corrigindo falhas"(LECHETA, 2013)

Entretanto, vale ressaltar que o código do sistema operacional Android não é em sua totalidade livre, e de acordo com (GARGENTA, 2011), existem alguns fragmentos de código de baixo nível que são proprietários, tais como a pilha de código para rede de telefone, wi-fi, bluetooth GPS entre outros. Ou seja, mesmo fazendo parte do sistema operacional, esses fragmentos de código não são licenciados como código aberto, que permite a legalidade das modificações para fins diversos.

Mesmo tendo essa liberdade de códigos abertos, o Android tem uma capacidade de segurança muito forte. Como é executado em um kernel do Linux, toda vez que um aplicativo for instalado em uma estação Android, é criado um novo usuário Linux para aquele programa, com diretórios que serão usados pelo aplicativo, mas somente para aquele usuário Linux. Como os aplicativos ficam completamente isolados precisa ser explicitamente autorizada pelo usuário, podendo ser negada a instalação do aplicativo, ou autorizada a instalação, mas controladas as permissões que este aplicativo poderá ter através de um mecanismo de permissão (PEREIRA; SILVA, 2009).

2.4.2 Versões do Android

A primeira versão comercial do Android foi lançada em setembro de 2008, e desde o lançamento até os dias de hoje tiveram várias versões, conforme ilustrado na Figura 12, a partir da versão 1.5 da Google atribui ao nome da versão um nome de doce e um número sequencial denominado API Level (GLAUBER, 2014). Nesse trabalho o aplicativo foi desenvolvido com suporte a partir da versão 4.0.3, de nome *IceCreamSandwich* e API Level 15.

Figura 12 – Versões do Android.



Fonte: (LIMA, 2017).

2.5 RASPBERRY PI 3

O Raspberry Pi, apresentado da Figura 13, é um computador de baixo custo com dimensões físicas de um cartão de crédito. Assim como computadores *desktop* modernos, possui saída *High-Definition Multimedia Interface* (HDMI), saída de áudio, interface *Universal Serial Bus* (USB) e interface Ethernet.

Além desses periféricos, possui diversos pinos denominados *General Purpose Input/Output pins* (GPIO pins), cuja função é servir de entrada ou saída de sinais (RASPBERRY..., 2019).

Figura 13 – Raspberry Pi 3 B+.

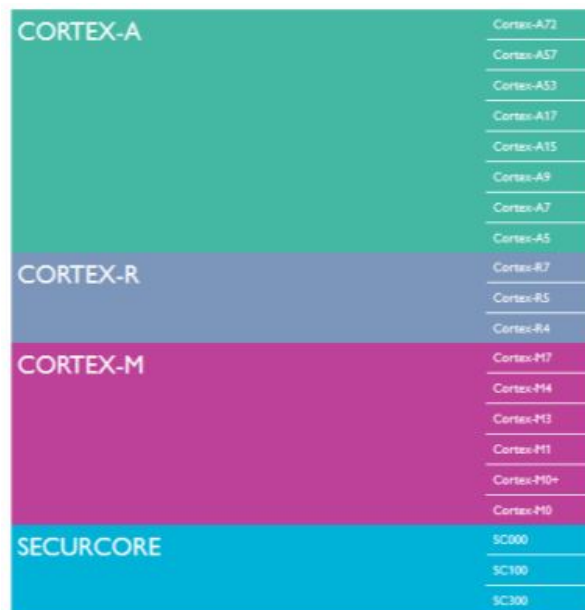


Fonte: (RASPBERRY..., 2019).

Contando com um processador ARM 1176 com relógio padrão de 700MGz e uma *Graphics Processing Unit* (GPU), ele é capaz de processar 1 milhão de pixels por segundo, tendo um grande poder de processamento. A premissa fundamental dos processadores pertencentes a essa família é o conjunto de instruções *Reduced Instructions Set Computing* (RISC), cujo objetivo é reduzir drasticamente o número de instruções necessárias para que o processador realize determinada ação (ARM..., 2019). A figura 14 esquematiza as diversas famílias de processadores ARM.

A família de processadores CORTEX-A é amplamente utilizada no projeto de *smartpho-*

Figura 14 – Famílias de processadores ARM.



