

UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Maria Helena Franciscatto

**SISTEMA DE RECOMENDAÇÃO PARA TRIAGEM DE DESORDENS  
DOS SONS DA FALA INFANTIL BASEADO EM UM MODELO DE  
CONSCIÊNCIA DE SITUAÇÃO**

Santa Maria, RS  
2018

**Maria Helena Franciscatto**

**SISTEMA DE RECOMENDAÇÃO PARA TRIAGEM DE DESORDENS DOS SONS DA FALA INFANTIL BASEADO EM UM MODELO DE CONSCIÊNCIA DE SITUAÇÃO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação, Área de Concentração em Computação, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Ciência da Computação**.

ORIENTADOR: Prof. João Carlos Damasceno Lima

COORIENTADORA: Prof.<sup>a</sup> Iara Augustin

Santa Maria, RS  
2018

Franciscatto, Maria Helena

Sistema de Recomendação para Triagem de Desordens dos Sons da Fala Infantil Baseado em um Modelo de Consciência de Situação / Maria Helena Franciscatto.- 2018.

93 p.; 30 cm

Orientador: João Carlos Damasceno Lima

Coorientadora: Iara Augustin

Dissertação (mestrado) - Universidade Federal de Santa Maria, Centro de Tecnologia, Programa de Pós-Graduação em Ciência da Computação, RS, 2018

1. Consciência de Situação 2. Sistemas de Recomendação  
3. Desordens dos Sons da Fala I. Damasceno Lima, João Carlos II. Augustin, Iara III. Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB 10/1728.

---

©2018

Todos os direitos autorais reservados a Maria Helena Franciscatto. A reprodução de partes ou do todo deste trabalho só poderá ser feita mediante a citação da fonte.

E-mail: mariafranciscatto@gmail.com

**Maria Helena Franciscatto**

**SISTEMA DE RECOMENDAÇÃO PARA TRIAGEM DE DESORDENS DOS SONS DA FALA INFANTIL BASEADO EM UM MODELO DE CONSCIÊNCIA DE SITUAÇÃO**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação, Área de Concentração em Computação, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Ciência da Computação**.

**Aprovado em 10 de julho de 2018:**

---

**João Carlos Damasceno Lima, Dr. (UFSM)**  
(Presidente/Orientador)

---

**Iara Augustin, Dra. (UFSM)**  
(Coorientadora)

---

**Roseclea Duarte Medina, Dra. (UFSM)**

---

**Cristiano André da Costa, Dr. (Unisinós) (videoconferência)**

Santa Maria, RS  
2018

## AGRADECIMENTOS

*Agradeço, em primeiro lugar, à minha família, minha irmã, meus tios e avós, especialmente meus pais, Sandra e Rudiberto, que estiveram do meu lado em todos os momentos, me ouvindo e aconselhando e que, mesmo com a distância, nunca deixaram que faltasse nada.*

*Agradeço ao meu namorado, Cezar, pelo companheirismo e lealdade nos tempos difíceis, por não ter me deixado desanimar nem vacilar, e por sempre me trazer coragem e alegria.*

*Agradeço ao meu orientador Prof. Dr. João Carlos Lima, pela disponibilidade, pelos ensinamentos, pelos momentos de descontração e por toda a ajuda que recebi durante o desenvolvimento desta dissertação. Também agradeço à professora Dra. Iara Augustin, pelo conhecimento a mim transmitido e pelas orientações quando tive dúvidas.*

*Agradeço aos professores Dr. Celio Trois e Dr. Vinicius Maran, pelo auxílio nas pesquisas, nas análises, na escrita e na revisão de artigos, por terem me proporcionado um grande aprendizado ao longo desta caminhada.*

*Agradeço às professoras Dra. Marcia Keske Soares, Dra. Karina Carlesso Pagliarin e ao graduando Augusto Moro, pelas informações e esclarecimentos na área da fonoaudiologia, e pela disponibilização dos dados que foram essenciais para a conclusão deste trabalho.*

*Agradeço aos colegas do Grupo de Sistemas de Computação Móvel (Isadora, Milene, William, Renan e Un Hee) e aos demais colegas e amigos do mestrado (Renata, Kadico, Rafael, Leonardo e Bernardo) pelos cafés, momentos e dúvidas compartilhados, pela alegria que me proporcionaram nesses anos. Agradeço, também, aos amigos da minha cidade natal, Patrícia, Jéssica, Guilherme Busatto, Andriely, Guilherme Argenta e Marcelo, pela amizade que me é tão importante, e por terem feito a minha caminhada mais leve.*

*Agradeço a todos que de alguma forma contribuíram para a conclusão desta dissertação. E, finalmente, agradeço a Deus por todas as graças recebidas e pela Sua presença indiscutível na minha vida.*

*Sê humilde para evitar o orgulho, mas  
voa alto para alcançar a sabedoria.*

*(Santo Agostinho)*

## RESUMO

### SISTEMA DE RECOMENDAÇÃO PARA TRIAGEM DE DESORDENS DOS SONS DA FALA INFANTIL BASEADO EM UM MODELO DE CONSCIÊNCIA DE SITUAÇÃO

AUTORA: Maria Helena Franciscatto

ORIENTADOR: João Carlos Damasceno Lima

COORIENTADORA: Iara Augustin

Consciência de Situação envolve a correta interpretação de uma situação, permitindo que um sistema responda ao ambiente observado e forneça suporte à tomada de decisões em muitos domínios de sistemas. A fonoaudiologia é um exemplo de domínio onde a consciência de situação pode proporcionar benefícios, uma vez que os profissionais devem monitorar o paciente para realizar ações terapêuticas. Ainda, o Raciocínio Baseado em Casos (CBR) pode contribuir para a área, visto que utiliza conhecimento prévio para resolução de problemas. No entanto, poucas propostas na área realizam interpretação de uma situação para auxiliar nessas tarefas, assim como a metodologia CBR é pouco abordada. Neste sentido, é proposto um sistema de recomendação baseado em um modelo de consciência de situação, que utiliza raciocínio baseado em casos para fazer recomendações úteis ao fonoaudiólogo com base no conhecimento adquirido sobre o paciente. O sistema de recomendação objetiva auxiliar a tomada de decisão clínica, e é descrito por meio de uma arquitetura composta por cinco módulos distintos, sendo que, destes, três são especificados na presente dissertação - *Percepção*, *Classificação* e *Compreensão*. Além de basear-se em Consciência de Situação e Raciocínio Baseado em Casos, a arquitetura incorpora conceitos da Teoria do Desenvolvimento Cognitivo como forma de incrementar cognitivamente a proposta e as recomendações ao profissional de fonoaudiologia. Nos testes realizados com relação ao Módulo de Classificação, o sistema foi capaz de classificar pronúncias infantis com uma precisão acima de 92%. No Módulo de Compreensão, padrões de erro na pronúncia infantil foram identificados, permitindo a recomendação de estratégias de reparo ao fonoaudiólogo e fornecendo suporte nos processos de triagem e planejamento terapêutico.

**Palavras-chave:** Consciência de Situação. Raciocínio Baseado em Casos. Teoria do Desenvolvimento Cognitivo. Sistemas de Recomendação. Desordens dos sons da fala. Reconhecimento de fala.

## ABSTRACT

### RECOMMENDER SYSTEM FOR SCREENING OF SPEECH SOUND DISORDERS BASED ON A SITUATION AWARENESS MODEL

AUTHOR: Maria Helena Franciscatto  
ADVISOR: João Carlos Damasceno Lima  
CO-ADVISOR: Iara Augustin

Situation Awareness (SA) involves the correct interpretation of situations, allowing systems to respond to the observed environment and providing support for decision making in many systems domains. Speech therapy is an example of domain where situation awareness can provide benefits, since practitioners should monitor the patient in order to perform therapeutic actions. Also, Case-Based Reasoning (CBR) may contribute to the area, since it uses prior knowledge to solve problems. However, few proposals in the area perform interpretation of a situation to aid in these tasks, just as the CBR methodology is little addressed. For this reason, a recommendation system based on a situation awareness model is proposed, which uses CBR to make useful recommendations to the Speech-Language Pathologists (SLPs) based on the knowledge acquired about the patient. The recommendation system aims to aid clinical decision making, and it is described through an architecture composed of five distinct modules, of which three are specified in the present dissertation - *Perception*, *Classification* and *Comprehension*. In addition to being based on Situation Awareness and Case-Based Reasoning, the architecture incorporates concepts from the Cognitive Development Theory, as a way of cognitively increasing the proposal and making recommendations that are useful to the SLPs. In the tests performed with respect to the Classification Module, the system was able to classify infant pronunciations with an accuracy above 92%. In the Comprehension Module, error patterns in infant pronunciation were identified, allowing the recommendation of repair strategies to the SLPs and providing support in the screening and therapeutic planning processes.

**Keywords:** Situation Awareness. Case-Based Reasoning. Cognitive Development Theory. Speech sound disorders. Speech recognition.



## LISTA DE FIGURAS

Figura 2.1 – Representação de contexto de acordo com o nível de abstração. ....	22
Figura 2.2 – Modelo de Consciência de Situação. ....	25
Figura 2.3 – Esquema para conceitos de assimilação e acomodação. ....	28
Figura 2.4 – Ciclo da metodologia CBR. ....	29
Figura 3.1 – Etapas do processo de seleção do estudo. ....	35
Figura 4.1 – Arquitetura do sistema de recomendação. ....	52
Figura 4.2 – Percepção Situacional e Módulo de Classificação. ....	55
Figura 4.3 – Módulo de Compreensão. ....	56
Figura 4.4 – Módulo de Assimilação ....	57
Figura 4.5 – Módulo de Acomodação. ....	59
Figura 5.1 – Banco de dados utilizado para gerenciar o corpus de fala. ....	62
Figura 5.2 – Espectrogramas referentes à palavra <i>travesseiro</i> falada corretamente (à esquerda) e incorretamente (à direita). ....	68
Figura 5.3 – Exemplo de operação do método LBP. ....	69
Figura 5.4 – Tabela do banco de dados contendo palavras alvo e estratégias de reparo. ....	76
Figura 5.5 – Recuperação de estratégias de reparo com base em um caso de entrada. ....	77

## LISTA DE GRÁFICOS

Gráfico 3.1 – Porcentagem de critérios de Consciência de Situação suportados pelas propostas. ....	45
Gráfico 3.2 – Total de ocorrências dos métodos nos trabalhos relacionados. ....	48
Gráfico 5.1 – Porcentagem de concordância (1-1, 0-0) e discordância (0-1, 1-0) entre a API <i>Google Cloud Speech</i> (GCS) e o especialista (SLP). ....	65
Gráfico 5.2 – Precisão de classificação a partir dos espectrogramas. ....	70
Gráfico 5.3 – Porcentagem de concordância (1-1, 0-0) e discordância (0-1, 1-0) entre o método Decision Tree (DT) e o especialista (SLP). ....	70
Gráfico 6.1 – Comparativo entre as propostas na área da fonoaudiologia. ....	79
Gráfico 6.2 – Critérios pouco abrangidos por trabalhos relacionados na área da fonoaudiologia. ....	81

## LISTA DE TABELAS

Tabela 2.1 – Objetivos conceituais de sistemas de recomendação.....	32
Tabela 3.1 – String de busca.....	34
Tabela 3.2 – Critérios de inclusão e exclusão.....	35
Tabela 3.3 – Integração de Consciência de Situação em sistemas de suporte à terapia da fala.....	44
Tabela 3.4 – Métodos encontrados em trabalhos relacionados.....	47
Tabela 5.1 – Avaliação de palavras-alvo com <i>Google Cloud Speech</i> .....	64
Tabela 5.2 – Avaliação de palavras-alvo com Decision Tree .....	71
Tabela 5.3 – Estratégias de reparo utilizadas no processo de aquisição da fala .....	73
Tabela 5.4 – Exemplo de categorização de palavras-alvo de acordo com estratégias de reparo .....	75
Tabela 6.1 – Critérios de SA abrangidos por esta dissertação .....	80

## LISTA DE ABREVIATURAS E SIGLAS

<i>1NN</i>	1-Nearest Neighbor
<i>AAL</i>	Ambient Assisted Living
<i>ANN</i>	Artificial Neural Network
<i>API</i>	Application Programming Interface
<i>ASR</i>	Automatic Speech Recognition
<i>ATD</i>	Assistive Technology Device
<i>CART</i>	Classification And Regression Tree
<i>CBR</i>	Case-Based Reasoning
<i>CBST</i>	Computer Based Speech Training
<i>CCD</i>	Cognitive Communication Disorder
<i>CDSS</i>	Clinical Decision Support System
<i>CLIP</i>	Clinical Speech Processing Chain
<i>CMS</i>	Cepstral Mean Subtraction
<i>DNN</i>	Deep Neural Network
<i>DTW</i>	Dynamic Time Warping
<i>EC</i>	Encontro Consonantal
<i>FSG</i>	Finite-State Grammar
<i>GCS</i>	Google Cloud Speech
<i>GMM</i>	Gaussian Mixture Model
<i>GMob</i>	Grupo de Sistemas de Computação Móvel
<i>HFCC</i>	Human Factor Cepstral Coefficient
<i>HMM</i>	Hidden Markov Model
<i>HNN</i>	Hierarchical Neural Network
<i>KDE</i>	Kernel Density Estimation
<i>KMC</i>	k-Means Clustering
<i>kNN</i>	k-Nearest Neighbors
<i>LBP</i>	Local Binary Pattern

<i>LM</i>	Language Model
<i>LMT</i>	Logistic Model Tree
<i>LPC</i>	Linear Predictive Coding
<i>LR</i>	Linear Regression
<i>LogR</i>	Logistic Regression
<i>LS/SVM</i>	LeastSquare Support Vector Machine
<i>MaxEnt</i>	Maximum Entropy
<i>MFCC</i>	Mel Frequency Cepstral Coefficient
<i>MHARS</i>	Mobile Human Activity Recognition System
<i>MIMO</i>	Multi-Input Multi-Output
<i>ML</i>	Machine Learning
<i>MLLR</i>	Maximum Likelihood Linear Regression
<i>MLP</i>	MultiLayer Perceptron
<i>MSG</i>	Modulation SpectroGram
<i>NB</i>	Naive Bayes
<i>NC</i>	Normalized Autocorrelation Coefficient
<i>NLP</i>	Natural Language Processing
<i>PCN</i>	Pronunciation Confusion Network
<i>PEP</i>	Phonological Error Pattern
<i>PLP</i>	Perceptual Linear Prediction
<i>PNN</i>	Probabilistic Neural Network
<i>PPGDCH</i>	Programa de Pós-Graduação em Distúrbios da Comunicação Humana
<i>PSSLM</i>	Phonologically-Structured Sparse Linear Model
<i>RASTA</i>	log-RelAtive SpecTrAl
<i>RF</i>	Randomic Forest
<i>RT</i>	Regression Tree
<i>SA</i>	Situation Awareness
<i>SLP</i>	Speech-Language Pathologist
<i>SOX</i>	Sound eXchange

<i>SVM</i>	Support Vector Machine
<i>TC</i>	Taxa de Concordância
<i>TD</i>	Taxa de Discordância
<i>TDC</i>	Teoria do Desenvolvimento Cognitivo
<i>WCPM</i>	Words Correct Per Minute
<i>WFST</i>	Weighted Finite-State Transducer

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	CONTEXTUALIZAÇÃO DO TEMA	15
1.2	DELIMITAÇÃO DO TEMA	16
1.3	PERGUNTA DE PESQUISA	16
1.4	METODOLOGIA	17
1.5	OBJETIVOS	17
<b>1.5.1</b>	<b>Objetivo Geral</b>	<b>17</b>
<b>1.5.2</b>	<b>Objetivos Específicos</b>	<b>17</b>
1.6	ORGANIZAÇÃO DO DOCUMENTO	18
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>19</b>
2.1	COMPUTAÇÃO MÓVEL, UBÍQUA E PERVASIVA	19
2.2	CONSCIÊNCIA DE CONTEXTO	20
2.3	CONSCIÊNCIA DE SITUAÇÃO	22
<b>2.3.1</b>	<b>Consciência de Situação na Área da Saúde</b>	<b>25</b>
2.4	TEORIA DO DESENVOLVIMENTO COGNITIVO	26
2.5	RACIOCÍNIO BASEADO EM CASOS	29
2.6	SISTEMAS DE RECOMENDAÇÃO	30
<b>3</b>	<b>REVISÃO SISTEMÁTICA</b>	<b>34</b>
3.1	PLANEJAMENTO	34
3.2	CONDUÇÃO	34
3.3	RELATÓRIO: SISTEMAS DE SUPORTE À TERAPIA DA FALA	36
<b>3.3.1</b>	<b>Integração de Consciência de Situação na Fonoaudiologia</b>	<b>42</b>
<b>3.3.2</b>	<b>Métodos Utilizados em Sistemas no Domínio da Fala</b>	<b>46</b>
3.4	MOTIVAÇÃO E JUSTIFICATIVA	49
<b>4</b>	<b>SISTEMA DE RECOMENDAÇÃO BASEADO EM CONSCIÊNCIA DE SITUAÇÃO</b>	<b>51</b>
4.1	CONTEXTUALIZAÇÃO	51
4.2	ARQUITETURA DE SISTEMA	51
<b>4.2.1</b>	<b>Unidade Móvel</b>	<b>52</b>
<i>4.2.1.1</i>	<i>Módulo de Percepção</i>	<i>53</i>
<b>4.2.2</b>	<b>Unidade Servidor</b>	<b>53</b>
<i>4.2.2.1</i>	<i>Módulo de Classificação</i>	<i>54</i>
<i>4.2.2.2</i>	<i>Módulo de Compreensão</i>	<i>55</i>
<i>4.2.2.3</i>	<i>Módulo de Assimilação</i>	<i>56</i>
<i>4.2.2.4</i>	<i>Módulo de Acomodação</i>	<i>58</i>
<b>5</b>	<b>AVALIAÇÃO E RESULTADOS</b>	<b>61</b>
5.1	MÓDULO DE PERCEPÇÃO SITUACIONAL	61
<b>5.1.1</b>	<b>Coleta das Amostras de Fala</b>	<b>61</b>
5.2	MÓDULO DE CLASSIFICAÇÃO	62
<b>5.2.1</b>	<b>Avaliação utilizando a API Google Cloud Speech</b>	<b>63</b>
<b>5.2.2</b>	<b>Pré-processamento dos dados de fala</b>	<b>65</b>
<b>5.2.3</b>	<b>Geração de Espectrogramas</b>	<b>67</b>
<b>5.2.4</b>	<b>Extração de Características</b>	<b>68</b>
<b>5.2.5</b>	<b>Resultados de Classificação</b>	<b>69</b>

5.3	MÓDULO DE COMPREENSÃO.....	72
5.3.1	<b>Agrupamento de Padrões</b> .....	<b>72</b>
5.3.2	<b>Recuperação de Casos</b> .....	<b>74</b>
5.3.3	<b>Recomendação de Estratégias de Reparo</b> .....	<b>76</b>
6	<b>COMPARAÇÃO COM TRABALHOS RELACIONADOS</b> .....	<b>79</b>
7	<b>CONSIDERAÇÕES FINAIS</b> .....	<b>83</b>
7.1	CONTRIBUIÇÕES DA PESQUISA.....	84
7.2	LIMITAÇÕES DA PESQUISA.....	85
7.3	TRABALHOS FUTUROS .....	85
7.4	PUBLICAÇÕES.....	86
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>87</b>



# 1 INTRODUÇÃO

Como introdução ao processo de desenvolvimento desta dissertação, o tema da pesquisa é contextualizado e delimitado nos tópicos seguintes. Na sequência, são apresentadas as principais metas almeçadas, as perguntas de pesquisa a serem respondidas, bem como a metodologia utilizada para o desenvolvimento da proposta. Ao final do capítulo, descreve-se a organização do presente documento.

## 1.1 CONTEXTUALIZAÇÃO DO TEMA

A popularização de dispositivos móveis e de sensoriamento permitiu o amadurecimento de aplicações conscientes de situação nas pesquisas recentes. Através da Consciência de Situação, tais aplicações e sistemas são capazes de agir de forma inteligente em uma grande variedade de domínios, como, por exemplo, setores médicos, setores de controle e automação, controle de tráfego aéreo e, especialmente, os setores de segurança e serviços de emergência. Obter Consciência de Situação, entretanto, é uma tarefa complexa que envolve alto nível de dinâmica: é preciso coletar informações sobre o ambiente (geralmente de diferentes fontes), tomar decisões baseadas nessa coleta e no conhecimento, agir de acordo com as decisões, coletar feedback do ambiente para complementar o conhecimento existente e, desta maneira, tomar decisões melhores no futuro (KOKAR; ENDSLEY, 2012).

Muitos sistemas conscientes de situação têm sido apresentados em ambientes de cuidados com a saúde, no que se refere à prevenção de riscos e monitoramento do paciente. Podemos citar, por exemplo, o uso de Dispositivos de Tecnologia Assistiva (ATDs, ou *Assistive Technology Devices*) conscientes de situação, responsáveis por observar o indivíduo através de sensores e fornecer, se necessário, medidas convenientes conforme a análise de seu comportamento. Em casos como este, as formas computacionais desempenham a função de “mediadoras entre o usuário e o ambiente” (KIRSTE, 2001), a fim de ajudar o indivíduo a atingir seus objetivos de forma independente.

No domínio de terapia da fala, percebe-se que existem poucas propostas que utilizam modelagem de conhecimento para melhorar tarefas como diagnóstico, planejamento de terapia e intervenção terapêutica (CHUCHUCA-MÉNDEZ et al., 2016). Segundo Frost e Gabrielli (2013), “os clínicos devem ser apoiados na obtenção de um bom nível de consciência situacional sobre a condição de seu paciente no momento do atendimento, quando as decisões precisam ser tomadas.” Neste sentido, os sistemas conscientes de situação representam ferramentas impactantes que devem auxiliar no processo de diagnóstico e decisão clínica.

Da mesma forma que a Consciência de Situação é pouco abordada em domínio fonoaudiológico, também encontram-se poucas propostas na área que fazem uso da metodologia de Raciocínio Baseado em Casos. Essa metodologia tem provado ser uma escolha favorável em contextos de saúde, já que possui boas capacidades de aprendizagem e sua habilidade de resolver problemas melhora à medida que novos casos são armazenados em um histórico ou base de dados (HUSAIN; PHENG, 2010).

Nesse sentido, o presente trabalho apresenta um sistema de recomendação que faz uso de Raciocínio Baseado em Casos e Consciência de Situação para auxiliar os profissionais de fonoaudiologia em tarefas terapêuticas, de triagem e diagnóstico. A fim de equipar o sistema de recomendação com habilidades de planejamento, são incorporados conceitos da Teoria do Desenvolvimento Cognitivo, proposta por Piaget (1970a). Por meio destes conceitos (Assimilação e Acomodação), o sistema é capaz de se adaptar e se organizar para atingir uma boa performance em uma grande variedade de situações.

Esta dissertação foi desenvolvida no Grupo de Sistemas de Computação Móvel (GMob) da Universidade Federal de Santa Maria e objetiva modelar uma arquitetura de sistema de recomendação consciente de situação. O foco do trabalho é especificar os Módulos de Percepção, Classificação e Compreensão da arquitetura, apontando as avaliações realizadas e resultados encontrados a partir do estudo.

## 1.2 DELIMITAÇÃO DO TEMA

O problema de pesquisa abordado nesta dissertação possui natureza investigatória, ou seja, objetiva coletar conhecimento sobre como as teorias aplicadas podem contribuir para a resolução de um determinado problema. A partir desse objetivo, e considerando que a literatura relacionada ao domínio da fala pouco menciona as teorias utilizadas neste trabalho, salienta-se que esta dissertação não visa desenvolver um produto tecnológico a ser usado em ambiente clínico e fonoaudiológico, mas o desenvolvimento de uma arquitetura de sistema validada por meio de estudo de caso e prototipação. O trabalho posiciona-se na fase de acumulação de conhecimento para que se estabeleça um sistema de recomendação adequadamente fundamentado.

## 1.3 PERGUNTA DE PESQUISA

Esta dissertação propõe-se a responder a seguinte pergunta de pesquisa: Como seria um sistema de recomendação com a finalidade de oferecer suporte aos profissionais da fonoaudiologia na triagem de desordens da fala infantil?

A hipótese defendida é que este sistema pode ser modelado utilizando três teorias (Teoria do Desenvolvimento Cognitivo, Raciocínio Baseado em Casos e Consciência da Situação) que, associadas, podem trazer bons resultados à fonoaudiologia. Com relação ao desenvolvimento cognitivo infantil, Jean Piaget (PIAGET, 1975) apresenta um desenvolvimento baseado em quatro estágios e os conceitos de assimilação e acomodação, processos pelos quais a mente toma informações do ambiente e as reorganiza para estabelecer um estado de equilíbrio. O Raciocínio Baseado em Casos é um método de decisão de resultado a partir da descrição do problema, onde casos antigos são utilizados para resolver novos casos. Já a Consciência de Situação é um modelo que utiliza conceitos de percepção, compreensão e projeção para interpretação de uma situação e tomada de decisões no ambiente observado.

## 1.4 METODOLOGIA

A pesquisa descrita nessa dissertação seguiu o método científico baseado em Estudo de Caso para prova de conceito do modelo de consciência de situação definido. Para tal, foram realizadas reuniões presenciais e virtuais regulares com orientadores e o grupo de área de aplicação do estudo (Fonoaudiologia), do Programa de Pós-Graduação em Distúrbios da Comunicação Humana (PPGDCH). A interação entre grupos de pesquisa possibilitou auxílio nas decisões de pesquisa e permitiu a modelagem de uma arquitetura de sistema, sob a qual baseou-se o Estudo de Caso.

## 1.5 OBJETIVOS

### 1.5.1 Objetivo Geral

O presente trabalho objetiva propor uma arquitetura de sistema de recomendação consciente de situação baseado na Teoria do Desenvolvimento Cognitivo e na metodologia de Raciocínio Baseado em Casos.

### 1.5.2 Objetivos Específicos

A partir do objetivo central, o trabalho tem as finalidades de:

- Realizar uma revisão sistemática da literatura, a fim de arrecadar conhecimento na área;
- Analisar as teorias propostas, de forma a avaliar como elas podem ser utilizadas na modelagem de uma arquitetura de sistema;
- Definir o Estudo de Caso para criação do sistema de recomendação;
- Validar a arquitetura de sistema proposta por meio do Estudo de Caso e prototipação;
- Analisar o desempenho dos métodos aplicados em fase de validação - no que se refere ao auxílio à fonoaudiologia e, especificamente, à triagem de distúrbios da fala infantil - para aceitar ou rejeitar a hipótese de pesquisa;
- Avaliar o impacto que um sistema de recomendação consciente de situação pode trazer para a área de aplicação da proposta;
- Gerar publicações para divulgação dos resultados finais e parciais da pesquisa no meio acadêmico.

## 1.6 ORGANIZAÇÃO DO DOCUMENTO

Esta dissertação encontra-se organizada da seguinte maneira: no Capítulo 2, a fundamentação teórica utilizada na elaboração da proposta é apresentada. O Capítulo 3 apresenta uma revisão sistemática, abrangendo sistemas de suporte à terapia da fala e a integração de consciência de situação nessas propostas. O Capítulo 4 apresenta o sistema de recomendação baseado em consciência de situação, demonstrado através de uma arquitetura e de módulos componentes. No Capítulo 5, são apresentados os métodos de avaliação da proposta e os resultados alcançados em cada módulo da arquitetura. O Capítulo 6 compara a proposta desta dissertação com os trabalhos relacionados descritos no Capítulo 3. Por fim, o Capítulo 7 apresenta as considerações finais sobre o trabalho, incluindo as contribuições da pesquisa e trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

Como forma de promover um maior entendimento acerca do trabalho desenvolvido, este capítulo objetiva explicar conceitos essenciais presentes na modelagem do sistema de recomendação proposto. Para tal, são apresentados os conceitos de Computação Móvel, Ubíqua e Pervasiva. Também, são esclarecidos aspectos relacionados à Consciência de Contexto e Consciência de Situação, bem como sistemas na área da saúde que incorporam Consciência de Situação. São apresentadas a metodologia de Raciocínio Baseado em Casos e a Teoria do Desenvolvimento Cognitivo, teorias fundamentais para o entendimento da proposta. Por fim, são descritos sistemas de recomendação e os tipos existentes.

### 2.1 COMPUTAÇÃO MÓVEL, UBÍQUA E PERVASIVA

A evolução das tecnologias computacionais tem impactado diária e crescentemente nas mais diversas atividades humanas. A exigência por serviços personalizados e inteligentes fez com que a computação voltasse seu foco para formas de integrar sistemas a tarefas cotidianas do usuário, no intuito de promover processos adaptativos adequados. A interação de sistemas computacionais com o usuário foi ganhando destaque e maturidade de tal maneira que a indústria voltou-se a formas de computação mais distribuídas e flexíveis: os computadores em grades e isolados mudaram-se para os nossos escritórios pessoais, para os nossos bolsos, roupas e corpos (LYYTINEN; YOO, 2002). À medida que se deu esta evolução, surgiram na literatura os conceitos de Computação Móvel, Ubíqua e Pervasiva, levando a computação tradicional a um nível superior.

A Computação Móvel, segundo Araujo (2003), “baseia-se no aumento da nossa capacidade de mover fisicamente serviços computacionais conosco”, ou seja, faz uso de comunicação sem fio para que o usuário possa utilizar serviços de maneira móvel, independentemente de sua localização. Este modelo computacional permite expandir as capacidades de comunicação e acesso, entretanto, apresenta a limitação de não se adaptar de acordo com o movimento. Em outras palavras, a Computação Móvel não permite reconhecer informação contextual no ambiente circundante e ajustar seu comportamento de acordo com o contexto compreendido.

Na busca por automação e adaptação baseada em contexto, surge o conceito de Computação Pervasiva, a qual assume uma série de entidades computacionais ou sensoriais invisíveis, que interagem tanto com o usuário quanto com o ambiente em que operam (YE; DOBSON; MCKEEVER, 2012). Trata-se, especificamente, de um modelo de computação caracterizado por aplicações distribuídas em todo lugar e, portanto, incorporadas ao ambiente de forma imperceptível ao usuário. Neste sentido, a Computação Pervasiva

permite que as aplicações possam utilizar dados de contexto do indivíduo e dos elementos ao seu redor para monitorar mudanças no estado do ambiente e fornecer serviços personalizados ou convenientes a uma determinada necessidade.

Para Mark Weiser (1999), entretanto, o poder computacional existente e a forma com que o mesmo é integrado na vida humana diária refletem apenas “um passo de transição para alcançar o real potencial da tecnologia de informação”. Por isso, o autor concebeu uma nova forma de pensar sobre a computação, definindo o conceito de Computação Ubíqua: “o método de melhorar o uso do computador disponibilizando muitos computadores em todo o ambiente físico, mas tornando-os efetivamente invisíveis para o usuário” (WEISER, 1993). O termo abrange um enfoque no ambiente humano natural de forma que os computadores devem desaparecer do plano de fundo. Do ponto de vista de Weiser, esta é a essência da Computação Ubíqua, pois quando a tecnologia desaparece desta maneira, é possível utilizá-la sem pensar e focar em novos objetivos (WEISER, 1999).

A ampla utilização de dispositivos móveis ao longo do tempo e a incorporação de sensores em diversos domínios permitiram a criação de ambientes inteligentes onde a tecnologia é invisível ao usuário, de forma semelhante ao ideal proposto pela Computação Ubíqua. Entretanto, o cenário idealizado por Weiser – onde a computação desaparece ao integrar-se a atividades diárias do usuário e age sem intervenção do mesmo – encontra-se ainda distante da realidade atual. De acordo com Neto (2015), esta distância deve-se ao fato de que “a forma de interação entre os aparatos tecnológicos e os seres humanos ainda demanda tempo para que o usuário aprenda a manusear novas tecnologias”, ou seja, os dispositivos ainda influenciam na forma com que as atividades do usuário são conduzidas, fazendo com que o indivíduo necessite se adequar ao modo de operação tecnológica (NETO, 2015).

A partir dos conceitos explanados, a Computação Pervasiva pode ser considerada a que mais se aproxima da realidade, visto que sua definição engloba dispositivos móveis e demais aparatos tecnológicos atualmente incorporados nos meios sociais. Neste modelo computacional, sistemas podem responder adequadamente a mudanças nos ambientes nos quais são inseridos através da compreensão de dados de contexto.

## 2.2 CONSCIÊNCIA DE CONTEXTO

Contexto é um termo definido classicamente por Dey (2001) como qualquer informação que pode ser usada para caracterizar a situação de uma entidade (pessoa, local ou objeto) que é considerada relevante para a interação entre o usuário e a aplicação, incluindo o próprio usuário e a aplicação (DEY, 2001).

Outra definição é dada por Yau, Wang e Karim (2002), que apontam contexto como “qualquer atributo detectável e relevante de um dispositivo, sua interação com outros dis-

positivos e/ou o seu ambiente circundante em um instante de tempo”. Ainda, contexto pode ser interpretado como um conjunto de aspectos envolvendo o local onde o indivíduo está, com quem ele está, e que recursos de encontram nas proximidades, ou seja, contexto não se trata apenas de localização, mas também inclui questões dinâmicas como iluminação, nível de ruído, conectividade e presença de outros indivíduos no ambiente (SCHILIT; ADAMS; WANT, 1994). A partir de tais definições, pode-se compreender contexto como um conjunto de dados que pode ser detectado e interpretado a fim de promover a interação humano-computador e fornecer comportamento adaptativo.

Uma exemplificação de contexto pode ser dada através do uso de sensores de GPS: as leituras produzidas por estes sensores representam dados brutos, porém, uma vez que estes dados representam uma localização geográfica, podem ser chamados de informações de contexto (PERERA et al., 2014). A partir de informações contextuais capturadas por sensores, é possível abstrair uma ou mais situações de interesse e obter conhecimento de eventos que estão ocorrendo no ambiente observado. Neste sentido, surge o termo Consciência de Contexto, que representa o primeiro passo em uma hierarquia para adquirir consciência de situação.

A discussão sobre consciência de contexto teve seu início através de Schilit e Theimer (1994), que definiram o termo como “a habilidade de aplicações descobrirem e reagirem a mudanças no ambiente em que estão situadas”. Podemos dizer, portanto, que a consciência de contexto observada no atual modelo computacional manifesta vantagens com relação à computação tradicional uma vez que é possível inferir a intenção do usuário, ou seja, pode-se obter uma ação coerente sem que o usuário informe qualquer tipo de dado de entrada a um dispositivo, como ocorre na computação tradicional (LIU; LI; HUANG, 2011).

Dey (2001) aponta que “um sistema é consciente de contexto se ele usa contexto para fornecer informações relevantes e/ou serviços ao usuário”. Assim, um sistema consciente de contexto deve ser capaz de interpretar de maneira adequada os dados capturados por sensores, visto que apresentam alta complexidade. O sistema não deve se deter, portanto, a particularidades como saber a frequência cardíaca de um usuário. É preciso que as informações capturadas sejam analisadas em um nível superior: a partir da frequência cardíaca, saber se o usuário está se exercitando, se está sofrendo um ataque cardíaco, entre outras possibilidades (YE; DOBSON; MCKEEVER, 2012).

Neste sentido, podemos dizer que a informação contextual é analisada em vários níveis, como apresentado na Figura 2.1. O contexto de baixo nível pode ser compreendido como a informação proveniente de sensores físicos, ou seja, aquela adquirida sem qualquer interpretação associada. Por outro lado, a informação contextual de alto nível é aquela que passa por um processo de raciocínio de contexto, visto que o ambiente em que o usuário se encontra inclui uma variedade de elementos de interesse que devem ser usados para o desenvolvimento de aplicações adaptáveis. Por fim, a partir da informação

Figura 2.1 – Representação de contexto de acordo com o nível de abstração.



Fonte: Adaptado de (BETTINI et al., 2010).

contextual de alto nível, é possível abstrair situações ocorrendo no ambiente e as relações entre elas. A identificação de uma situação é componente crucial para que sistemas possam fornecer recursos flexíveis e eficientes ao usuário, considerando suas preferências e necessidades.

### 2.3 CONSCIÊNCIA DE SITUAÇÃO

*Situação* é um termo que abrange uma ampla variedade de definições na literatura. O termo é definido por Ye, Dobson e McKeever (2011) como sendo “uma abstração dos eventos que ocorrem no mundo real, derivada de contexto e hipóteses sobre como os contextos observados se relacionam a fatores de interesse para projetistas e aplicações”. De forma similar, Bettini et al. (2010) apontam situação como “abstrações semânticas provenientes de pistas contextuais de baixo nível, conhecimento humano e interpretação do mundo”. Dey (2001) define situação como “uma descrição dos estados de entidades relevantes”, ao passo que Yau et al. (2003) também fazem sua definição, apontando que situação é “o conjunto de atributos de contextos passados e/ou ações de dispositivos individuais que são relevantes para determinar ações futuras do dispositivo”.

Por sua vez, Consciência de Situação (SA, ou *Situation Awareness*) é um termo que expressa, segundo Endsley (1995), “a percepção dos elementos do ambiente dentro de um volume de tempo e espaço, a compreensão do seu significado e projeção de seus efeitos em um futuro próximo”. Esta definição sugere que, através da consciência de situação,



aplicações e sistemas são capazes de compreender eventos circundantes, utilizando o raciocínio de vários contextos no ambiente. Logo, se um sistema é consciente de situação, ele tem habilidade de projetar respostas no ambiente de forma adaptativa, aprimorando a experiência do usuário em determinado meio e durante determinada atividade.

Embora Consciência de Situação não seja um conceito identificado recentemente, a literatura relacionada somente veio a receber consideráveis estímulos ao final dos anos 80, quando o tema ganhou verdadeira notoriedade (STANTON; CHAMBERS; PIGGOTT, 2001). O termo tem suas origens na história da teoria militar, mais especificamente durante a Primeira Guerra Mundial, na qual Oswald Boelke faz referência à importância de adquirir consciência do inimigo antes que o mesmo pudesse obter consciência semelhante (GILSON, 1995). Desde a identificação do conceito, deu-se um significativo intervalo de tempo até que o termo ganhasse destaque e maturidade, à medida que problemas tangíveis foram constatados e houve a necessidade de aprimorar a percepção situacional. Esta necessidade, inicialmente, surgiu na indústria da aviação, que procurou investigar formas de tomar consciência dos eventos que ocorriam durante o voo, e a partir desta consciência, planejar ações futuras.

O desenvolvimento da consciência situacional ocorreu de forma simultânea ao crescimento de problemas interligados a fatores humanos, uma vez que exigem habilidades de percepção e tomada de decisões. De acordo com Endsley (1995), “os profissionais devem lidar com o desempenho humano em tarefas que são principalmente físicas ou perceptuais, assim como devem considerar o comportamento humano envolvendo tarefas cognitivas de alta complexidade [...]” Hoje, a Consciência de Situação tem sido reconhecida como uma questão importante e ainda não resolvida em muitos domínios diferentes, os quais incluem redes de comunicação móvel, redes sociais, sistemas de segurança e monitoramento, controle de epidemias, transportes inteligentes, entre outros (KOKAR; ENDSLEY, 2012). Sua grande potencialidade está no fato de que ela fornece orientação para tomadas de decisões apropriadas. Além disso, ela representa um componente crítico para o sucesso de ações em sistemas complexos e dinâmicos, onde uma ação mal planejada por levar a resultados graves (OOSTHUIZEN; PRETORIUS, 2015).

De acordo com Feng, Teng e Tan (2009), enquanto a consciência de contexto trata-se de explorar o contexto de um usuário para ajudá-lo a ter uma interação mais eficaz com o sistema e permitir que este sistema altere seu comportamento de acordo com o contexto observado, a consciência de situação “se concentra na modelagem do ambiente de um usuário para ajudá-lo a ser consciente de uma situação atual”. Logo, sistemas conscientes de situação fazem uso de consciência de contexto, e devem ser capazes de lidar com a complexidade de dados e a relação entre eles a fim de exercer controle sobre situações e fornecer benefícios à vida humana.