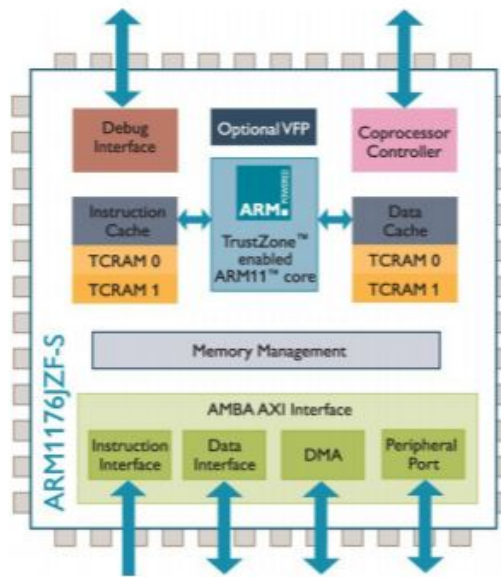
Fonte: (ARM..., 2019).

nes, notebooks e televisores digitais. A família CORTEX-R encontra aplicações no projeto de sistemas de freio automotivo, controladores de armazenamento em massa, e controladores de rede. A família CORTEX-M é utilizada para implementação de airbags, sensores inteligentes, e microcontroladores. Aplicações de alto nível de segurança e processadores *Field-programmable Gate Array* (FPGAs) são projetados com processadores da família SECURCORE.

O processadores ARM1178JZF-S, usado do Raspberry Pi, pertence a família CORTEX-A e é bastante interessante para o uso em sistemas embarcados. A frequência de operação pode ser aumentada para até 1Ghz sem a necessidade do uso de um dissipador de calor. Seu consumo de energia é de aproximadamente 0.625mW/Ghz. Como pode ser observado na Figura 15, o processador conta com arquitetura de núcleo único, tendo interfaces para debug, controlador de coprocessador, instruções, dados, *Direct Memory Access* (DMA) e periféricos.

O Raspberry pode ser utilizado para armazenar um servidor, pois ele trabalha com processamento sobre um sistema operacional que executa uma ação de entrada, saída e armazenamento. Diferente dos microcontroladores que trabalham com ciclo de clock sem o controle interno de um sistema operacional (CRUZ; LISBOA, 2014). A fundação Raspberry desenvolveu uma distribuição e *open-source* em Linux embarcado chamada Raspbian, que nada mais é que um sistema operacional composta por pacotes de software com desenvolvimento ativo e constate de atualizações com ênfase na melhoria de estabilidade e desempenho.

Figura 15 – Diagrama funcional do processador ARM1176JZF-S.



Fonte: (ARM..., 2019).

2.6 INFORMAÇÕES ELETRÔNICAS DO USUÁRIO

A internet, representa a maior das Tecnologias da informação e da comunicação (TIC) (MUSEN; BEMMEL, 1997). De acordo com o documento do Conselho Econômico e Social das Nações Unidas (ONU), no contexto da saúde, as TIC vêm contribuir para a saúde dos cidadãos, uma vez que servem para melhorar os processos de tomada de decisões, contribuindo para as trocas de experiências de domínios específicos e também aos cuidados pessoais, bem como a assistência qualificada e igualmente, para reforçar a eficácia das instituições de saúde. Contribuem, ainda para o desenvolvimento do sistemas de informação de organizações de saúde, para a criação e implementação de redes públicas e particulares de saúde, para a construção de sistemas expertos, entre outros (PINTO, 2006).

O Protuário Eletrônico do Paciente (PEP), é um exemplo desse contexto. No Brasil existem diretrizes para aplicação dessa tecnologia como Manual de Certificação para Sistema de Registro Eletrônico em Saúde da Sociedade Brasileira de Informática em Saúde (SBIS) (LEÃO et al., 2009), as resoluções *n.º* 1821/2007 e 1638/2002 do Conselho Federal de Medicina (CFM) (TONELLO; NUNES; PANARO, 2013) e as leis brasileiras (Constituição, Código Penal e Civil).

No ponto de vista de Sabatini (SABATINI, 2002), com os dispositivos móveis, o Pron- tuário do Paciente, antes um documento passivo, difícil de ser entedido e distante do paciente, passou a ser percebido como "um instrumento ativo, uma central de serviços de informação, um promotor de saúde e de prevenção de problemas, e um educador de pacientes e divulgador de informações confiáveis sobre medicina e saúde". Resumindo, o prontuário, que antes tinha fun-

ção de documentar as doenças e informações da saúde do paciente hoje, se tornou complexo, passando a ter um papel de fundamental importância na sociedade contemporânea, tendo como principal funções o compartilhamento de informações entre diferentes profissionais.

2.7 SERVIDOR E CLIENTE

O modelo de comunicação utilizado na internet é baseado em duas atribuições principais para os elementos envolvidos na comunicação: cliente e servidor. O servidor deve estar sempre disponível, à espera da iniciativa do cliente. O cliente aciona o servidor sempre que precisa fazer a comunicação. Esse princípio difere essencialmente dos modelos de comunicação. No modelo cliente-servidor, a comunicação sempre se inicia no cliente.

O cliente é um software, normalmente acionado por um usuário, razão pela qual é comum que tenha uma interface gráfica amigável. Cabe ao cliente iniciar a comunicação com o servidor, seja acionada diretamente pelo usuário ou de forma automática, em resposta a um evento ou uma ação externa.