A modelagem de sistemas conscientes de situação baseia-se, geralmente, nos princípios do modelo proposto por Endsley, na década de 90. O Modelo de Endsley determina

que, para obter consciência situacional, é preciso que três estágios ou níveis sejam alcançados (Figura 2.2). São eles:

- nível 1 - *Percepção*: Etapa básica para obter consciência de situação. Refere-se à percepção de todos os elementos espaciais no ambiente (entidades, objetos, eventos, sistemas, fatores ambientais, etc.) e seus estados atuais (locais, circunstâncias, ações) que devem ser avaliados de acordo com a relevância que possuem para tomadas de decisões. Referente a esta etapa, Endsley e Garland (2000) apontam que “sem percepção básica de informações importantes, as probabilidades de formar uma imagem incorreta da situação aumentam dramaticamente”, ou seja, percepção é o elemento fundamental para que a consciência situacional seja adquirida;
- nível 2 - *Compreensão*: Este nível trata da capacidade de estabelecer relação entre informações e compreender a situação, ou seja, entender como os elementos percebidos no ambiente irão impactar nos objetivos do usuário. Uma entidade que possui este nível de consciência situacional é capaz de derivar significados relevantes a partir dos dados percebidos no primeiro nível (ENDSLEY; GARLAND, 2000). O nível de compreensão envolve processos de reconhecimento de padrões, interpretação das informações adquiridas e avaliação, a fim de criar um quadro abrangente do mundo que seja significativo às ações do usuário;
- nível 3 - *Projeção*: O nível de projeção é o nível mais alto no modelo de consciência de situação, e refere-se a projetar ações futuras dos elementos. Ao antecipar ou prever o futuro próximo, um sistema é capaz de fornecer respostas eficazes às eventuais ocorrências em um ambiente. A projeção é alcançada a partir da identificação e dinâmica dos elementos, e da compreensão da situação (níveis 1 e 2). A partir da utilização dessas informações frente ao tempo, é possível determinar como o ambiente operacional pode ser afetado. Ainda, conhecendo o estado do ambiente e/ou entidade, operadores podem realizar tomadas de decisões, ou seja, podem decidir o que fazer sobre a situação e executar quaisquer ações necessárias (ENDSLEY; GARLAND, 2000).

Inúmeras pesquisas têm sido conduzidas envolvendo o estudo e desenvolvimento de sistemas conscientes de situação, métodos de reconhecimento e tomada de decisões. Kurmis et al. (2013), por exemplo, investigaram a identificação de situação em tempo real, a fim de fornecer serviço automatizado e adaptativo em um ambiente veicular sem a intervenção do usuário. Semelhante a esta proposta, Aiello et al. (2014) propõem uma abordagem para ambiente veicular a fim de prevenir mortes acidentais de crianças presas em veículos estacionados, a qual se baseia na detecção do ambiente, no processamento de dados de sensores e na tomada de ações em conformidade. A consciência situacional também está presente, por exemplo, na modelagem de sistemas de recomendação, como

Figura 2.2 – Modelo de Consciência de Situação.



Fonte: Adaptado de (ENDSLEY, 1995).

abordado por Woerndl et al. (2011), que apresentam um modelo para proatividade em sistemas de recomendação móveis e cientes de contexto.

### 2.3.1 Consciência de Situação na Área da Saúde

Na grande área da Computação Consciente de Situação, o cuidado com a saúde vem ganhando destaque ao longo dos anos. O enfoque das pesquisas frequentemente volta-se à idealização de sistemas inteligentes que possam apoiar a autonomia humana, através da prevenção de doenças e tomada de ações ao ser identificada qualquer situação de risco. O cuidado com a população idosa é um exemplo de questão emergente relacionada à saúde: o envelhecimento da população é um dos fatores que apoiam, por exemplo, o desenvolvimento de sistemas para Ambientes de Vivência Assistida (AAL, do inglês *Ambient Assisted Living*), os quais têm a função de monitorar o status de saúde do paciente através da análise de dados recolhidos por meio de tecnologias baseadas em sensores (CULMONE et al., 2014).

Assim como a saúde física e mental dos idosos exige atenção e cuidado constante para prevenção de acidentes e comportamentos de risco, há muitos outros casos em que faz-se necessária uma estratégia de monitoramento adequada. Crianças, por exemplo, são naturalmente ativas e colocam-se, eventualmente, em situações inseguras que podem ocasionar acidentes indesejados (quedas, queimaduras, entre outros). Nestes casos, a tarefa de monitoramento pode ser facilitada por meio de sistemas assistivos conscientes de situação. O emprego destes sistemas computacionais para detecção autônoma de atividades e a correlação destas com parâmetros fisiológicos pode minimizar a necessidade de intervenção do próprio paciente ou do terapeuta na identificação de situações de risco (SCHEUNEMANN, 2016). Ainda, se o comportamento e condições físicas puderem ser analisadas, as informações adquiridas podem ser úteis para a promoção da saúde e prevenção de doenças, como a obesidade recorrente de hábitos sedentários (GOTO et al., 2013).

Como exemplo de pesquisa que envolve modelagem e reconhecimento de situações para cuidados com a saúde, podemos citar o sistema MHARS (*Mobile Human Activity Recognition System*) proposto por Filho et al. (2016), um sistema para monitoramento de pacientes com doenças crônicas, que reconhece atividades do usuário e a intensidade nas quais elas são realizadas. O sistema é capaz de interagir com diferentes tipos de sensores e utiliza dados de aceleração e frequência cardíaca para inferir situações de saúde e realizar ações de resposta.

Outro exemplo provém de Megalingam et al. (2016), que propuseram o sistema MediSuit (composto por dispositivo vestível) para monitoramento contínuo de parâmetros de saúde. São coletados dados de eletrocardiograma, nível de saturação de oxigênio no sangue, nível de inclinação do paciente, sua temperatura e frequência cardíaca. A partir destes dados, o sistema realiza o processamento de sinais e armazenamento de parâmetros de saúde, enviando mensagens de alerta na ocorrência de qualquer anormalidade no estado do usuário.

Por fim, o trabalho de Lopes et al. (2016) também exemplifica consciência de situação na área da saúde. Os autores propuseram uma arquitetura de software consciente de situação onde o profissional de saúde pode monitorar os sinais vitais do paciente e confirmar se o efeito desejado da administração dos medicamentos está sendo alcançado. Se a situação compreendida indicar algum problema, alertas podem ser enviados para uma central de alerta em hospital, bem como para o dispositivo móvel do profissional.

A gestão de ambientes de saúde pode ser considerada, portanto, uma operação crítica, já que os serviços devem ser fornecidos com o objetivo de proporcionar e *manter* o bem estar de um paciente. Sistemas aplicados nestes contextos fazem-se essenciais para prevenção de doenças, comportamentos de risco, e detecção de situações de emergência. Acima de tudo, sistemas conscientes de situação na área da saúde representam ferramentas poderosas no suporte à decisão clínica, ajudando profissionais em tarefas de diagnóstico e planejamento de terapia.

## 2.4 TEORIA DO DESENVOLVIMENTO COGNITIVO

O desenvolvimento cognitivo é um processo que passa por contínua reorganização de estruturas mentais, o qual é explicado por meio de diferentes teorias (GELDERBLUM; KOTZÉ, 2008). Uma dessas teorias provém de Jean Piaget, biólogo suíço, que apontou o desenvolvimento de estruturas lógicas como a base da cognição e aprendizagem do indivíduo.

Epistemologias clássicas como racionalismo e empirismo partem de um princípio comum que aborda a relação existente entre o objeto de conhecimento e o sujeito. Piaget, porém, manifestou sua discordância a partir deste ponto de vista, afirmando que o

conhecimento não se origina dos objetos ou do sujeito, mas sim da interação entre os dois (GARCIA, 1997). A epistemologia genética de Piaget (PIAGET, 1970b) defende que todo indivíduo passa por um processo de desenvolvimento baseado em quatro estágios: sensório-motor (até 2 anos), pré-operatório (2 a 7 anos), operatório-concreto (7 a 11 anos) e operatório-formal (12 anos). Essas quatro etapas estabelecem níveis para a construção da inteligência na criança, caracterizada pela sua capacidade de estabelecer raciocínio dedutivo e abstração.

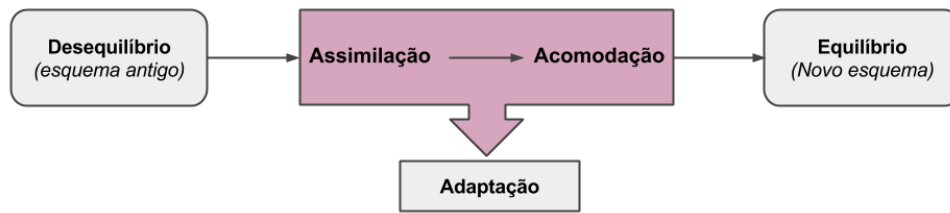
Sobre a origem da inteligência na criança, Piaget esclarece que a inteligência não aparece num determinado momento do desenvolvimento mental como um mecanismo inteiramente montado e distinto dos que o precederam; pelo contrário, apresenta uma continuidade com os processos adquiridos ou mesmo inatos, provenientes da associação habitual e do reflexo (PIAGET, 1975). Assim sendo, de acordo com o teórico, as crianças não herdam capacidades mentais prontas, mas suas atividades intelectuais visam adaptação ao ambiente através de uma construção gradativa (DIAS, 2010). Em resumo, a constituição da inteligência se dá por meio da interação da criança com o mundo - suas experiências - que ampliam sucessivamente suas capacidades cognitivas. Para Piaget, ainda, a inteligência está enraizada em dois atributos biológicos: *organização* e *adaptação*. Organização é a tendência de cada organismo integrar processos ou esquemas em estruturas mais complexas, ao passo que adaptação é a tendência inata da criança interagir com seu ambiente (SIMATWA, 2010), mudando seus esquemas para servirem a uma nova informação.

Este processo de tomar informações do ambiente, organizando-as e adaptando-as de forma a manter um estado de equilíbrio coerente, ocorre de duas maneiras (Figura 2.3): *assimilação* e *acomodação* (GELDERBLOM; KOTZÉ, 2008). *Assimilação* é o processo cognitivo onde o indivíduo integra um novo dado (objeto ou experiência) às suas estruturas cognitivas existentes, aplicando seu conhecimento atual para interpretar a realidade externa. Em outras palavras, na assimilação a criança vai procurar compreender a natureza do novo objeto e tentará aplicá-lo a cada um de seus esquemas mentais (PIAGET, 1975).

A *acomodação*, por outro lado, é o processo no qual ocorre a diferenciação de um esquema existente e inserção de novos elementos entre os que já os constituem (PIAGET, 1975). Gelderblom e Kotzé (2008) afirmam que “a acomodação acontece quando o conhecimento atual não se encaixa no ambiente e novos esquemas de conhecimento devem ser criados”, ou seja, na acomodação, a criança incorpora o objeto assimilado, transformando sua estrutura anterior.

Uma definição complementar e sucinta de assimilação e acomodação é dada por Dias (2010), que declara: "O indivíduo incorpora o objeto enquanto meio de conhecimento, fenômeno este chamado de assimilação. Assim, o sujeito assimila o objeto. Em um segundo momento, o sujeito transforma sua estrutura anterior para incorporar o objeto já assimilado, o que é chamado de acomodação." (DIAS, 2010). Evidencia-se, assim, que a assimilação e a acomodação são os dois pólos de uma interação entre o organismo e o

Figura 2.3 – Esquema para conceitos de assimilação e acomodação.



Fonte: Adaptado de (PIAGET, 1970a).

meio, constituindo a condição para o desenvolvimento cognitivo. Apesar de serem processos distintos, tornam-se cada vez mais complementares um ao outro, à medida em que o pensamento da criança evolui (PIAGET, 1970a).

A Teoria do Desenvolvimento Cognitivo (TDC) de Piaget tem sido aplicada em pesquisas para equipar as metodologias com habilidades de adaptação e cognição. Kuo et al. (2015), por exemplo, utilizaram estratégias de assimilação e acomodação em um sistema multiagente para modificar e adaptar a estrutura cognitiva dos agentes de acordo com o ambiente. No trabalho de Cheu, Ng e Quek (2009), um sistema neural fuzzy é proposto para geração automática de modelos linguísticos. A identificação da estrutura do algoritmo baseia-se na visão cognitiva de Piaget, onde a linguagem é tratada como um aspecto de conhecimento derivado de cognição não linguística, e a inteligência é derivada da ação. De acordo com os autores, o caráter distribuído de uma rede neural lembra a definição de esquema proveniente de Piaget, assim, a ideia de esquema na proposta abrange um dispositivo adaptativo que assimila automaticamente as entradas similares aos esquemas existentes e acomoda novos esquemas como uma função de entradas desconhecidas.

Por fim, como exemplo de aplicação da TDC em sistemas computacionais, podemos citar a proposta de Farias et al. (2013), que desenvolveram um modelo de agente BDI-Fuzzy para a plataforma JSON. A proposta objetiva avaliar qualitativamente e subjetivamente os valores das bolsas sociais originados no fornecimento e recebimento de serviços não econômicos. O modelo se baseia nos conceitos apresentados na teoria de Piaget, que formaliza a dinâmica das interações sociais como um sistema de intercâmbio de serviços entre indivíduos de uma sociedade que, por sua vez, é formada pelas interações que ocorrem entre os indivíduos.

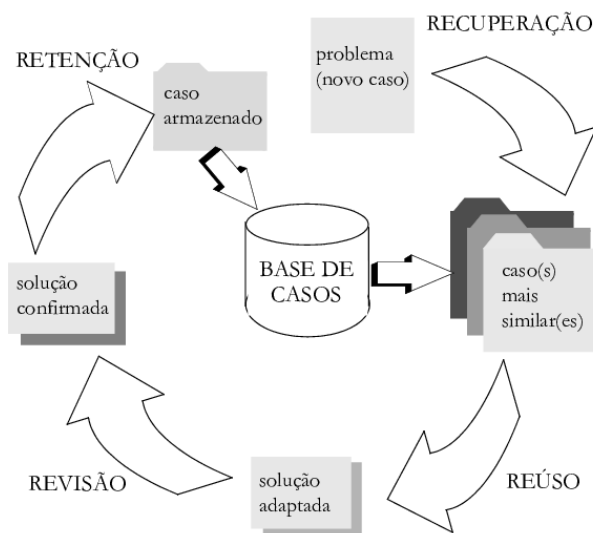
É correto dizer, considerando as pesquisas citadas, que a TDC proposta por Piaget é utilizada não somente em domínio educacional e biológico, mas se apresenta como uma teoria multidisciplinar e crescentemente aplicada na área tecnológica. As múltiplas abordagens da teoria e os conceitos nela integrados são investigados em uma grande quantidade de propostas, a fim de complementar e otimizar as metodologias existentes. Nesta dissertação, especificamente, a TDC objetiva incorporar as habilidades fornecidas pela metodologia de Raciocínio Baseado em Casos, a qual é explanada a seguir.

## 2.5 RACIOCÍNIO BASEADO EM CASOS

O termo Raciocínio Baseado em Casos (CBR, ou *Case-Based Reasoning*) descreve uma metodologia vinda da área de Inteligência Artificial que se inspira no raciocínio humano para resolução de problemas. As origens da metodologia foram estimuladas pelo desejo de entender como as pessoas se recordam de informações e, posteriormente, reconheceu-se que elas comumente resolvem problemas através da lembrança de como resolviam problemas semelhantes no passado (WATSON, 1999).

A definição clássica de Raciocínio Baseado em Casos é dada por Riesbeck e Schank (1989), que apontam que “um raciocínio baseado em casos resolve problemas usando ou adaptando soluções para velhos problemas”, ou seja, um novo problema é resolvido encontrando um caso passado semelhante, e reutilizando a informação e conhecimento deste caso na nova situação problemática (AAMODT; PLAZA, 1994). A metodologia é comumente descrita por um ciclo de quatro atividades, conforme mostra a Figura 2.4: *recuperar* casos que se assemelham à descrição do problema, *reutilizar* uma solução existente para um caso semelhante, *revisar* esta solução de forma a atender o novo problema, e *manter* esta solução, se a mesma for confirmada (WATSON, 1999).

Figura 2.4 – Ciclo da metodologia CBR.



Fonte: (WANGENHEIM; WANGENHEIM; RATEKE, 2013).

De acordo com Aamodt e Plaza (1994), "o paradigma CBR cobre uma gama de diferentes métodos para organizar, recuperar e indexar o conhecimento retido em casos anteriores". De acordo com os autores, casos podem ser (1) mantidos como experiências concretas, (2) usados em conjunto para formar um caso generalizado, (3) armazenados como unidades de conhecimento separadas, (4) divididos em subunidades e distribuídos dentro de uma estrutura de conhecimento ou (5) indexados por um vocabulário

pré-definido, estrutura hierárquica ou de índice.

Sistemas que fazem uso de Raciocínio Baseado em Casos têm sido utilizados com sucesso em domínio médico, devido à suas propriedades inteligentes e humanas para diagnosticar o caso de um paciente (BISWAS et al., 2015). Ahmed, Begun e Funk (2012) apresentam estudos de casos em sistemas CDSS (*Clinical Decision Support System*, ou Sistemas de Suporte à Decisão Clínica) como uma visão geral de pesquisas envolvendo CBR. Em sistemas dessa natureza, a abordagem CBR apresenta bons resultados, uma vez que fornece ao clínico casos antigos similares para ajudá-lo a tomar decisões melhor informadas. A proposta de Raman et al. (2010) também apresenta o uso de CBR em contexto clínico, onde o objetivo é a detecção precoce de câncer de mama. Para isto, os autores desenvolveram um algoritmo de classificação utilizando CBR, a fim de fornecer suporte à decisão clínica na realização da biópsia da mama.

Um sistema que utiliza CBR possui manutenção simplificada, já que conhecimento novo pode ser inserido adicionando novos casos na biblioteca ou no histórico de casos. Além disso, a abordagem é compatível com outros métodos de Inteligência Artificial, de forma que os sistemas que utilizam CBR podem aproveitar outras técnicas para melhorar seu desempenho (AHMED; BEGUM; FUNK, 2012). Pode-se citar, por exemplo, a abordagem híbrida de Husain e Pheng (2010), que utiliza CBR aliado ao Raciocínio Baseado em Regras para recomendação de terapia de bem-estar, através da sugestão de soluções para os usuários. Biswas et al. (2015) desenvolveram um modelo de classificação híbrida baseada em CBR e redes neurais artificiais, onde a rede neural treinada desempenha o papel principal na medida de similaridade e melhora a capacidade de previsão do método CBR. Ainda, algumas propostas utilizaram CBR em conjunto com Consciência de Situação (NWIABU et al., 2012), a fim de investigar uma consciência baseada em experiência ao invés de regras.

No contexto da fala, foco da presente dissertação, o uso do Raciocínio Baseado em Casos também tem sido investigado, por exemplo, no trabalho de Guillén e Ursey (2005), que procurou investigar a detecção de depressão em um indivíduo a partir das propriedades de seu discurso usando uma abordagem CBR. Contudo, no que se refere à terapia da fala e ao suporte ao profissional de fonoaudiologia, percebe-se a falta de propostas neste domínio que integra Raciocínio Baseado em Casos como método para resolução de problemas.

## 2.6 SISTEMAS DE RECOMENDAÇÃO

A definição de sistemas de recomendação é amplamente conhecida da perspectiva de comércio eletrônico (*e-commerce*), na qual representam ferramentas que permitem a personalização rápida e automatizada de sites de vendas, adaptando-os para mostrar pro-



dados de acordo com interesse e necessidade específicos do cliente (SIVAPALAN et al., 2014). Entretanto, sistemas de recomendação são utilizados em uma grande variedade de domínios e representam, de forma geral, uma estratégia de tomada de decisões em espaços de informação complexos, através da sugestão de itens que podem ser valiosos ao usuário de acordo com o conhecimento que se tem sobre este usuário (RASHID et al., 2002).

O processo de recomendação possui três fases distintas: coleta de informações, aprendizagem e a recomendação propriamente dita. A *coleta de informações* é feita com o intuito de gerar um perfil ou modelo para predição de tarefas com base nestas informações. A etapa de *aprendizagem* refere-se à aplicação de um algoritmo de aprendizagem para filtrar as características obtidas na fase de coleta de informações. Por fim, na fase de *recomendação ou previsão* são recomendados itens ou serviços com base na informação coletada, a qual pode ser baseada em memória, em modelo, ou adquirida através da observação do sistema sobre as atividades do usuário (ISINKAYE; FOLAJIMI; OJOKOH, 2015).

De acordo com Aggarwal (2016), sistemas de recomendação são classificados de acordo com o modelo usado para fazer previsões ao usuário (Tabela 2.1). Neste contexto, o autor apresenta os seguintes modelos de sistemas de recomendações:

- **sistemas de recomendação colaborativos:** sistemas de recomendação colaborativos, como o próprio nome diz, são aqueles que utilizam poder colaborativo de classificações feitas por múltiplos usuários para fazer recomendações a um usuário específico (AGGARWAL, 2016). Trabalham com dados provenientes de interações usuário-item (como comportamento de compra ou avaliação) e abrangem, de forma geral, técnicas baseadas em memória e modelo, técnicas de agrupamento, regras de associação, árvores de decisão, redes neurais artificiais, entre outros métodos;
- **sistemas de recomendação baseados em conteúdo:** utilizam algoritmo dependente de domínio que dá mais ênfase à análise de atributos de itens para realizar previsões, ou seja, a recomendação é feita baseada nas características de um ou mais itens avaliados pelo usuário no passado (ISINKAYE; FOLAJIMI; OJOKOH, 2015). Nestes sistemas, são frequentemente usados modelos probabilísticos tais como Naive Bayes, árvores de decisão e redes neurais;
- **sistemas de recomendação baseados em conhecimento:** este tipo de sistema é frequentemente utilizado quando o interesse do usuário é difícil de capturar a partir de atributos de itens e dados de histórico (avaliações passadas, por exemplo). Assim, a recomendação é baseada em *requisitos do usuário* e frequentemente exige o uso de bases de conhecimento contendo dados sobre regras e funções de similaridade a serem utilizadas durante o processo de recuperação (AGGARWAL, 2016);

Tabela 2.1 – Objetivos conceituais de sistemas de recomendação.

Abordagem	Objetivo conceitual	Dados de entrada
Colaborativa	Recomendações baseadas em uma abordagem colaborativa que aproveita as avaliações e ações de múltiplos usuários e do usuário alvo	Avaliações do usuário + avaliações comunitárias
Baseada em conteúdo	Recomendações baseadas no conteúdo (atributos) favorecido em avaliações ou ações passadas	Avaliações do usuário + atributos de itens
Baseada em conhecimento	Recomendações baseadas em especificação explícita do tipo de conteúdo (atributos) desejado	Especificação do usuário + atributos de itens + conhecimento de domínio
Demográfica	Recomendações baseadas em dados demográficos sobre o usuário	Avaliações do usuário + informações demográficas
Híbrida	Recomendações que combinam capacidades de dois ou mais métodos de recomendação em um único modelo	Diferentes fontes de dados de entrada

Fonte: Adaptado de (AGGARWAL, 2016).

- **sistemas de recomendação demográficos:** baseiam-se na hipótese de que usuários com características sociais, culturais e regionais similares compartilham gostos e necessidades também similares (ZANKER et al., 2007). Desta forma, a recomendação é feita com base em informação demográfica do usuário;
- **sistemas de recomendação baseados em conjuntos e híbridos:** representam uma combinação de múltiplas abordagens de recomendação, a fim de superar limitações existentes em uma ou outra abordagem de recomendação e atingir melhores performances (ISINKAYE; FOLAJIMI; OJOKOH, 2015).

Sistemas de recomendação têm sido amplamente abordados na literatura em contextos de saúde a fim de fornecer serviços de bem-estar ao usuário (SEZGIN; OZKAN, 2013). Lee e Kim (2015), por exemplo, propuseram um sistema de recomendação de serviços de saúde baseado em ontologias e CBR, onde os dados são classificados em sintomas, doenças, departamentos, entre outras categorias. A recomendação considera o contexto do paciente coletado em ambiente de computação na nuvem. Tseng et al. (2015) desenvolveram um sistema que age como um assistente virtual de saúde para auxiliar o usuário a entender e lidar com suas condições de saúde. O sistema avalia o estado de saúde do usuário através de exame físico em instituição médica e fornece serviço personalizado em termos de recomendação de dieta e exercícios. Ainda, a proposta de Zhang et al. (2016) representa outra abordagem baseada em recomendação na saúde. Os autores propuseram um sistema de suporte à decisão clínica que recomenda estratégias terapêuticas de acordo com a situação de saúde do paciente, utilizando o algoritmo kNN (*k-Nearest Neighbors*) para realizar o processo de aprendizagem de tratamento.

No contexto da presente dissertação, o sistema de recomendação proposto enquadra-se como *baseado em conhecimento*, visto que reúne conhecimento de domínio para fornecer uma recomendação ao profissional de fonoaudiologia sobre decisões a serem tomadas acerca da situação de um paciente. Mais especificamente, a metodologia de Raciocínio Baseado em Casos realiza a recomendação, por meio de métricas de similaridade do conhecimento de domínio. Em sistemas de recomendação baseados em casos, os casos são usados como alvos ou pontos de ancoragem para orientar a pesquisa em combinação com métricas de similaridade, e o resultado retornado frequentemente é usado como um novo caso alvo (AGGARWAL, 2016).

### 3 REVISÃO SISTEMÁTICA

Nesta seção são abordados trabalhos relacionados a esta dissertação, os quais abrangem sistemas de suporte à terapia da fala. Como forma de selecionar estudos relevantes para avaliar o problema de pesquisa previamente descrito na Seção 1, foi realizada uma revisão sistemática seguindo a metodologia descrita por Staples e Niazi (2007), na qual são realizadas atividades de *planejamento* (gerar o protocolo da revisão), *condução* (identificar, selecionar e sintetizar estudos) e *relatório* (analisar e apresentar os resultados da revisão).

#### 3.1 PLANEJAMENTO

Como fonte de dados da revisão sistemática foram utilizados os motores de busca Scopus, ACM Digital Library, IEEE Xplore e Science Direct. Na Tabela 3.1 a seguir encontra-se a string de busca desenvolvida para a pesquisa, a qual considera os termos nos campos título, resumo e palavras-chave do artigo.

Tabela 3.1 – String de busca.

(TITLE-ABS-KEY "Speech therapy" OR TITLE-ABS-KEY "Apraxia of speech" OR TITLE-ABS-KEY "Phonological system" OR TITLE-ABS-KEY "Situation-aware speech therapy" OR TITLE-ABS-KEY "Speech therapy recommendation" OR TITLE-ABS-KEY "Disfluency detection" OR TITLE-ABS-KEY "Speech-language disorders" OR TITLE-ABS-KEY "Communication disorders" OR TITLE-ABS-KEY "Speech sound disorders" OR TITLE-ABS-KEY "Speech-language diagnostic")
--

Para seleção dos estudos foram elaborados critérios de inclusão (I) e exclusão (E), conforme mostra a Tabela 3.2. A partir da definição destes critérios, foi dado início ao processo de condução da metodologia, a fim de selecionar estudos relevantes de acordo com o objetivo da revisão. A etapa de condução é abordada na sequência.

#### 3.2 CONDUÇÃO

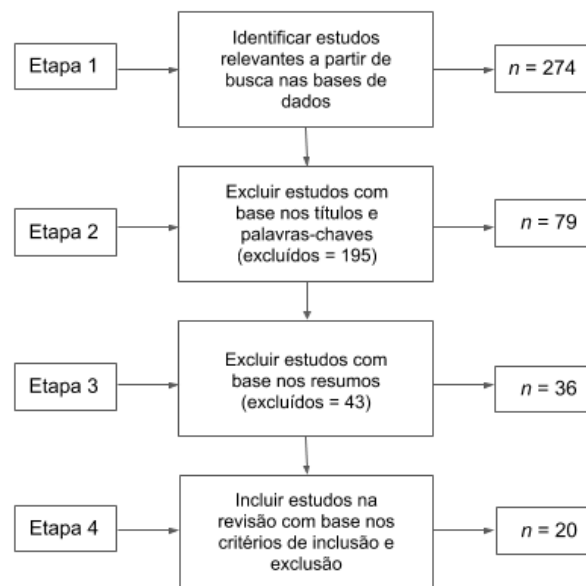
A pesquisa nos motores de busca foi realizada em junho de 2017 e, utilizando ferramentas de filtros, retornou 274 trabalhos. Desse total, 195 foram excluídos após a leitura

Tabela 3.2 – Critérios de inclusão e exclusão.

(I-1) Estudos contendo termos da área da fonoaudiologia no título, resumo ou palavras-chaves.
(I-2) Estudos publicados em workshops, conferências ou periódicos.
(I-3) Artigos completos (4 ou mais páginas).
(E-1) Estudos não relacionados à área da fonoaudiologia.
(E-2) Trabalhos em área de pesquisa que não seja computação.
(E-3) Trabalhos anteriores a 2010.
(E-4) Trabalhos que não estejam em língua inglesa.
(E-5) Estudos não disponíveis integralmente na web.

dos títulos, restando 79 estudos para serem analisados. Destes, 43 foram excluídos após a leitura dos resumos e, dos 36 estudos restantes, 20 foram incluídos nesta revisão com base nos critérios de inclusão e exclusão. O processo de seleção dos trabalhos para esta revisão encontra-se detalhado na Figura 3.1.

Figura 3.1 – Etapas do processo de seleção do estudo.



Fonte: Da autora.

Os 20 trabalhos selecionados na revisão sistemática permitem conhecer o estado da arte de métodos utilizados para reconhecimento de voz e suporte à decisão clínica, e são avaliados segundo critérios de Consciência de Situação, conforme exposto na Tabela 3.3. Por meio das análises realizadas, é possível conhecer quais aspectos de consciência situacional representam desafios no domínio da fonoaudiologia. O tópico seguinte apresenta detalhadamente cada um dos trabalhos selecionados, os métodos observados em cada um, e as principais características abordadas pelos diferentes autores.

### 3.3 RELATÓRIO: SISTEMAS DE SUPORTE À TERAPIA DA FALA

Robles Bykbaev et al. (2016) apresentam um sistema especialista para geração automática de diretrizes para a concepção de terapia fonoaudiológica. A proposta se deve ao fato de que, geralmente, patologistas de fala e linguagem (SLPs, ou *Speech-Language Pathologists*) possuem sobrecarga de trabalho, a qual poderia ser amenizada por meio de um sistema especialista. O sistema se baseia em ontologia e contém entidades que permitem modelar os marcos linguísticos adquiridos durante o desenvolvimento da criança e atividades terapêuticas. Nessa base, o sistema especialista é capaz de selecionar e sugerir as melhores atividades ou estratégias de intervenção para o perfil de um paciente específico, com base em suas habilidades, limitações e necessidades. A validação do sistema utiliza dados clínicos de casos reais de pacientes e elementos de informação de terapia da fala. São implementadas várias interfaces e relatórios web para consulta e gerenciamento de informações utilizando os principais elementos e conceitos de fala e linguagem.

Abad et al. (2013) apontam que pacientes com afasia<sup>1</sup> geralmente possuem dificuldade de se lembrar de nomes ou palavras, o que pode ser tratado através de exercícios de nomeação de palavras. Dessa forma, é proposta uma tecnologia de reconhecimento automático de fala baseada em um reconhecedor híbrido, utilizada em um sistema online. Os dados de entrada consistem do áudio do paciente e suas informações pessoais. A tecnologia de reconhecimento de fala processa o áudio e trabalha com entradas de contexto para decidir se a pronúncia foi correta ou incorreta. Para que haja precisão por parte do módulo de reconhecimento, foi utilizado um banco de dados composto por sessões de terapia de nomeação de palavras, a fim de treinar o detector na adaptação baseada nas particularidades da fala do paciente. A plataforma contém módulos que permitem rastrear a evolução da situação do paciente, além de permitir que os fonoaudiólogos criem exercícios para acesso posterior.

Parnandi et al. (2015) apresentam um sistema para administração remota de terapia fonoaudiológica para crianças com apraxia de fala<sup>2</sup>. O sistema consiste de uma aplicação móvel que grava os enunciados da criança, uma interface de gerenciamento de terapia para que o terapeuta atribua exercícios remotamente e monitore o desempenho da criança, além de um motor de processamento de fala. É realizado treinamento de modelos acústicos a partir de um corpus de fala para que o sistema adapte-se à fala da criança. Na proposta, o terapeuta tem acesso a uma série de relatórios sobre o desempenho do paciente e pode adaptar o programa de treinamento, se necessário.

Outro sistema semelhante chama-se EchoWear, proposto por Dubey et al. (2015),

---

<sup>1</sup>distúrbio de comunicação causado pelo dano de uma ou mais áreas do cérebro, afetando funcionalidades de fala e linguagem tais como compreensão auditiva, produção de fala, expressão oral e fluidez de escrita/leitura.

<sup>2</sup>desordem neurológica da fala infantil que prejudica a capacidade de pronúncia.

um sistema baseado em smartwatch para monitorar remotamente exercícios de fala conforme prescrito por um SLP. O sistema funciona em conjunto com um dispositivo móvel que sinaliza o processo de gravação de fala. O dispositivo móvel envia os dados de fala para um servidor na nuvem, onde um mecanismo de análise de fala processa o sinal através de uma cadeia de processamento de fala clínica baseada em conhecimento para obter métricas relevantes. É através destas métricas que o SLP monitora a qualidade da fala do paciente e infere se houve melhora. Ainda, o sistema de inferência usa as métricas computadas para fornecer relatórios de saúde automáticos para os SLPs.

Gabani et al. (2011) exploram o uso de um método automatizado para analisar narrativas infantis de forma a identificar a presença ou ausência de deficiência de linguagem. A tarefa é prever automaticamente o status do idioma das crianças a partir de transcrições ortográficas de seus enunciados gravados por áudio. Para isto, são utilizados modelos de linguagem e técnicas de processamento de linguagem natural em um ambiente de aprendizagem de máquina para derivar recursos adicionais. Dessa forma, é possível identificar informações que ajudem o terapeuta na tarefa de determinar o status de linguagem da criança. Os modelos de linguagem são treinados a partir de corpus de fala composto por conjuntos de dados diferentes, de forma que os modelos de linguagem aprendidos possam ser posteriormente utilizados com base nos novos dados. Uma vez que um modelo é treinado e as probabilidades condicionais são calculadas, usa-se o modelo para calcular as probabilidades de sequência de palavras.

Devido a grande preocupação no auxílio a pessoas com distúrbios de fala apontada por Schipor, Pentiu e Schipor (2010), os autores propõem um sistema CBST (*Computer Based Speech Training*) a fim de facilitar a avaliação do fonoaudiólogo, ajudando-o na intervenção terapêutica. O sistema especialista baseia-se em um guia terapêutico formalizado por meio de paradigma de Lógica Fuzzy, método que permite lidar com aspectos imprecisos da realidade. A arquitetura do sistema CBST proposto é baseada nas interações entre seis blocos funcionais: criança, fonoaudiólogo, programa de monitor de laboratório, sistema especialista, modelo 3D e programa de monitor infantil. O sistema especialista, se ativado, faz sugestão sobre alguns parâmetros de treinamento como frequência, duração e conteúdo da sessão, de acordo com as variáveis de entrada. É possível armazenar a evolução precisa de cada criança, de forma que o sistema pode se adaptar de acordo com esta evolução. Os modelos são treinados e atualizados de acordo com uma base de conhecimento composta por mais de 230 regras e 22 variáveis linguísticas. Dessa forma, o sistema gera conjuntos de exercícios constituídos a partir de informação contextual específica de cada paciente.

Grzybowska e Klaczynski (2014) apresentam um programa de software que pode ser usado para apoiar a terapia fonoaudiológica, que faz uso de tecnologia *Automatic Speech Recognition* (ASR) para identificar a identidade do locutor. O objetivo da proposta é verificar se o som articulado é o mesmo que os modelos previamente gravados. É reali-

zada a gravação do sinal de fala do usuário, transformado em sinal elétrico. O áudio passa por uma etapa de pré-ênfase para melhorar a taxa média de sinal-ruído. A qualidade da pronúncia é medida por meio do algoritmo *Dynamic Time Warping* (DTW) e o algoritmo *k-Nearest Neighbors* (kNN) realiza a classificação do sinal gravado.

Grossinho et al. (2016) apontam que a terapia da fala pode ser usada para resolver distúrbios dos sons da fala, os quais consistem em um atraso na aquisição de fonemas, sílabas ou palavras. Neste contexto, os autores propõem uma solução de reconhecimento de fonemas para um ambiente interativo de terapia de fala, elaborado segundo orientações de SLPs. Este ambiente, chamado VisualSpeech, aborda a produção de fonemas e inclui elementos motivacionais para a criança. A abordagem de reconhecimento consiste em extrair recursos de produções de fala e usar esses recursos em um algoritmo de classificação.

No campo de processamento da fala, a segmentação possui um papel importante, pois há necessidade de detectar e isolar algumas partes do discurso. Devido a esta questão, Iliya e Neri (2016) apresentam uma técnica baseada em sistema neural para segmentar enunciados de fala, a fim de dar suporte à terapia fonoaudiológica. Foram desenvolvidos e comparados dois modelos de segmentação para detectar e identificar inconsistências nos sinais de fala desordenados. O sistema é treinado e a taxa de aprendizagem é calculada, de forma que a arquitetura se adapta a várias vozes e sujeitos e detecta a parte da fala que é de interesse do terapeuta. A base de conhecimento utilizada como recurso de treinamento consistiu de um conjunto de frames de sinais de fala contendo pronúncias desordenadas e corretas. Além disso, conhecimento especialista foi utilizado para guiar ações a serem tomadas ao longo do processo de pesquisa.

Ward et al. (2016) apontam a falta de SLPs no atendimento à população, questão que pode ser otimizada com tecnologia ASR para automatizar a avaliação inicial. Dessa forma, os autores desenvolvem um sistema de prova de conceito com base em conhecimento especializado para identificar e avaliar PEPs (padrões de erros fonológicos, ou *Phonological Error Patterns*) no discurso infantil. O sistema proposto visa realizar a triagem de crianças para avaliação de SLP com base no valor de seu PEP (se é de baixo risco, risco moderado ou alto risco). Os dados de entrada são compostos pelo discurso de crianças realizando tarefas de nomeação de imagens. Através da detecção de PEP o SLP pode monitorar remotamente a terapia do paciente. Dicionário de fonemas foram criados em consulta com SLPs para cada palavra-alvo, consistindo de fonemas alvo seguidos pelas substituições (PEPs) esperadas para cada posição de fonema. O algoritmo de Viterbi é então usado para inferir a sequência mais provável de fonemas.

Le et al. (2016) apontam barreiras presentes na terapia de fala tradicional, como a falta de SLPs e o alto custo de tratamento. Nesse contexto, a terapia computadorizada pode apresentar uma alternativa acessível e de baixo custo para auxiliar no tratamento do paciente. De forma a fornecer um feedback significativo aos terapeutas, os autores pro-



põem um sistema automatizado para avaliar três aspectos da inteligibilidade do discurso do paciente afásico: clareza, fluidez e prosódia. Como primeiro passo, é aplicada uma avaliação de fala em aplicativo rodando em tablet, na qual o paciente deve descrever a imagem que vê, produzindo uma frase. São utilizados dois corpus de fala, um contendo amostras de fala de pessoas saudáveis e outro com amostras de fala de indivíduos afásicos. É feita anotação do corpus de fala afásico consistindo de duas tarefas: uma para produzir rótulos em nível de palavra para treinamento ASR e a segunda para produzir pontuações para modelagem da inteligibilidade da fala. A fase seguinte é de geração de transcrição automática, que objetiva obter uma transcrição detalhada do que o paciente falou incluindo alinhamentos de palavras e sílabas. Essa transcrição é pré-requisito para extração de recursos para classificação da inteligibilidade da fala.

Kumar e Kumar (2016) apresentam uma ferramenta para terapia de fala com o objetivo de melhorar a inteligibilidade do discurso do indivíduo com disartria<sup>3</sup>. O trabalho emprega transformação de características entre a população com fala normal e fala disártrica, e utiliza conjuntos de dados destas duas populações para avaliar a eficácia da metodologia proposta. Os dados provêm da base de dados Nemours, utilizada como base de conhecimento para melhorar a inteligibilidade do discurso disártrico. A fala transformada é resintetizada e avaliada subjetivamente em termos de inteligibilidade, naturalidade e taxa de erro de palavras. Uma avaliação objetiva também é feita para avaliar a qualidade do discurso antes e depois do aprimoramento. Uma melhora significativa da inteligibilidade do discurso do paciente foi obtida através da utilização dos algoritmos propostos.

Awad et al. (2015) apresentam uma aplicação móvel Android destinada a pacientes com distúrbios de fala e linguagem e, principalmente, pacientes com doença de Parkinson<sup>4</sup>, visto que terapia de fala computadorizada é capaz de suprir uma série de limitações apresentadas pela terapia tradicional. O dispositivo captura o sinal de fala do indivíduo e o software usa informações fornecidas pelo terapeuta para fornecer feedback visual e auditivo para indicar ao paciente com hipertonia<sup>5</sup> ou hipotonia<sup>6</sup> a necessidade de controlar o volume da própria voz. O aplicativo registra continuamente o estado da fala do paciente (alto, baixo ou aceitável) a fim de fornecer relatórios ao terapeuta. A interface apresenta setas que simbolizam a necessidade de aumentar ou diminuir o volume da fala. Se o paciente está falando abaixo do volume mínimo desejado, irá tocar o som de uma multidão para instigá-lo a falar mais alto. Para os pacientes que falam muito alto, um som relaxante deve fazer com que o paciente fale menos, combatendo a hiperfonia habitual. O aplicativo desenvolvido possui uma variedade de configurações que podem ser definidas pelo terapeuta, além de outros recursos de personalização fornecidos através de classes Android.