O servidor é também um software que mantém uma porta de comunicação aberta à espera do cliente. Sua localização, seja pelo endereço ou nome, deve ser conhecida por todos os clientes que querem acessá-lo. Um servidor pode receber um grande número de solicitações simultâneas de clientes, por isso, normalmente, executa em um computador de alto desempenho. Por causa da demanda de alto desempenho, os computadores de alto desempenho são denominados de servidores, apesar de o papel de servidor ser desempenhado por um software. Isso não impede que um dispositivo de *Internet of things* (IoT), com o Raspberry pi, execute a função de servidor.

REFERÊNCIAS

- AIKENS, James E et al. Diabetes self-management support using mhealth and enhanced informal caregiving. **Journal of Diabetes and its Complications**, Elsevier, v. 28, n. 2, p. 171–176, 2014. Citado na página 15.
- ALCAIM, Abraham; SANTOS, C Dos. **Fundamentos do processamento de sinais de voz e imagem**. Rio de Janeiro: Editora Interciência, 2011. Citado 2 vezes nas páginas 14 e 28.
- ALSOS, Ole Andreas; DAS, Anita; SVANÆS, Dag. Mobile health it: The effect of user interface and form factor on doctor–patient communication. **International journal of medical informatics**, Elsevier, v. 81, n. 1, p. 12–28, 2012. Citado na página 15.
- ARM Processors. 2019. Disponível em: <<https://developer.arm.com>>. Acesso em: 24 set. 2019. Citado 3 vezes nas páginas 31, 32 e 33.
- BEHLAU, Mara; PONTES, Paulo; MORETI, Felipe. **Higiene vocal: cuidando da voz**. [S.l.]: Thieme Revinter Publicações LTDA, 2018. Citado na página 17.
- BEHLAU, Mara et al. **Epidemiology of voice disorders in teachers and nonteachers in Brazil: prevalence and adverse effects**. 2012. 665.e9 p. Citado na página 14.
- BENAVIDES, Ana Montero et al. Analysis of voice features related to obstructive sleep apnoea and their application in diagnosis support. **Computer Speech & Language**, Elsevier, v. 28, n. 2, p. 434–452, 2014. Citado na página 14.
- BOERSMA, Paul. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In: AMSTERDAM. **Proceedings of the institute of phonetic sciences**. [S.l.], 1993. v. 17, n. 1193, p. 97–110. Citado na página 28.
- CATANI, Guilherme Simas do Amaral; CARVALHO, Bettina. **Proposed computerized protocol for epidemiological study of patients undergoing microsurgery of the larynx**. [S.l.]: Thieme Publicação Ltda, 2012. 346–352 p. Citado na página 14.
- CHEVEIGNÉ, Alain De; KAWAHARA, Hideki. Yin, a fundamental frequency estimator for speech and music. **The Journal of the Acoustical Society of America**, ASA, v. 111, n. 4, p. 1917–1930, 2002. Citado na página 29.
- CORRÊA, Maria Cristina Silva Montenegro. Anatomia e fisiologia. 2016. Citado 2 vezes nas páginas 25 e 26.
- CRUZ, Ariadne Arrais; LISBOA, Emerson Fausto. Webhome—automação residencial utilizando raspberry pi. **Revista Ciência e Tecnologia**, v. 17, n. 31, 2014. Citado na página 32.
- DAJER, Maria Eugenia. **Análise de sinais de voz por padrões visuais de dinâmica vocal**. 2010. Tese (Doutorado) — Universidade de São Paulo, 2010. Citado na página 21.
- GARGENTA, Marko. **Learning android**. [S.l.]: "O'Reilly Media, Inc.", 2011. Citado na página 30.
- GERHARD, David. **Pitch extraction and fundamental frequency: History and current techniques**. [S.l.]: Department of Computer Science, University of Regina Regina, Canada, 2003. Citado na página 28.

GLAUBER, Nelson. **Dominando o Android: do básico ao avançado**. [S.l.]: Novatec Editora, 2014. Citado 2 vezes nas páginas 29 e 30.

GÓMEZ-PUERTA, José A et al. Laryngeal assessment by videolaryngostroboscopy in patients with rheumatoid arthritis. **Reumatología Clínica (English Edition)**, Elsevier, v. 10, n. 1, p. 32–36, 2014. Citado na página 14.

HILLMAN, Robert E et al. **Ambulatory monitoring of disordered voices**. [S.l.]: Los Angeles, 2006. 795–801 p. Citado na página 15.

LEÃO, Beatriz de Faria et al. Manual de certificação para sistemas de registro eletrônico em saúde (s-res). 2009. Citado na página 33.