---

<sup>3</sup>desordem neuro-motora que prejudica os músculos utilizados para produção e articulação da fala.

<sup>4</sup>desordem degenerativa que afeta o controle físico do paciente, evidenciado frequentemente sob a forma de distúrbios da fala.

<sup>5</sup>no contexto da fala, refere-se à condição muscular que causa rigidez nas cordas vocais.

<sup>6</sup>no contexto da fala, refere-se à condição muscular que causa fraqueza nas cordas vocais.

Black, Tepperman e Narayanan (2011) afirmam que avaliar a alfabetização infantil pode oferecer um feedback valioso ao professor, além de ajudar a criança a adquirir habilidades de leitura. Por isso, os autores propõem-se a modelar e prever a capacidade geral da leitura das crianças com base em gravações de áudio. Em uma primeira etapa, avaliadores humanos especialistas avaliaram a capacidade de leitura das crianças. Em uma segunda etapa, foram realizadas experiências de aprendizado de máquina que permitiram prever a pontuação dos avaliadores usando recursos extraídos automaticamente do áudio. Na fase de extração de características, para cada palavra alvo foi construído um dicionário com ajuda de professores e linguistas especializados. As pronúncias inaceitáveis foram feitas substituindo as pronúncias corretas por erros comuns de letra-para-som. Foi executada tecnologia ASR baseada em regressão linear a partir dos dicionários de pronúncia sobre cada enunciado no conjunto de desenvolvimento. O sistema final otimizado foi capaz de aprender tanto as pontuações altas de um avaliador individual quanto as pontuações médias dos avaliadores, com o mesmo nível de concordância com o qual os avaliadores concordam entre si.

Distinguir pessoas com desordens comunicativas de indivíduos com fala saudável é um desafio importante na prática clínica. Por esse motivo, Gaspers et al. (2012) exploram recursos adicionais extraídos do desempenho de tarefas para distinguir indivíduos saudáveis de indivíduos afásicos ou com *Cognitive Communication Disorder* (CCD). Como primeiro passo, dados foram coletados de indivíduos afásicos, pacientes com CCD e indivíduos saudáveis, cada um com cinco tarefas de geração de palavras. As respostas dos participantes foram transcritas e anotadas. As listas de palavras resultantes foram utilizadas para extrair vetores de características contendo 44 recursos diferentes. Destes, os recursos mais relevantes foram escolhidos para treinar e avaliar um conjunto de classificadores. A saída dos classificadores foi a decisão entre (1) sujeito afásico e saudável, (2) paciente CCD e saudável, ou (3) paciente CCD e afásico. Foram utilizados recursos adicionais (características temporais, lexicais e sublexicais) para complementar a tarefa de classificação e, posteriormente, a precisão de cada classificador foi comparada.

Kim, Kim e Kim (2015) apresentam um método para avaliar automaticamente a inteligibilidade da fala de pacientes com disartria. O método proposto consiste de duas etapas principais: representação de recursos e previsão. Na primeira, os discursos de indivíduos saudáveis e disártricos são convertidos em sequências fonéticas. Na segunda etapa, é utilizada técnica ASR e de alinhamento para mapear características como coincidência, substituição e exclusão em uma sequência fonética. Através de um modelo PSSLM (*Phonologically-Structured Sparse Linear Model*) são representados padrões de pronúncia específicos para a inteligibilidade do discurso disártrico e previsões são feitas usando os padrões de pronúncia selecionados.

Toki, Pange e Mikropoulos (2012) reportam o desenvolvimento de um sistema grego especialista, denominado APLO, destinado a clínicos e professores envolvidos na aprendi-

zagem da fala. O sistema incorpora conhecimento especialista de avaliação de linguagem integrando redes neurais e regras simbólicas. O avaliador coleta uma variedade de informações sobre as habilidades verbais da criança, através de uma lista de perguntas e respostas. Posteriormente, o sistema fornece resultados automatizados com relação à avaliação das habilidades de fala. O projeto do sistema APLo se baseia em um sistema *Multi-Input Multi-Output (MIMO)* construído a partir de subsistemas menores, incluindo as condições para distúrbios fonológicos, de articulação, de fluência, de linguagem, entre outros. O subsistema de marcos de idade trata do conjunto de regras responsáveis pela aplicação de condições que determinam se uma criança está tendo uma performance de acordo com o esperado e de acordo com sua idade, ou seja, o sistema realiza raciocínio de acordo com as situações que se desenvolvem ao longo do tempo. Regras *if-then* são aplicadas a fim de produzir resultados de acordo com as informações do paciente. Quando os subsistemas completam as regras da aplicação, os resultados são incluídos no relatório de diagnóstico, juntamente com possíveis direções de intervenção. O feedback dado pelo sistema pode ser usado por SLPs e outros profissionais para planejar tarefas de terapia ou aprendizagem no contexto de fala e linguagem.

Caballero-Morales e Trujillo-Romero (2013) apresentaram uma abordagem que integra vários padrões de pronúncia e múltiplas respostas provenientes de um sistema ASR para aprimoramento da fala disártrica. Essa integração é realizada ao ponderar respostas de um sistema ASR, definindo restrições no modelo de linguagem. Foi realizada rotulagem ortográfica e fonética de recursos acústicos estendidos da base de dados Nemours, usada para avaliação do abordagem, a qual contém dados de indivíduos com fala normal e dados de fala disártrica. Uma versão estendida de metamodelos foi usada na proposta para modelagem de padrões de inserção, onde cada fonema é associado a um contexto de inserção. Os padrões das múltiplas respostas do sistema ASR são ponderados e integrados por meio de um algoritmo genético que oferece vantagens para lidar com dados incompletos e inconsistentes. Foi feita adaptação do sistema de linha de base para cada falante disártrico, e os modelos de linguagem para a linha de base juntamente com os metamodelos foram estimados a partir das transcrições ortográficas e fonéticas de todas as frases no banco de dados. As experiências mostraram que a abordagem evolutiva melhorou significativamente a precisão ASR quando comparada com metamodelos não otimizados.

Hariharan et al. (2013) propuseram um sistema especialista híbrido para detectar patologia da prega vocal usando amostras de voz compactadas (de baixa qualidade). O sistema inclui compressão de voz, extração de recursos, métodos de ponderação de características baseadas em agrupamento e classificação. Os métodos de ponderação de características baseados em agrupamento realizaram tarefas de normalização de dados e remoção de ruído, a fim de melhorar o desempenho dos classificadores. Amostras de voz das bases de dados MEEI e MAPACI foram utilizadas a fim de avaliar a robustez e

independência dos algoritmos para as bases de dados. Modelos acústicos foram treinados para utilização em quatro classificadores. Os métodos de ponderação/pré-processamento baseado em agrupamento melhoraram a capacidade discriminativa dos recursos brutos e o desempenho dos classificadores.

Bolaños et al. (2011) afirmam que avaliações de leitura são importantes para identificar estudantes que precisam de ajuda imediata, além de ajudar professores e demais profissionais a monitorar o progresso dos alunos. Por isso, os autores apresentam FLORA, um programa de computador acessível que usa reconhecimento de fala para fornecer medidas sobre a capacidade de leitura das crianças. O sistema apresenta passagens de texto de nível escolar para as crianças, e a partir das leituras realizadas, calcula o número de palavras corretas por minuto (WCPM, ou *Words Correct Per Minute*), uma medida padrão da fluidez da leitura oral. Os resultados (pontuações) são apresentadas analisando o contexto escolar do aluno. O sistema contém uma interface do estudante e uma interface do professor e a arquitetura é cliente-servidor. No lado servidor, há uma aplicação Java responsável por receber o áudio do cliente, armazená-lo e enviá-lo ao módulo de processamento. A proposta conta com um reconhecedor de fala automático baseado em C++ e modelagem acústica. Há, também, um rastreador de leitura chamado “ReadToMe”, que permite monitorar o WCPM do aluno junto com as porcentagens correspondentes a cada nível.

### 3.3.1 Integração de Consciência de Situação na Fonoaudiologia

Salfinger, Retschitzegger e Schwinger (2013) apresentaram uma série de critérios baseados em componentes de sistemas conscientes de situação. Estes critérios referem-se à capacidade de sistemas estabelecer ou obter consciência de situação, bem como manter SA ao longo do tempo. De acordo com os autores, a fim de *obter* Consciência de situação, o sistema deve (1) considerar a composição dos dados de entrada, (2) explicitar o conhecimento da informação que será gerenciada através de um modelo de domínio, (2) avaliar situações de interesse a partir da utilização de métodos específicos e (4) fornecer suporte ao operador humano em ações que devem ser tomadas.

De maneira semelhante, Salfinger, Retschitzegger e Schwinger (2013) apontam que, para *manter* Consciência de Situação, a proposta deve ser capaz de (1) capturar e rastrear a evolução de uma situação, (2) oferecer capacidade de projeção, (3) incorporar informação contextual no processo de interpretação de situações, (4) lidar com dados incompletos ou inconsistentes, (5) fornecer adaptação dos métodos utilizados para avaliar Consciência Situacional, (6) permitir ajuste do sistema observando fatores como desempenho e tempo de resposta, (7) utilizar uma base de conhecimento, (8) incorporar inteligência humana aliada ao processamento de máquina, (9) suportar personalização do sistema e

(10) fornecer explicação e exploração ao operador humano acerca do processamento conduzido.

Os critérios apontados pelos autores foram utilizados para analisar abordagens em diferentes domínios de aplicações, incluindo tráfego rodoviário, vigilância marítima, assistência ao motorista e monitoramento do espaço aéreo. Neste tópico, os critérios que avaliam a integração de Consciência de Situação são utilizados para analisar e comparar os trabalhos relacionados apresentados anteriormente. A tabela 3.3 sumariza as análises realizadas e uma estatística geral sobre a integração de SA nas propostas encontra-se no gráfico 3.1.

Ao analisar os trabalhos relacionados e comparar os critérios por eles sustentados, evidenciam-se algumas características como, por exemplo, o uso de bases de conhecimento pela maioria das propostas (75%). No domínio da fala, bases de conhecimento são comumente caracterizadas por dados de histórico e corpus de fala, utilizados para treinamento e avaliação das metodologias propostas. Através de um base de conhecimento, o sistema pode se tornar melhor em suas habilidades de avaliação na medida que novos dados são incorporados. Em outras palavras, o desempenho de um sistema pode ser melhorado quando conhecimento é usado para treinar métodos, guiando sua execução de acordo com o contexto de cada ambiente.

O uso de bases de conhecimento e o aprendizado a partir de experiências passadas estão intimamente ligados com a habilidade de adaptação. Este fato se torna evidente ao observar que, assim como o critério "Base de Conhecimento", o critério "Adaptação de SA" também possui grande abrangência nas propostas (75%). Isto demonstra que, conforme há mudanças de contexto no ambiente (atualização de base de conhecimento ou mudanças causadas por ações do operador humano, por exemplo), as propostas são capazes de adaptar seus métodos a essas mudanças. Especificamente, a maioria dos estudos demonstrou esta capacidade de adaptação ao incorporar métodos de aprendizado de máquina para fornecer reconhecimento e adaptação a diferentes amostras de fala.

Outro critério importante para avaliação de Consciência de Situação é a incorporação de informação contextual. No geral, 55% dos trabalhos relacionados incluiu informações de contexto em suas metodologias, como forma de melhorar a interpretação de situação. No domínio da fonoaudiologia, o contexto do paciente é incorporado para otimizar processos de classificação e reconhecimento de fala, e é frequentemente caracterizado por fatores como idade, escolaridade e gênero. A incorporação de informação contextual, além de influenciar no processo de interpretação de situações, pode ser usada para fornecer feedback personalizado a pacientes, terapeutas e outros indivíduos envolvidos.

Enquanto alguns critérios de Consciência de Situação encontram-se consolidados em sistemas no domínio da fala (como Base de conhecimento e Adaptação de SA), outros movem-se rumo à consolidação, como "Incompletude e inconsistência". Considerando todos os trabalhos relacionados, 40% deles abrangem ao menos parcialmente esta carac-

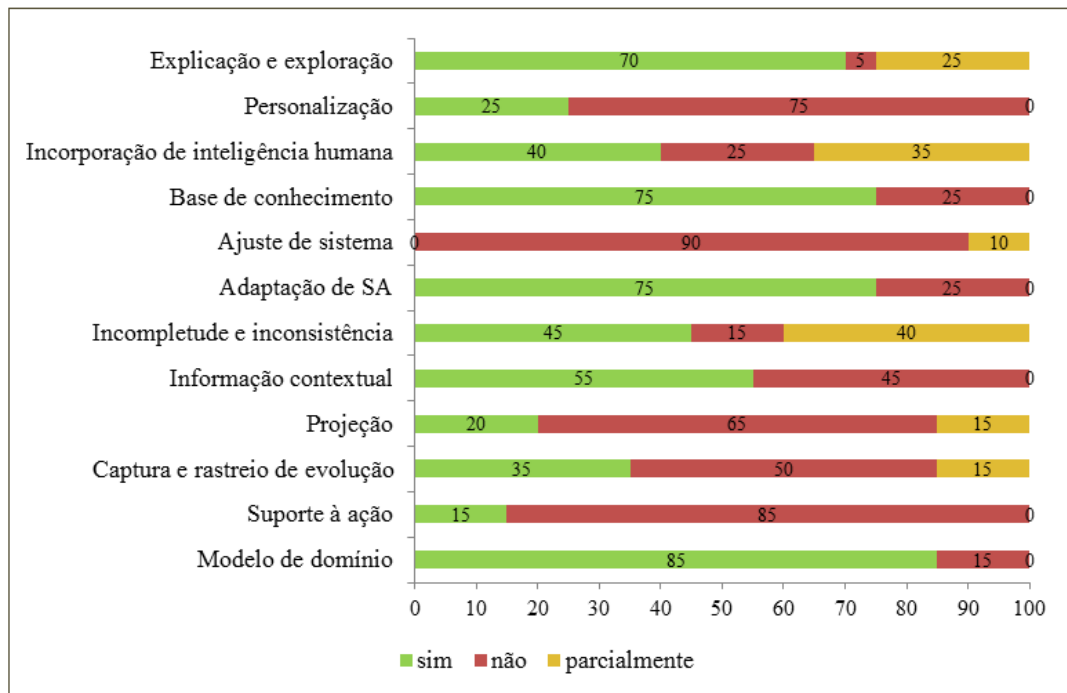
Tabela 3.3 – Integração de Consciência de Situação em sistemas de suporte à terapia da fala.

Proposta	Obtenção de SA					Conservação de SA					Evolução de Uso			
	Dados de Entrada	Modelo de Domínio	Avaliação de Situação	Suporte à Ação	Evolução do Ambiente			Evolução do Sistema			Incorporação de Inteligência Humana	Personalização	Explicação e Exploração	
					Captura e Rastreamento de Evolução	Projeção	Informação Contextual	Incompletude e Inconsistência	Adaptação de SA	Ajuste de Sistema				Base de Conhecimento
Robles-Bybaev et al., 2016	Heterog.	sim	Ontologias + regras	sim	parcial	sim	sim	não	não	sim	sim	não	sim	sim
Grossinho et al., 2016	Heterog.	não	NB, SVM e KDE	não	não	não	sim	sim	sim	não	não	parcial	sim	sim
Iliya e Neri, 2016	Homog.	sim	ANNs e SVM	não	não	não	sim	sim	sim	parcial	sim	parcial	não	sim
Ward et al., 2016	Homog.	sim	Decodificador HMM e HNN	não	parcial	não	sim	sim	sim	não	sim	sim	não	parcial
Le et al., 2016	Heterog.	sim	Árvore de decisão C4.5, LogR, NB, SVM e RF	não	não	sim	sim	sim	sim	sim	sim	parcial	sim	sim
Kumar e Kumar, 2016	Heterog.	sim	LPCs e DTW	não	não	não	parcial	parcial	parcial	não	sim	não	não	parcial
Parmadi et al., 2015	Homog.	sim	HMM, SVM, MLP e MaxEnt	não	sim	não	sim	sim	sim	não	sim	sim	sim	sim
Dubey et al., 2015	Homog.	sim	CLIP	não	sim	não	parcial	parcial	parcial	não	não	parcial	não	sim
Awad et al., 2015	Homog.	sim	Algoritmo desenvolvido pelos autores	não	sim	sim	parcial	parcial	parcial	não	não	parcial	sim	não
Kim, Kim e Kim, 2015	Heterog.	sim	Decodificador HMM e PSSLM	não	não	parcial	sim	não	não	não	sim	sim	não	sim
Grzybowska e Klaczynski, 2014	Homog.	sim	DTW e kNN	não	não	não	parcial	parcial	parcial	não	não	parcial	não	parcial
Caballero-Morales e Trujillo-Romero, 2014	Heterog.	sim	HMMs	não	não	sim	sim	sim	sim	não	sim	não	não	sim
Abad et al., 2013	Heterog.	sim	HMM + MLP	não	sim	não	parcial	parcial	parcial	sim	parcial	sim	não	sim
Haritharan et al., 2013	Heterog.	sim	LS-SVM + kNN + PNN + CART	não	não	não	sim	sim	sim	não	sim	não	não	sim
Gaspers et al., 2012	Heterog.	sim	TNN, LMT, MLP, NB e SVM	não	não	não	parcial	parcial	parcial	não	não	não	não	sim
Toki, Pange e Mikropoulos, 2012	Heterog.	sim	Regras "se-então" e redes neurais	sim	sim	sim	sim	não	não	não	não	sim	não	parcial
Gabani et al., 2011	Heterog.	não	LM, NLP e ML	não	parcial	parcial	sim	sim	sim	sim	sim	não	não	parcial
Black, Tepperman e Narayanan, 2011	Homog.	não	LR, HMM	não	não	parcial	parcial	parcial	parcial	não	sim	sim	não	sim
Boiaros et al., 2011	Heterog.	sim	ASR e HMMs/GMMs	não	sim	sim	parcial	parcial	parcial	não	sim	parcial	não	sim
Schipor, Pentiu e Schipor, 2010	Heterog.	sim	Lógica Fuzzy	sim	sim	sim	sim	sim	sim	não	sim	sim	sim	sim

Fonte: Da autora.

terística, ao passo que 45% abrangem-na totalmente, lidando com ruídos e qualidade de som nas gravações de áudio utilizadas. Os desafios de pesquisa mais evidentes ao analisar integração de SA nas propostas incluem os critérios de Projeção, Suporte à ação, Ajuste de sistema e Personalização, já que apenas uma pequena parte dos trabalhos foi capaz de garantir esses atributos.

Gráfico 3.1 – Porcentagem de critérios de Consciência de Situação suportados pelas propostas.



Fonte: Da autora.

Personalização refere-se à habilidade de customizar um sistema para que ele possa fornecer serviços a diferentes usuários. Na área da fonoaudiologia, apenas 25% dos trabalhos cobriram esse critério. Considerando que cada operador tem suas necessidades e preferências, é ideal que recursos de personalização sejam melhor explorados, especialmente em contextos de saúde, onde as ações terapêuticas são tomadas em prol do paciente. Da mesma forma, pode-se observar que "Ajuste de sistema" é um critério não abordado por 90% dos estudos, demonstrando que estratégias de otimização não costumam ser incorporadas nos processos metodológicos. Visto que sistemas conscientes de situação devem lidar com informações contextuais que sofrem alterações ao longo do tempo, é imprescindível que os métodos utilizados possam incorporar meios de garantir tomadas de decisões, mantendo o desempenho esperado.

Suporte à ação representa outro desafio na fonoaudiologia, já que apenas 15% dos trabalhos cobriram esse critério. Isto significa que, na maioria das vezes, as aplicações não são capazes de estabelecer uma relação entre a situação interpretada e ações apropriadas a ela, fazendo com que operador humano tome decisões sem qualquer tipo

de suporte. Na área da saúde, a qual cobre o domínio da fonoaudiologia, eventos deste tipo se tornam fatores prejudiciais em tarefas envolvendo tratamento e diagnóstico, onde agilidade e precisão são esperadas na tomada de decisões clínica.

Captura e rastreamento de situações em evolução é outra necessidade a ser explorada. Embora não tão urgente quanto as demais citadas acima, metade dos trabalhos analisados não foi capaz de integrar este critério. A habilidade de identificar a evolução de uma situação é essencial para inferir estados futuros, raciocinar sobre como eles podem afetar o ambiente, e conseqüentemente projetar ações da maneira adequada. Desta forma, capturar e rastrear modelos de evolução são fatores que contribuem para a capacidade de projeção de um sistema.

Finalmente, considerando os critérios mais relevantes para avaliação de Consciência de Situação, observa-se que o nível de Projeção (último nível abordado no Modelo de Consciência de Situação de Endsley), não foi satisfatoriamente incorporado pelas propostas na fonoaudiologia. Apenas 4 dos 20 trabalhos cobriram totalmente o critério, prevendo situações futuras e desencadeando ações automaticamente. Três dos trabalhos lidaram com o nível de Projeção de forma parcial, ao passo que o restante não cobriu o critério de nenhuma maneira (60%). De forma geral, a baixa adesão de Projeção na literatura relacionada demonstra que a habilidade dos sistemas agirem sem intervenção humana ainda representa um desafio a ser explorado. Embora as habilidades de estimar o futuro e tomar ações convenientes em uma determinada situação sejam desejáveis em qualquer domínio de aplicação, elas se tornam mais urgentes em áreas que envolvem cuidados com a saúde e prevenção, onde se espera que sistemas automatizados possam lidar com as barreiras existentes nas tarefas clínicas tradicionais.

### **3.3.2 Métodos Utilizados em Sistemas no Domínio da Fala**

Em uma análise dos trabalhos relacionados à presente proposta de sistema, foram identificados métodos aplicados no domínio da fala e quais se destacam por sua frequente utilização. A tabela 3.4 demonstra os métodos encontrados nos estudos abordados no início desta seção, ao passo que o gráfico 3.2 mostra quais destes métodos são frequentemente empregados em tarefas que envolvem processamento e reconhecimento de fala.

Modelos de Markov (HMMs) representam os métodos mais utilizados em propostas que envolvem reconhecimento de fala, frequentemente empregados em tarefas de modelagem acústica e decodificação. Um dos motivos que explicam essa questão é que os Modelos de Markov possuem características dinâmicas de modelagem temporal, as quais fornecem robustez contra flutuações de comprimento de palavras em larga escala (HASEGAWA-JOHNSON et al., 2006).

Como indicado no gráfico 3.2, coeficientes MFCC também são utilizados com frequên-

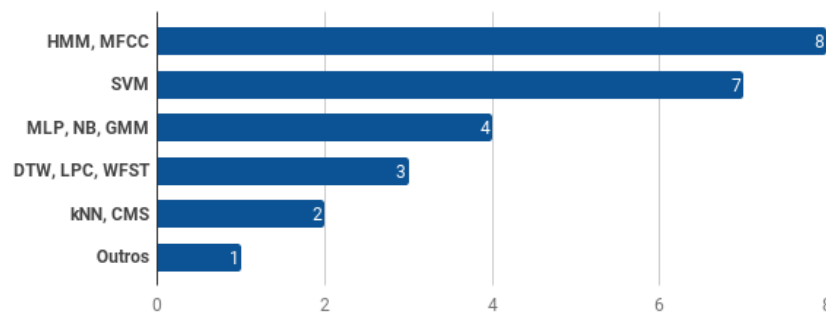


Tabela 3.4 – Métodos encontrados em trabalhos relacionados.

Método	Aplicação em Trabalhos Relacionados
Modelo Oculto de Markov (HMM)	Modelagem temporal do sinal de fala, decodificação, modelagem acústica
Perceptron Multicamadas (MLP)	Rede neural usada para classificação de padrões e processamento de fala
Máquina de Suporte Vetorial (SVM)	Algoritmo de classificação e segmentação de fala
Cadeia de Processamento de Fala Clínica (CLIP)	Cadeia de software modular que utiliza aprendizagem de máquina para obter métricas de qualidade de fala
Modelos de Linguagem (LM)	Método probabilístico e de reconhecimento de fala
Processamento de Linguagem Natural (NLP)	Derivação de características adicionais da fala
Lógica Fuzzy	Criação de modelos precisos da realidade para auxiliar na decisão clínica
Algoritmo DTW	Medição de similaridade de padrões acústicos
Algoritmo K-Vizinhos mais Próximos (k-NN)	Método de classificação
Naive Bayes	Método probabilístico e de classificação
Estimativa de Densidade Kernel (KDE)	Método probabilístico e de classificação
Rede Neural Artificial (ANN)	Segmentação de fala desordenada
Rede Neural Hierárquica (HNN)	Treinamento de modelos acústicos
Modelo de Mistura Gaussiana (GMM)	Emissão de probabilidades de HMMs e modelagem acústica
Rede Neural Profunda (DNN)	Modelagem acústica
Árvore de Decisão C4.5	Método de classificação usado para avaliar conjuntos de treinamento
Regressão Logística (LogR)	Método de classificação usado para avaliar conjuntos de treinamento
Floresta Aleatória (RF)	Método de classificação usado para avaliar conjuntos de treinamento
Codificação Preditiva Linear (LPC)	Técnica de transformação de fala e extração de características
Regressão Linear (LR)	Previsão automática de pontuações dadas por avaliadores
Coefficiente Cepstral de Frequência Mel (MFCC)	Extração de características acústicas e parametrização de dados
Algoritmo 1NN	Método de classificação
Árvore de Modelo Logístico (LMT)	Método de classificação
Transdutor Ponderado de Estado Finito (WFST)	Método de alinhamento para capturar inconsistências na fala, método de decodificação e modelagem de matriz de confusão
Máquina de Suporte Vetorial "LeastSquare"(LS-SVM)	Método de classificação para investigar a eficácia de características brutas e ponderadas
Rede Neural Probabilística (PNN)	Método de classificação para investigar a eficácia de características brutas e ponderadas
Árvore de Regressão e Classificação (CART)	Método de classificação para investigar a eficácia de características brutas e ponderadas
Subtração da Média Cepstral (CMS)	Processamento de dados ruidosos
Predição Linear Perceptual (PLP)	Extração de características
RASTA	Extração de características
Modulação de Espectrograma (MSG)	Extração de características
Entropia Máxima (MaxEnt)	Classificação de fala
Redes Bayesianas	Aprendizagem de máquina
Algoritmo LogitBoost	Aprendizagem de máquina
Coefficiente Cepstral de Fator Humano (HFCC)	Parametrização do sinal de fala
Coefficiente Delta e Delta-Delta	Parametrização do sinal de fala
Coefficiente Normalizado de Auto-Correlação (NC)	Extração de características
Agrupamento k-Means (KMC)	Método de ponderação de características baseado em agrupamento
Fuzzy C-Means	Método de ponderação de características baseado em agrupamento
Agrupamento Subtrativo	Método de ponderação de características baseado em agrupamento
Rede de Confusão de Pronúncia (PCN)	Captura de possíveis variações na pronúncia
Modelo Linear Esperso Fonologicamente Estruturado (PSSLM)	Avaliação automática de inteligibilidade de fala desordenada
Gramática de Estado Finito (FSG)	Verificação automática da pronúncia
Algoritmo genético	Otimização de técnica baseada em HMM
Regressão Linear de Máxima Verossimilhança (MLLR)	Adaptação não supervisionada de métodos e variações para distribuições gaussianas
Árvore de Regressão (RT)	Agrupamento de distribuições gaussianas
Maximização da Expectativa	Agrupamento de distribuições gaussianas

cia, no que se refere a tarefas de pré-processamento e extração de características de segmentos de fala. Máquinas de Suporte Vetorial (SVMs) aparecem em segundo lugar em número de ocorrências. Este método, comumente usado para classificação e reconhecimento de padrões, apresenta boas precisões em testes de desempenho e tem comprovado sua eficácia ao detectar a porção de fala que é de interesse do terapeuta (ILIYA; NERI, 2016). Ainda, SVMs representam uma boa técnica ao lidar com redução e deleção de consoantes (HASEGAWA-JOHNSON et al., 2006).

Gráfico 3.2 – Total de ocorrências dos métodos nos trabalhos relacionados.



Fonte: Da autora.

MLP (*MultiLayer Perceptron*), Naive Bayes e GMM (*Gaussian Mixture Model*) são métodos que ocorrem 4 vezes cada dentro os trabalhos relacionados. O primeiro trata-se de uma rede neural comumente utilizada para tarefas de classificação. De forma similar, Naive Bayes realiza classificação por meio de teoremas de probabilidade e apresenta bom desempenho em situações de alta dimensionalidade de dados (GROSSINHO et al., 2016). Por fim, GMMs são utilizados em algumas propostas para realizar modelagem acústica e emissão de probabilidades para HMMs.

Métodos com três ocorrências cada envolvem DTW (*Dynamic Time Warping*), LPC (*Linear Predictive Coding*) e WFST (*Weighted Finite-State Transducer*). O algoritmo DTW é uma técnica usada para facilitar a comparação de características entre enunciados de fala (KUMAR; KUMAR, 2016). LPCs são utilizados para tarefas de extração de características, já que representam uma forma simples e rápida de estimar os principais parâmetros de sinais de fala (GROSSINHO et al., 2016). O método WFST, por sua vez, é empregado para modelagem de pronúncia e reconhecimento de fala de vocabulário extenso, tendo provado sua efetividade ao reduzir erros de inserção em dados de fala (KIM; KIM; KIM, 2015).

KNN (*k-Nearest Neighbors*) e CMS (*Cepstral Mean Subtraction*) são métodos que obtiveram destaque duas vezes cada, dentro os estudos analisados. O primeiro, kNN, é um algoritmo baseado em instância e um classificador não-paramétrico que possui simplicidade de compreensão e implementação (HARIHARAN et al., 2013). O segundo, CMS, representa uma técnica de normalização aplicada em dados ruidosos. Todos os demais

métodos encontrados obtiveram apenas uma ocorrência cada.

Considerando os métodos utilizados no domínio da fala, percebe-se que a metodologia CBR permanece um desafio a ser explorado, já que não houve menção deste método em nenhum dos trabalhos analisados. Ainda, no que se refere a seu uso conjunto com a Teoria do Desenvolvimento Cognitivo, destaca-se o diferencial da proposta da presente dissertação. Complementarmente, por meio dos resultados observados na análise comparativa de trabalhos relacionados, poucas características de consciência de situação são integradas nos estudos, de forma que é possível encontrar subsídios para a implementação de um sistema consciente de situação no domínio da fonoaudiologia. Assim, no próximo tópico são apresentadas motivações para a construção de uma arquitetura de sistema de recomendação baseado em Consciência de Situação, que objetiva fornecer suporte aos profissionais da fonoaudiologia em atividades envolvendo triagem e diagnóstico de distúrbios dos sons da fala infantil.

### 3.4 MOTIVAÇÃO E JUSTIFICATIVA

Como justificativa para o desenvolvimento da arquitetura de sistema proposta, tem-se que a terapia de fala computadorizada pode ajudar a superar barreiras existentes na terapia tradicional, como a falta de profissionais disponíveis para atendimento e a dificuldade de monitorar adequadamente o paciente. O sistema de recomendação se justifica ao considerar as seguintes características:

- **Fraca integração de Consciência de Situação:** A Consciência de Situação é um fator imprescindível para a tomada de decisões em uma grande variedade de ambientes. No domínio da fala, entretanto, observou-se que critérios que caracterizam SA são pouco abrangidos, como mostrado no item 3.1.1. Torna-se necessário, portanto, a exploração de um sistema voltado à área fonoaudiológica que possa atender requisitos de percepção, compreensão e projeção de ações, especificados no modelo de Endsley (ENDSLEY, 1995);
- **Metodologia CBR raramente empregada:** Como exposto na seção anterior, uma grande variedade de métodos é utilizada na área da fonoaudiologia no desenvolvimento de sistemas automatizados. Entretanto, observa-se a falta de propostas que utilize a metodologia CBR para resolução de problemas. O Raciocínio Baseado em Casos, além de possuir excelentes capacidades de aprendizagem, possui bom desempenho ao lidar com contextos de saúde, motivo pelo qual deve ser melhor explorado no suporte à fonoaudiologia;
- **Inclusão de aspectos cognitivos:** A Teoria do Desenvolvimento Cognitivo integra a arquitetura de sistema de recomendação através de estratégias cognitivas (assimila-

ção e acomodação), as quais devem favorecer o raciocínio empregado e permitir que o sistema possa lidar com situações inesperadas, adaptando seu processamento. A TDC integrada representa um diferencial explícito em comparação com trabalhos anteriores analisados, visto que não abordam aspectos desta teoria.

A fim de sustentar a proposta, uma abordagem conceitual foi desenvolvida e definiu-se uma arquitetura de sistema, apresentada detalhadamente na seção seguinte.

## 4 SISTEMA DE RECOMENDAÇÃO BASEADO EM CONSCIÊNCIA DE SITUAÇÃO

Nessa seção, será contextualizado o Sistema de Recomendação Baseado em Consciência de Situação, o qual utiliza-se de conceitos da Teoria do Desenvolvimento Cognitivo e da metodologia CBR. Será apresentada a arquitetura proposta e seu funcionamento, incluindo os módulos que a compõem.

### 4.1 CONTEXTUALIZAÇÃO

De forma geral, sistemas de saúde que envolvem consciência de situação consistem de dois componentes: um responsável pela detecção da situação do paciente (por meio do uso de sensores) e outro responsável por selecionar a intervenção apropriada. Neste último, a situação reconhecida é usada para inferir as causas do comportamento do paciente e, com base nesta inferência, eleger uma intervenção (YORDANOVA et al., 2016).

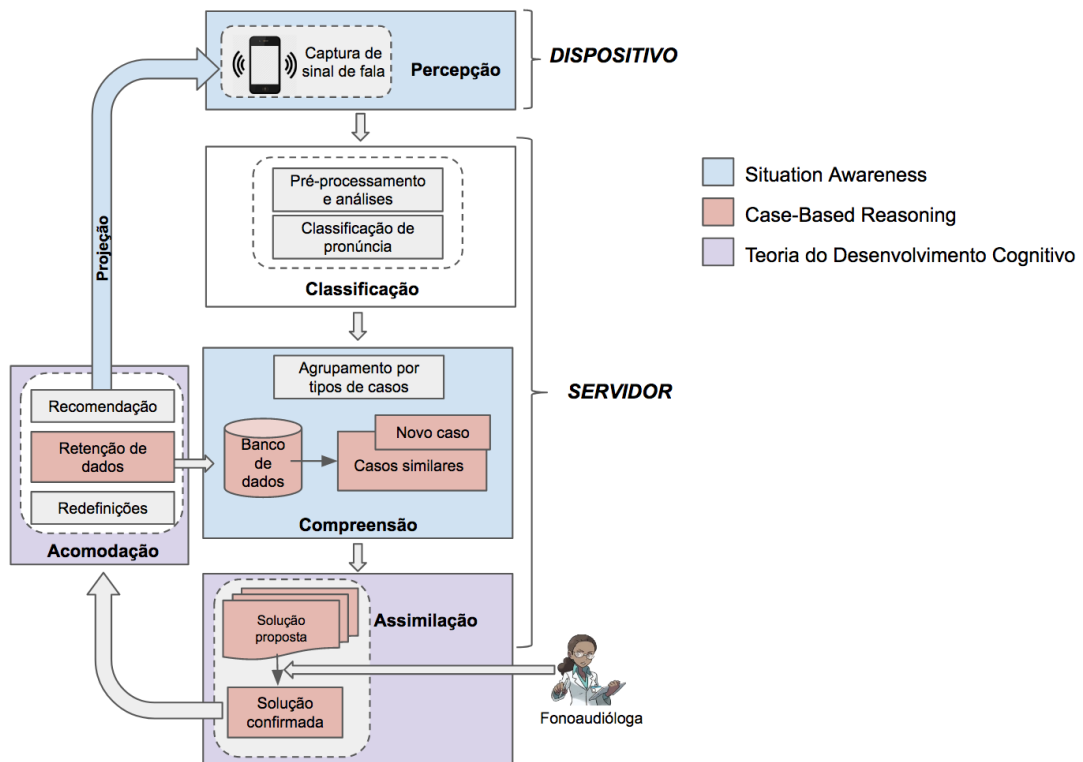
No contexto da terapia da fala, Martín-Ruiz, Duboy e De La Cruz (2013) apontam que, geralmente, os pediatras não dispõem de tempo suficiente para realizar o monitoramento adequado do desenvolvimento neuro-evolutivo infantil. Ao realizar tal monitoramento, pode-se evitar doenças que comprometem fatores biológicos, psicológicos e sociais da criança. Neste sentido, o sistema proposto visa utilizar conceitos de consciência de situação para fornecer serviço relevante na área da fonoaudiologia, melhorando a triagem de distúrbios da linguagem no cuidado infantil, e oferecendo apoio aos terapeutas no diagnóstico de transtornos de linguagem e na tomada de decisões.

### 4.2 ARQUITETURA DE SISTEMA

A arquitetura desenvolvida (Figura 4.1) se baseou em três teorias específicas: consciência de situação, teoria do desenvolvimento cognitivo e raciocínio baseado em casos. A escolha de teorias para compor a arquitetura se justifica pelo fato de que uma faz-se complementar à outra. Os conceitos de assimilação e acomodação presentes na teoria de Piaget agregam aspectos cognitivos na estrutura, o raciocínio baseado em casos fornece reuso de soluções e ampliação da base de conhecimento, essenciais para a aquisição de capacidades em um sistema clínico. Por sua vez, a consciência de situação permite que o sistema proposto possa responder adequadamente ao ambiente no qual é inserido, através da integração de três aspectos fundamentais: Percepção, Compreensão e Projeção.

Na arquitetura proposta, o nível de *Percepção* é atingido através da coleta de sinais

Figura 4.1 – Arquitetura do sistema de recomendação.



Fonte: Da autora.

de fala durante tarefas de nomeação. Desta forma, o sistema se torna consciente dos elementos dispostos no ambiente e seus estados atuais, avaliando sua relevância para a tomada de decisões. A capacidade de *Compreensão* deve ser adquirida através de tarefas de avaliações realizadas com ferramentas de reconhecimento e classificação de fala. Também, ao longo do ciclo CBR, o nível de compreensão do sistema deve ser expandido, visto que o mesmo baseia-se em casos previamente diagnosticados para realizar sugestões. Finalmente, o nível de *Projeção* deve ser alcançado a partir da identificação e compreensão da situação do paciente, onde o fonoaudiólogo recebe como retorno a sugestão do sistema.

A arquitetura do sistema é composta por duas unidades principais - unidade móvel e unidade servidor - que são explanadas a seguir.

#### 4.2.1 Unidade Móvel

A *Unidade Móvel* abrange quaisquer dispositivos móveis, tais como tablets ou smartphones, e é responsável por utilizar tais dispositivos para coletar dados de fala do público alvo e processá-los através de ferramentas de reconhecimento e classificação. Ao coletar dados do ambiente observado, o sistema é capaz de atingir o primeiro nível de SA - *Percepção* -

e utilizar as informações adquiridas para classificar o paciente como indivíduo saudável ou portador de desordem da fala.

#### 4.2.1.1 Módulo de Percepção

Compreender uma situação envolve a identificação de um conjunto apropriado de elementos de percepção, modelos de compreensão de nível superior e operadores de previsão (CAMERON et al., 2012). Assim, um sistema consciente de situação deve estabelecer, em primeiro lugar, uma etapa de percepção situacional, a qual abrange a captura de pistas sensoriais do meio onde o sistema está inserido.

Capturar dados sensoriais é a primeira etapa para o estabelecimento de compreensão situacional. Trata-se de dados brutos (por exemplo, dados numéricos mostrados por determinado sensor) que, aliados a uma informação de contexto, possibilitam compreender a situação ou evento que está ocorrendo no ambiente.

Conforme demonstrado na figura 4.1, a percepção situacional da arquitetura proposta ocorre por meio de captura de áudio (sinais de fala), que permite estabelecer as pistas sensoriais iniciais requeridas para consciência de situação. Nesta etapa, são coletados dados de pronúncia de um grupo de indivíduos, a fim de determinar, em etapas posteriores, a presença ou ausência de desordens da fala. Os dados capturados passam por uma série de etapas, que incluem pré-processamento do sinal de fala, conversão dos áudios para imagens (representações visuais) a fim de detectar padrões de fala, extração de características e, por fim, classificação da pronúncia como correta ou incorreta. Desta forma, após a extração de dados realizada na Unidade Móvel, os dados são enviados à Unidade Servidor.