LECHETA, Ricardo R. **Google Android-3ª Edição: Aprenda a criar aplicações para dispositivos móveis com o Android SDK**. [S.l.]: Novatec Editora, 2013. Citado na página 30.

LIMA, Welton Dias. Android e a influência do sistema operacional linux. **Tecnologias em Projeção**, v. 8, n. 1, p. 100–111, 2017. Citado na página 31.

MEHTA, Daryush D et al. **Smartphone-based detection of voice disorders by long-term monitoring of neck acceleration features**. Cambridge, 2013. 1–6 p. Citado na página 14.

MOORE, Keith L; DALLEY, Arthur F. **Anatomía con orientación clínica**. [S.l.]: Ed. Médica Panamericana, 2009. Citado 2 vezes nas páginas 19 e 20.

MUSEN, Mark A; BEMMEL, Jan H van. **Handbook of medical informatics**. [S.l.]: Bohn Stafleu Van Loghum Houten, the Netherlands, 1997. Citado na página 33.

PELLANDA, Eduardo Campos. **Comunicação móvel no contexto brasileiro**. 2009. 11 p. Citado na página 15.

PEREIRA, Lucio Camilo Oliva; SILVA, Michel Lourenço da. **Android para desenvolvedores**. [S.l.]: Brasport, 2009. Citado na página 30.

PINHEIRO, Paulo Ricardo Guedes. Ciclos evolutivos das telecomunicações. **Tutorial Teleco**, v. 7, n. 06, 2004. Citado na página 17.

PINHO, Sílvia Maria Rebelo; KORN, Gustavo Polacow; PONTES, Paulo. **Músculos intrínsecos da laringe e dinâmica vocal**. [S.l.]: Thieme Revinter Publicações LTDA, 2019. Citado na página 21.

PINTO, Virgínia Bentes. Prontuário eletrônico do paciente: documento técnico de informação e comunicação do domínio da saúde 10.5007/1518-2924.2006 v11n21p34. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 11, n. 21, p. 34–48, 2006. Citado na página 33.

RABINER, Lawrence et al. A comparative performance study of several pitch detection algorithms. **IEEE Transactions on Acoustics, Speech, and Signal Processing**, IEEE, v. 24, n. 5, p. 399–418, 1976. Citado na página 28.

RABINER, Lawrence R; SCHAFER, Ronald W. **Digital processing of speech signals**. Nova Jersey: Prentice-hall Englewood Cliffs, 1978. v. 100. Citado na página 14.

RASPBERRY Pi. 2019. Disponível em: <<https://www.raspberrypi.org>>. Acesso em: 24 set. 2019. Citado na página 31.

ROSA, Marcelo de Oliveira. **Laringe digital**. 2002. Tese (Doutorado) — Universidade de São Paulo, 2002. Citado 4 vezes nas páginas 21, 22, 23 e 24.

ROTHFIELD, Kenneth P; RUSSO, Sebastian G. Videolaryngoscopy: should it replace direct laryngoscopy? a pro-con debate. **Journal of clinical anesthesia**, Elsevier, v. 24, n. 7, p. 593–597, 2012. Citado na página 14.

SABATINI, RME. Preservando a confiabilidade médica na internet. **Revista Check-up**, 2002. Citado na página 33.

SOUCHARD, Philippe-Emmanuel. **Respiração**. [S.l.]: Summus Editorial, 1987. Citado na página 20.

SOUCHARD, Philippe Emmanuel. **O diafragma**. [S.l.]: Summus editorial, 1989. Citado na página 20.

STARCOUNTER GLOBASTATS. Operating System Market Share Worlwide. 2018. Disponível em: <<https://gs.statcounter.com/os-marker-share>>. Acesso em: 13 nov. 2018. Citado na página 29.

TANENBAUM, Andrew S. **Sistemas operacionais modernos 3. ed.** [S.l.]: São Paulo, SP: Pearson Prentice Hall, 2009. Citado na página 29.

TONELLO, Izângela Maria Sansoni; NUNES, Risia Meressa da Silva; PANARO, Aline Peres. Prontuário do paciente: a questão do sigilo e a lei de acesso à informação. **Informação & informação**, v. 18, n. 2, p. 193–210, 2013. Citado na página 33.

TORRES, Begoña. Anatomía funcional de la voz. **Capítulo 1 del libro: Medicina del Canto**. URL: <http://www.medicinadelcant.com/cast/llibre.htm#>, Josep Rumbau i Serra, 2007. Citado 6 vezes nas páginas 17, 18, 19, 21, 24 e 27.

VIEIRA, Júlio Doin. Laringe: Falsas cordas vocais e as cordas verdadeiras. **Arquivos Catarinenses de Medicina**, v. 32, n. 1, 2003. Citado na página 25.