#### 4.2.2 Unidade Servidor

A *Unidade Servidor* é responsável por tomar cada avaliação de paciente realizada na Unidade Móvel como entrada para o Raciocínio Baseado em Casos. Este processo ocorre, primeiramente, através de uma classificação da fala coletada, onde serão definidas palavras pronunciadas de forma correta e incorreta. Os resultados obtidos na classificação são utilizados para identificar e agrupar padrões de erros, a fim de formar tipos de casos fonológicos. Assim, dá-se início ao ciclo CBR, onde os tipos de casos identificados são *recuperados* e tomados como base para avaliação de um novo caso de entrada.

De acordo com sua adequação, os casos recuperados podem ser *reutilizados* para categorizar o novo caso com base no padrão de erro apresentado. A probabilidade da categorização é apresentada ao profissional fonoaudiólogo, que *revisa* as sugestões e dá

a confirmação da situação identificada. Em última etapa do ciclo CBR, as informações processadas e classificadas são *armazenadas* no repositório do servidor, de forma que a base de conhecimento possa se manter constantemente atualizada com novos casos. Ainda, estas informações são avaliadas para estabelecer recomendações apropriadas, as quais serão retornadas ao usuário.

É importante acrescentar que, juntamente com o ciclo CBR realizado na unidade Servidor, são integradas na arquitetura práticas que aplicam conceitos da Teoria do Desenvolvimento Cognitivo e de Consciência de Situação, objetivando otimizar o raciocínio empregado e a adaptação do sistema. As teorias utilizadas na arquitetura serão abordadas detalhadamente a seguir, na especificação do módulos que compõem a unidade Servidor.

#### 4.2.2.1 *Módulo de Classificação*

No Módulo de Classificação, os dados de fala coletados no Módulo de Percepção são pré-processados e a pronúncia do paciente é classificada como certa ou errada. Primeiramente, em uma avaliação inicial, os áudios coletados são analisados sem aplicação de qualquer técnica de pré-processamento, ou seja, sem filtragem, corte ou remoção de ruído. Esta avaliação inicial é feita por meio de uma ferramenta de código aberto (neste caso, a ferramenta Google Cloud Speech) e objetiva verificar a qualidade do som adquirido e o desempenho da ferramenta empregada.

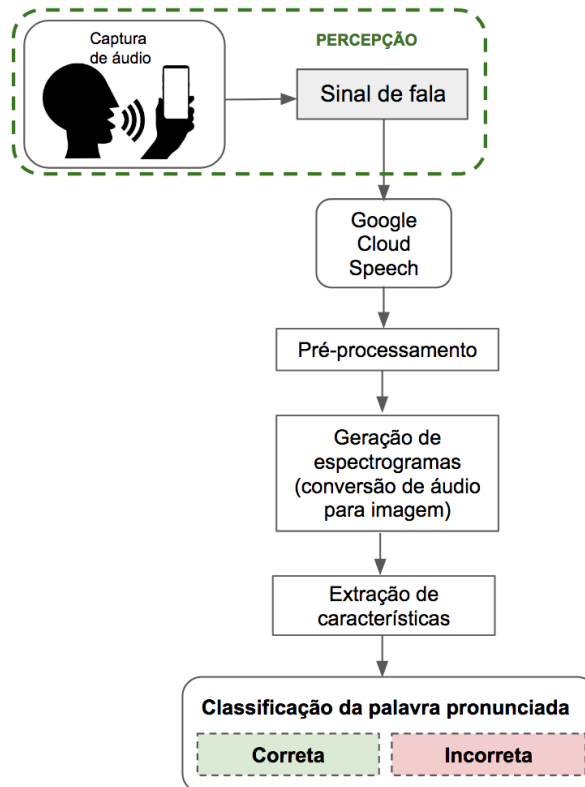
Na sequência, o Módulo de Classificação, através de ferramentas de pré-processamento, objetiva otimizar a qualidade do som coletado, removendo eventuais ruídos de fundo e realizando cortes conforme necessário. Realizar uma etapa de pré-processamento é muito importante, pois influencia nos resultados de classificação e no desempenho dos classificadores utilizados.

Após pré-processamento, os áudios são utilizados para gerar espectrogramas, ou seja, são convertidos em imagens ou representações visuais das frequências dos sons. Ao gerar espectrogramas do som, o Módulo de Classificação é capaz de distinguir padrões na fala do indivíduo e, conseqüentemente detectar características que definem erros de pronúncia. A partir da geração dos espectrogramas, então, são aplicadas técnicas de extração de características, as quais serão usadas como entrada para o processo de classificação de pronúncia.

Por fim, a palavra pronunciada pela criança deve ser classificada como correta ou incorreta pelo método de classificação utilizado. Após a classificação, inicia-se o processamento do Módulo de Compreensão da arquitetura, responsável por compreender padrões de erro na pronúncia do indivíduo.



Figura 4.2 – Percepção Situacional e Módulo de Classificação.



Fonte: Da autora.

#### 4.2.2.2 Módulo de Compreensão

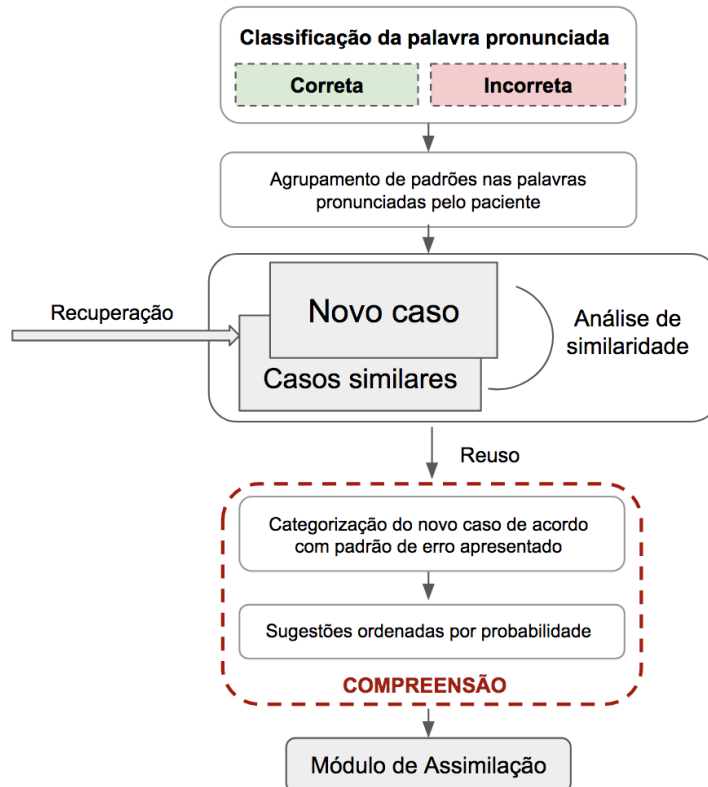
No Módulo de Compreensão, ocorre a comparação do caso atual/novo com casos anteriores já avaliados. Em outras palavras, é nesta etapa que ocorre as atividades de *Recuperação* e *Reuso*, atividades iniciais presentes no ciclo da metodologia CBR.

Conforme demonstrado na figura 4.3, o processamento neste módulo se inicia a partir da classificação realizada no módulo anterior. As pronúncias de um paciente classificadas como incorretas são avaliadas a fim de identificar padrões de erro, os quais serão agrupados para estabelecer tipos de casos fonológicos. Esses tipos de casos identificados são utilizados para categorizar um novo caso de acordo com seu padrão de erro.

Ao assumir um novo caso dentro do ciclo CBR, assume-se que este caso necessita de investigação, portanto, é realizada a *recuperação* de casos existentes na base de conhecimento do sistema. Uma vez que a utilidade de um caso recuperado não pode ser avaliada diretamente a priori, a semelhança entre descrições de problemas é usada para estimar a utilidade esperada dos casos sob o novo caso (KUO et al., 2015). Assim, após o processo de recuperação, avalia-se a similaridade entre o novo caso e os casos recuperados, a fim de identificar, em primeiro lugar, quais casos retêm o conhecimento mais útil para tipificar o caso de entrada.

Em etapa posterior, os casos similares recuperados serão *reutilizados*, ou seja, se-

Figura 4.3 – Módulo de Compreensão.



Fonte: Da autora.

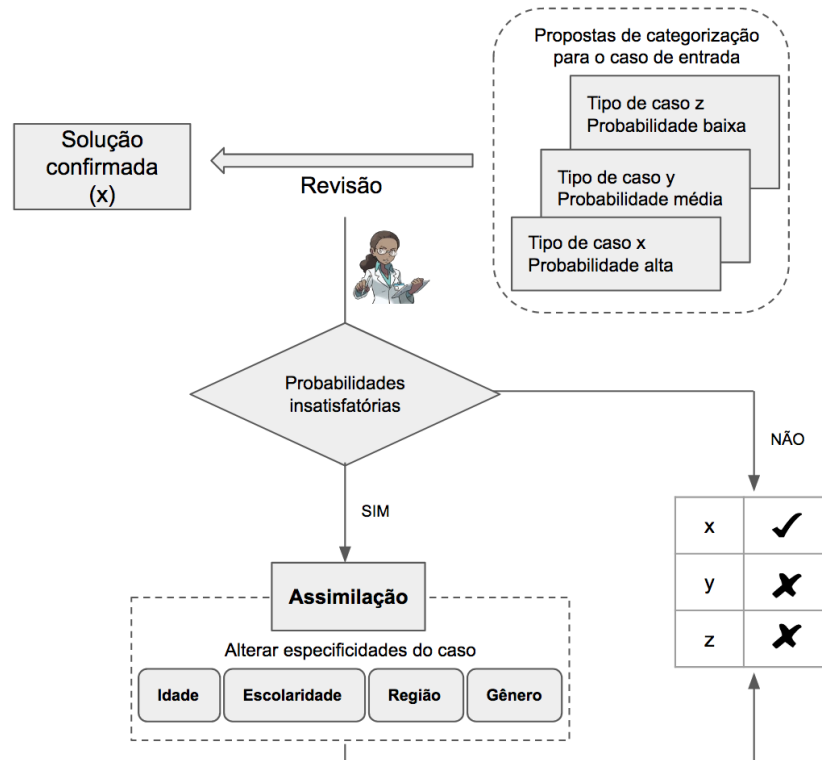
rão sugeridos como resolução para o caso de entrada conforme a similaridade de padrões de erro. O caso de entrada, então, será categorizado de acordo com o padrão de erro que apresenta. Este processo de categorização caracteriza o nível de Compreensão de SA, visto que, ao alcançar este nível, o sistema torna-se capaz de determinar o significado da informação do primeiro nível e integrá-lo para formar uma imagem abrangente do meio ambiente (NWIABU et al., 2011). De forma geral, a saída do Módulo de Compreensão é representada pela tipificação do caso dado como entrada, na qual as sugestões de tipos devem ser retornadas ao especialista humano por ordem de probabilidade de ocorrência. As sugestões são geridas pelo Módulo de Assimilação, abordado a seguir.

#### 4.2.2.3 Módulo de Assimilação

É desejável que sistemas conscientes de situação possam estimar ações adequadas a uma determinada situação e sugeri-las ao operador humano. Portanto, é no Módulo de Assimilação da arquitetura proposta que as sugestões retornadas no módulo anterior são avaliadas de forma a encontrar o tipo adequado ao caso de entrada. Nesta etapa, ocorre o processo de *revisão* do ciclo CBR, ou seja, as propostas são revisadas e a solu-

ção mais adequada é confirmada para utilização no novo caso. O Módulo de Assimilação é demonstrado na figura 4.4.

Figura 4.4 – Módulo de Assimilação



Fonte: Da autora.

Um sistema consciente de situação bem sucedido deve combinar poder computacional de máquina com cognição e intuição humanas (SALFINGER; RETSCHITZEGGER; SCHWINGER, 2013). Desta forma, há a participação de especialistas na revisão das classificações, como forma de utilizar o conhecimento humano para guiar o sistema no processo de aprendizagem.

No processo de revisão, o profissional fonoaudiólogo analisa os tipos sugeridos e as probabilidades com as quais eles foram inferidos. Se as probabilidades forem consideradas satisfatórias, o especialista escolhe o tipo mais apropriado e dá feedback ao sistema. Por outro lado, se as probabilidades forem insatisfatórias, ocorre a estratégia de *assimilação*.

Como visto na Seção 2, a assimilação descrita na Teoria do Desenvolvimento Cognitivo de Piaget é o processo onde o indivíduo tenta integrar um novo dado às estruturas cognitivas existentes (PIAGET, 1975). Assim, o processo de assimilação pode ser compreendido como uma *estratégia*, na qual é gerado um plano para encaixar novos dados na estrutura cognitiva atual. No contexto do Módulo de Assimilação, trata-se de alterar especificidades do caso quando sua probabilidade de ocorrência for baixa: o tipo sugerido a partir de casos prévios pode não se encaixar efetivamente ao novo caso, mas pode ser considerado se alguma particularidade do caso for alterada.

Para elucidar esta questão, pode-se citar um novo caso ao qual o sistema propôs ao especialista o enquadramento do tipo *A*, com base no padrão de erro *x* e garantindo baixa probabilidade. O especialista, então, revisa o caso e percebe que o padrão de erro *x* encontrado no caso não é comumente enquadrado ao tipo *A*, porém, o tipo *A* pode ser adequado ao caso se o padrão é comum de determinada região onde o paciente reside (por exemplo, o sotaque de um indivíduo pode determinar se sua pronúncia apresenta um padrão de erro ou não, dependendo da região em que ele se encontra). Desta forma, alterando especificidades de um caso através de uma estratégia de assimilação, pode-se incrementar o raciocínio do sistema e as variáveis analisadas no processamento.

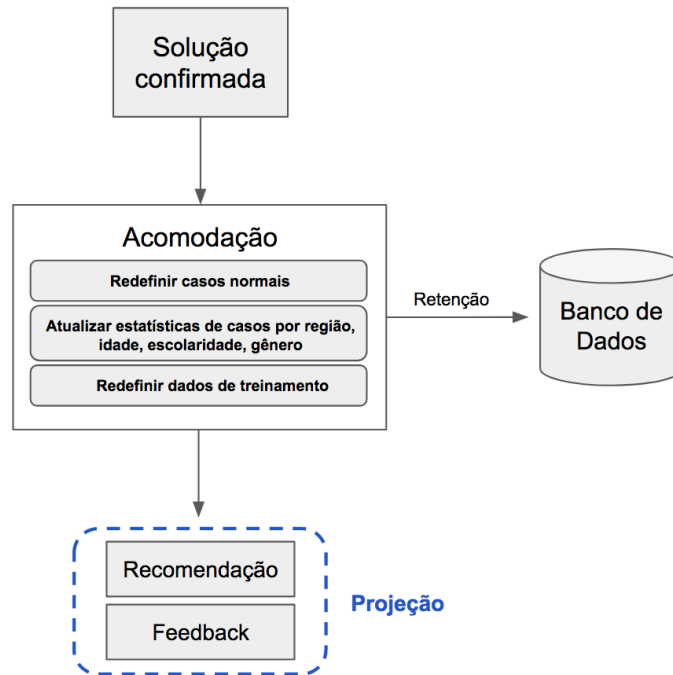
De forma geral, a avaliação dos casos feita pelo sistema e a revisão das sugestões realizada pelo especialista humano permitem que o nível de compreensão adquirido no módulo anterior seja incrementado, visto que ocorre uma comparação de diferentes soluções e a descoberta de qual é a mais relevante para a situação atual. A partir da confirmação de um tipo de caso dada pelo profissional fonoaudiólogo, é possível derivar relações entre as informações adquiridas e interpretar como a situação compreendida irá impactar nos objetivos do usuário. A partir desta sequência de atividades, o Módulo de Acomodação lida com a solução confirmada e seu impacto na estrutura atual do sistema.

#### 4.2.2.4 *Módulo de Acomodação*

O Módulo de Acomodação é responsável por receber a solução confirmada obtida no Módulo de Assimilação e avaliar como ela pode ser processada a fim de gerar a recomendação adequada ao usuário. A solução confirmada, neste contexto, refere-se à situação do paciente interpretada a partir do tipo de caso no qual ele se enquadra. Tem-se, portanto, o enquadramento do caso-problema como um novo dado a ser incluído na estrutura do sistema ou esquema atual e a necessidade de estabelecer qual a ação mais indicada para lidar com o caso. Nesta etapa, ocorre o processo de *Acomodação*: a modificação da estrutura atual para inclusão da informação previamente assimilada (PIAGET, 1970a). Em outras palavras, depois que o caso do paciente foi assimilado (diferenciado dos demais casos em histórico e devidamente categorizado segundo os padrões da fonoaudiologia), é necessário acomodar este caso à estrutura do sistema, modificando-a através da execução de três tarefas:

- **redefinição de casos normais:** Como visto anteriormente, o caso-problema dado como entrada no ciclo CBR é interpretado para que se tenha compreensão da situação do paciente. Em outras palavras, o caso avaliado é classificado de acordo com tipos fonológicos que abrangem uma situação de normalidade (o paciente não possui desordens da fala) ou anormalidade (paciente que possui atraso de fala ou erros questionáveis em sua pronúncia). Ao obter essa classificação é preciso avaliar se o

Figura 4.5 – Módulo de Acomodação.



Fonte: Da autora.

conceito de caso normal e sua composição na base de dados do sistema necessita de mudanças. Por exemplo, se um caso avaliado é de tipo considerado normal dentro de seu contexto e não há, na base de dados, registro de casos normais da mesma natureza, então a composição de casos normais deve ser alterada de forma a incluir o caso recente. Da mesma forma, se a natureza do caso avaliado e sua respectiva solução já constam no histórico no sistema, a redefinição de casos não é necessária;

- **atualização de estatísticas do sistema:** Antes do caso-problema ser dado como entrada no ciclo CBR, o sistema mantém algumas estatísticas sobre os casos avaliados anteriormente (por exemplo, "70% de substituições de fonema x em pacientes de determinada faixa etária referem-se a casos de atraso de fala relacionado ao desenvolvimento"). Após avaliar o novo caso e estabelecer uma solução adequada a ele, o sistema deve recalcular as estatísticas com base nas novas informações e, se necessário, realizar modificações atentas a características de região, idade, etc;
- **redefinição de dados de treinamento:** O ciclo CBR lida com casos de forma mais eficiente à medida que novas informações são armazenadas no banco de dados. O sistema é treinado para sugerir soluções ao novo caso de acordo com situações semelhantes previamente analisadas por ele. Assim, após a solução do caso ser confirmada, é necessário redefinir os casos usados como base para o raciocínio. Com uma base de conhecimento sempre atualizada, o sistema se torna eficiente em compreender situações futuras e sugerir as soluções mais convenientes.

Concluído o processo de Acomodação, o caso (abrangendo todas as suas características) é *armazenado* ou *retido* no banco de dados do sistema, completando o ciclo da metodologia CBR. Por fim, ocorre a etapa de Projeção, ou seja, a tomada de ações requeridas como resposta à situação interpretada. Nesta última etapa do módulo, são definidas as recomendações que serão dadas ao fonoaudiólogo, ou seja, quais atividades ele pode desenvolver com o paciente ou quais estratégias ele pode tomar a fim de lidar com o caso identificado. A partir da recomendação retornada pelo sistema, o especialista pode tomar decisões melhor informadas com base no perfil do paciente e pode dar um feedback ao sistema, informando se as recomendações foram úteis.

## 5 AVALIAÇÃO E RESULTADOS

Nesta seção, são apresentados detalhes de avaliação da arquitetura e os resultados obtidos em cada módulo componente. Em uma primeira etapa, são demonstradas avaliações abrangendo o Módulo de Percepção Situacional, ao passo que em etapas posteriores, avaliações e resultados dos módulos de Classificação e Compreensão são apresentados.

### 5.1 MÓDULO DE PERCEPÇÃO SITUACIONAL

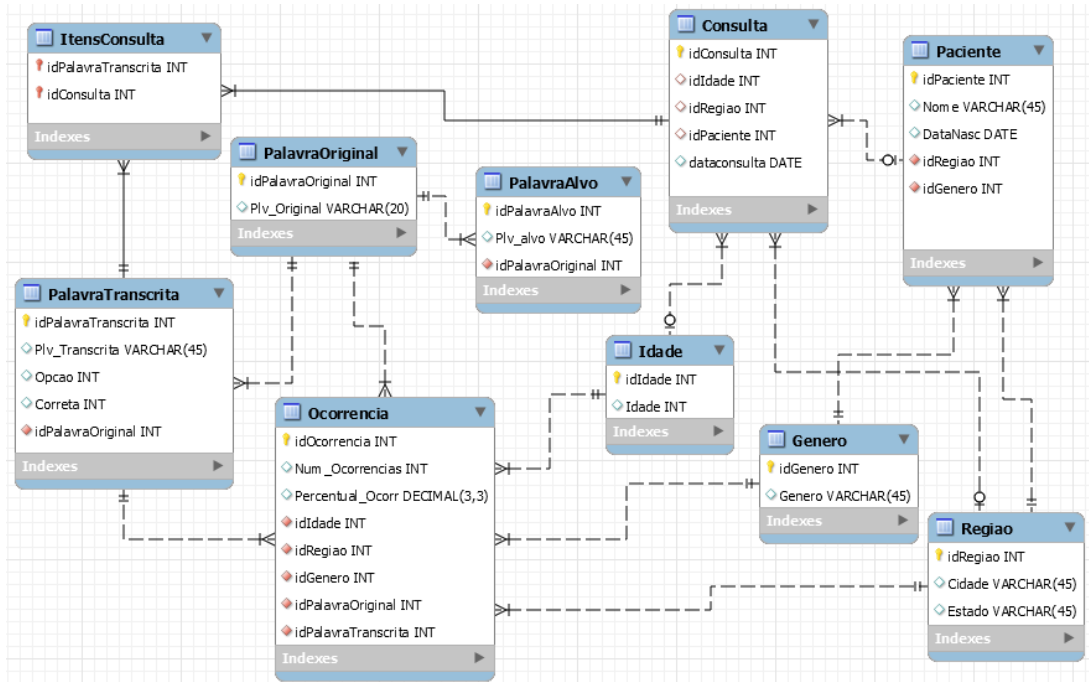
Para compreender efetivamente uma situação corrente, um sistema consciente de situação deve estabelecer uma fase de percepção, que abrange a captura de pistas sensoriais do ambiente circundante. A seguir, serão apresentados detalhes sobre a coleta de dados do ambiente, mais especificamente, dados de fala de crianças em avaliação fonoaudiológica.

#### 5.1.1 Coleta das Amostras de Fala

A fim de atingir a percepção de situação na arquitetura proposta, sinais de fala foram coletados por meio de avaliações fonoaudiológicas realizadas com o público infantil. Os dados são provenientes de avaliações realizadas por equipe de fonoaudiólogas do Programa de Pós-Graduação em Distúrbios da Comunicação Humana (PPGDCH) da Universidade Federal de Santa Maria, e abrangem um total de 133.143 casos de avaliações fonoaudiológicas realizadas com 1.362 crianças de idade entre 3 a 8 anos e 11 meses realizando tarefas de nomeação de imagens correspondentes a 84 palavras-alvo. Durante as tarefas de nomeação, cada criança foi apresentada a uma imagem (referente a uma palavra-alvo), a partir da qual deveria pronunciar a palavra correspondente. Aos arquivos de áudio provenientes de todas as avaliações foram armazenados em banco de dados, modelado conforme a figura 5.1 e organizado a partir de uma série de informações:

- informações da criança (paciente), tais como nome, data de nascimento, região, gênero e idade;
- informações de consulta (avaliação), como data de avaliação, paciente avaliado, palavra-alvo avaliada e palavra transcrita a partir da avaliação;
- informações sobre a palavra-alvo avaliada (foi definido o número 1 para os casos onde a pronúncia foi realizada corretamente e 0 para pronúncia realizada incorretamente);

Figura 5.1 – Banco de dados utilizado para gerenciar o corpus de fala.



Fonte: Da autora.

- informações estatísticas sobre as palavras avaliadas, no que se refere ao número de ocorrências durante as avaliações.

A partir dessas informações coletadas, deu-se início ao processamento conduzido no Módulo de Classificação, o qual abrange etapas de pré-processamento de dados e análises. Como primeira análise realizada no Módulo de Classificação, a ferramenta *Google Cloud Speech* foi utilizada para realizar reconhecimento de voz em avaliações de 20 palavras-alvo específicas. Detalhes sobre a condução do experimento são abordados na sequência.

## 5.2 MÓDULO DE CLASSIFICAÇÃO

O Módulo de Classificação que compõe a arquitetura proposta abrange uma série de etapas que, por sua vez, incluem desde avaliações realizadas sem pré-processamento de dados, até a classificação da pronúncia utilizando dados pré-processados. No próximo item, são abordados resultados obtidos a partir dos dados não processados, com o intuito de avaliar a qualidade da fala coletada e a eficácia da ferramenta de análise.



### 5.2.1 Avaliação utilizando a API *Google Cloud Speech*

Neste tópico é abordada a análise realizada com dados de fala não processados, utilizando a API (*Application Programming Interface*) de reconhecimento de voz *Google Cloud Speech* (neste trabalho, a ferramenta será abreviada por meio da sigla *GCS*). A API *GCS* realiza o reconhecimento de voz convertendo áudio para texto através da tecnologia de aprendizado de máquina. Mais especificamente, a ferramenta aplica modelos avançados de rede neural e é capaz de realizar transcrições de voz em uma ampla variedade de idiomas. Uma vez que a API é uma maneira simples para os desenvolvedores integrarem recursos de reconhecimento de fala em suas aplicações (BALLINGER et al., 2010), pesquisas recentes usaram essa tecnologia em suas metodologias. Pode-se mencionar, por exemplo, a proposta de Mohamed, Hassanin e Othman (2014), que abordou um ambiente educacional para pessoas cegas e portadoras de deficiência.

A análise sem pré-processamento de dados objetivou verificar a qualidade da fala coletada, bem como investigar a adequação da API *GCS* para realizar avaliações fonoaudiológicas iniciais, as quais visam classificar os pacientes como indivíduos que possuem desordens da fala ou indivíduos saudáveis. Para isso, um conjunto de 20 palavras-alvo em português (brasileiro) foi selecionado do corpus de fala coletado para avaliar a capacidade de pronúncia infantil por meio da API *GCS*. A API analisou cada arquivo de áudio proveniente das tarefas de nomeação, informando ao profissional de fonoaudiologia se a pronúncia realizada pela criança foi correta ou incorreta, e retornando a transcrição do que foi compreendido.

De um total de 31.752 avaliações realizadas com 20 palavras-alvo, a *GCS* retornou um resultado de transcrição para 11.641 delas. Para o restante das avaliações (20.111), a ferramenta não foi capaz de compreender a sentença falada. A tabela 5.1 mostra os resultados obtidos a partir dos 11.641 casos em que houve resposta da API utilizada. Foi considerado, para cada palavra-alvo:

- **GCS1SLP1:** número de avaliações em que a API *GCS* considerou a pronúncia CORRETA (1) e a profissional de fonoaudiologia (SLP) considerou CORRETA (1).
- **GCS1SLP0:** número de avaliações em que a API *GCS* considerou a pronúncia CORRETA (1) e a profissional de fonoaudiologia (SLP) considerou INCORRETA (0).
- **GCS0SLP1:** número de avaliações em que a API *GCS* considerou a pronúncia INCORRETA (0) e a profissional de fonoaudiologia (SLP) considerou CORRETA (1).
- **GCS0SLP0:** número de avaliações em que a API *GCS* considerou a pronúncia INCORRETA (0) e a profissional de fonoaudiologia (SLP) considerou INCORRETA (0).

Foi considerada uma taxa de concordância *TC* composta por casos em que a API e o terapeuta consideraram a pronúncia da criança correta somada ao número de casos

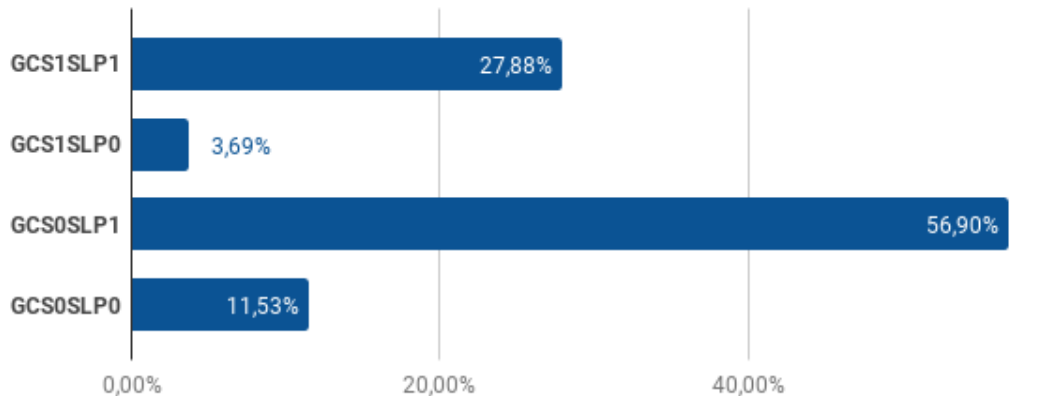
Tabela 5.1 – Avaliação de palavras-alvo com *Google Cloud Speech*

Palavra-Alvo	Fonética	Avaliações	GCS1SLP1	GCS1SLP0	GCS0SLP1	GCS0SLP0
Caminhão	kamiŋəo	631	310	10	293	18
Cachorro	kaʃoɣo	586	249	21	265	51
Bebê	bebe	716	271	5	432	8
Casa	kazə	473	148	2	292	31
Jacaré	zakare	616	182	42	294	98
Cama	kəmə	462	135	3	309	15
Cavalo	kavalu	501	146	6	310	39
Coelho	koeʎo	499	136	6	306	51
Jornal	ʒornaw	667	174	42	311	140
Cabelo	kabelu	797	203	10	547	37
Sofá	sofa	394	97	2	261	34
Bicicleta	bisikletə	543	132	38	221	152
Relógio	ɣelɔzu	462	110	12	246	94
Gato	gato	454	107	3	313	31
Batom	batõw	530	122	2	394	14
Galinha	galiŋə	622	142	14	442	24
Cobra	kɔbrə	578	130	30	277	141
Microfone	mikrofone	831	180	146	224	281
Folha	foʎə	554	119	17	382	36
Barriga	baɣigə	725	152	18	507	48
<b>TOTAL</b>		<b>11.641</b>	<b>3.245</b>	<b>429</b>	<b>6.624</b>	<b>1.343</b>

Fonte: Da autora.

em que ambos consideraram a pronúncia incorreta. Assim, tem-se  $TC = GCS1SLP1 + GCS0SLP0$ . Da mesma forma, estabeleceu-se uma taxa de discordância  $TD$  composta por casos em que a API e o terapeuta consideraram resultados diferentes para o enunciado analisado, assim,  $TD = GCS1SLP0 + GCS0SLP1$ .

Gráfico 5.1 – Porcentagem de concordância (1-1, 0-0) e discordância (0-1, 1-0) entre a API *Google Cloud Speech* (GCS) e o especialista (SLP).



Fonte: Da autora.

Conforme mostrado no gráfico 5.1, para o total de 11.641 avaliações que a API GCS retornou um resultado de transcrição, TC atingiu 39.41% dos casos, ao passo que houve uma TD de 60.59%. Pode-se observar que os casos em que a pronúncia da criança foi considerada errada pela API e correta pelo terapeuta (GCS0SLP1) atingem a maior porcentagem dentre o comparativo demonstrado: 56,90%. Observa-se, também, que a taxa de discordância atingiu um valor bastante alto (maior que 50%), demonstrando que a maioria das respostas provenientes da API não coincidem com as respostas dadas pelo profissional de fonoaudiologia responsável pelas avaliações de pronúncia.

Considerando os resultados obtidos na avaliação realizada pela ferramenta GCS, foi possível inferir uma composição de dados ruidosos e inconsistentes no corpus de fala utilizado. Assim, estratégias alternativas fizeram-se necessárias no Módulo de Classificação, como forma de otimizar a robustez dos dados e a classificação da pronúncia coletada. Tais estratégias são abordadas na sequência.

## 5.2.2 Pré-processamento dos dados de fala

As taxas de concordância e discordância obtidas a partir da avaliação realizada com a API GCS apresentaram resultados insatisfatórios, ou seja, a ferramenta não foi capaz de classificar corretamente a pronúncia infantil com base nas avaliações realizadas por terapeutas. A qualidade dos dados coletados na fase de Percepção justificam o resultado

---

**Algoritmo 1** Algoritmo de pré-processamento de áudio.
 

---

**Entrada:** áudio coletado da palavra alvo  $A$

**Entrada:** porcentagem de tolerância em relação à duração média da palavra  $P$

```

1:  $A \leftarrow \text{noiseRemoval}(A)$ ;
2:  $WAD \leftarrow \text{getWordAverageDuration}(A)$ ;
2:  $\text{ruído} \leftarrow \text{getNoiseLevel}(A)$ ;
3: faça
4:    $\text{inicio}, \text{fim} \leftarrow \text{detectLastSilent}(A, \text{ruído})$ ;
5:    $\text{ruído}++$ ;
6:    $\text{dur} = \text{fim} - \text{inicio}$ ;
7: enquanto ((( $\text{dur} > WAD * P$ ) OU ( $\text{dur} < WAD * P$ )) E ( $\text{dur} \neq 0$ ));
8: se ( $\text{dur} == 0$ ) então
9:   retorna NULL;
10: fim se
11:  $A \leftarrow \text{splitChildWord}(A, \text{inicio}, \text{fim})$ ;
12: retorna  $A$ 

```

---

obtido, pois os áudios das avaliações de fala eram ruidosos e de baixo volume, dificultando seu entendimento. Outra possível razão para a pequena TC obtida, é que os áudios frequentemente continham uma sequência inteira pronunciada pela criança, e não apenas a palavra desejada (alvo da avaliação).

Visto que não foi aplicada nenhuma técnica de pré-processamento, filtragem ou ajuste nesses dados analisados, considerou-se que a taxa de concordância poderia atingir valores mais altos com a inclusão de estratégias para lidar especificamente com dados incompletos, ruidosos e inconsistentes. Assim, foi incluída no Módulo de Classificação uma etapa de pré-processamento de dados, a fim de melhorar a classificação de fala de acordo com a avaliação do especialista. Os sinais de fala coletados foram pré-processados por meio de duas ferramentas de código aberto: SOX (v14.4.1) e FFmpeg (v14.4.1).

SOX<sup>1</sup> (*Sound eXchange*) é uma ferramenta conveniente para utilização em análises de fala, visto que permite realizar facilmente conversões de arquivo (MATHEW; ANSELAM; PILLAI, 2014). A fim de lidar com o ruído presente nos áudios coletados, este recurso foi utilizado com as opções *noiseprof* e *noisered*. A opção *noiseprof* é responsável por calcular o perfil do áudio a partir de um trecho de silêncio que contém o ruído. Este trecho de silêncio/ruído geralmente é encontrado no início ou ao final do arquivo de áudio, e é usado pelo recurso *noisered*, que efetivamente remove o ruído de acordo com o perfil identificado no primeiro passo.

Assim como SOX, a ferramenta FFmpeg<sup>2</sup> também foi utilizada para pré-processamento de dados. O software é composto por um conjunto de bibliotecas que fornecem funcionalidades para codificação, decodificação, conversão de formatos e filtragem de conteúdo de áudio e vídeo (RADOVIĆ et al., 2017). No Módulo de Classificação, o software foi utilizado

---

<sup>1</sup>Disponível em: <http://sox.sourceforge.net/>

<sup>2</sup>Disponível em: <https://ffmpeg.org/>

para detectar sentenças da seguinte forma: primeiro, o comprimento do áudio foi medido; na ocorrência de um comprimento maior que o comprimento médio de determinada palavra, procurou-se momento de silêncio usando a opção *silencedetect*. No caso de detecção de múltiplos momentos de silêncio, considerou-se o áudio como uma sentença, assim, o áudio foi cortado para isolar apenas a palavra desejada. O corte de áudio foi realizado com as opções *ss* e *t* da ferramenta FFmpeg. O método de pré-processamento aplicado encontra-se descrito no Algoritmo 1.

### 5.2.3 Geração de Espectrogramas

Após realizar pré-processamento dos áudios coletados, observou-se a necessidade de identificar a palavra pronunciada pela criança no momento da avaliação fonoaudiológica. Consequentemente, foi necessário distinguir padrões de erro na pronúncia que permitissem diferenciar uma palavra pronunciada incorretamente de uma palavra pronunciada corretamente (palavra alvo sugerida pelo especialista), para fins de classificação.

Uma nova tendência utilizada no reconhecimento de fala é o uso de espectrogramas para extração de características, onde o espectrograma representa visualmente a frequência de tempo do sinal de fala (PRASOMPHAN, 2015). Em outras palavras, um espectrograma mostra o áudio em forma de imagem, o que permite distinguir padrões que mostram diferentes características da fala do indivíduo. As propriedades principais de um espectrograma são a frequência e a amplitude de som que, quando vistos de forma conjunta, criam uma assinatura de som única que pode ser utilizada para detectar padrões de erro no sinal de fala.

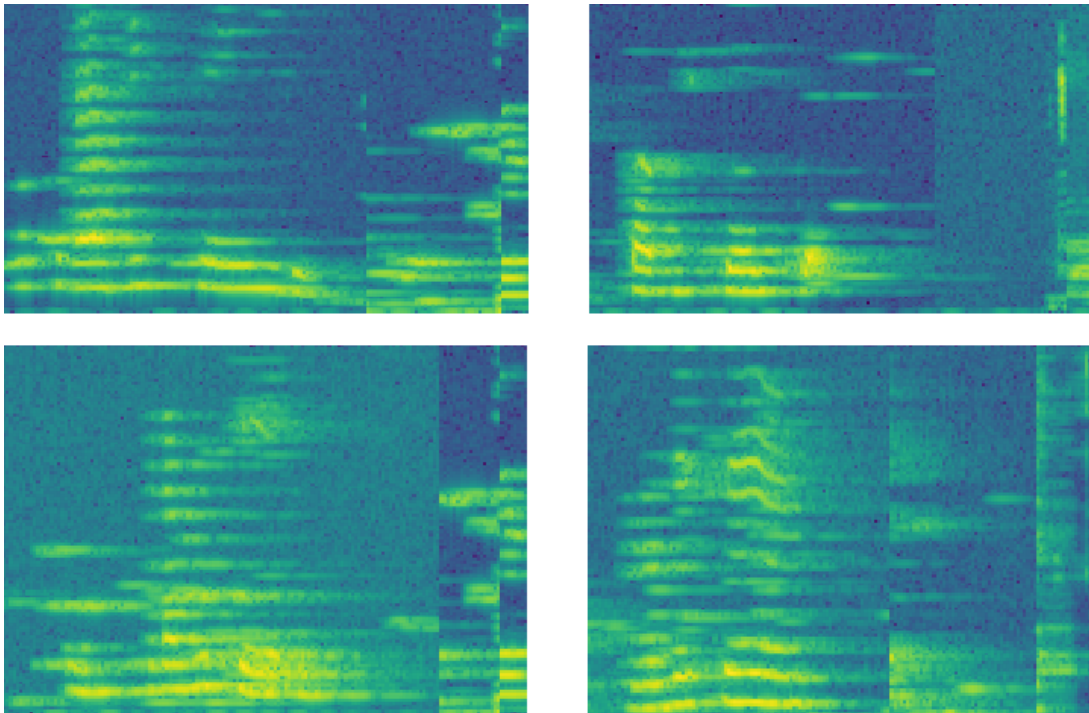
Para gerar espectrogramas dos sinais de fala, foi usado *matplotlib* (v2.2.2)<sup>3</sup>, um pacote Python para plotagem, que gera gráficos de qualidade a partir de poucos comandos (ARI; USTAZHANOV, 2014). Com o pacote *matplotlib*, foi possível gerar imagens para representar visualmente as frequências do som da fala ao longo do tempo. Na Figura 5.2 a seguir são apresentados exemplos de espectrogramas gerados a partir da pronúncia correta e incorreta da mesma palavra (travesseiro).

Para reduzir a complexidade computacional envolvida, os espectrogramas foram convertidos para escala de cinza. Após esta etapa, características foram extraídas das imagens, a fim de diferenciar padrões nas palavras pronunciadas. Os passos efetuados na etapa de extração de características são descritos na sequência.

---

<sup>3</sup>Disponível em: <https://matplotlib.org>

Figura 5.2 – Espectrogramas referentes à palavra *travesseiro* falada corretamente (à esquerda) e incorretamente (à direita).



Fonte: Da autora.

#### 5.2.4 Extração de Características

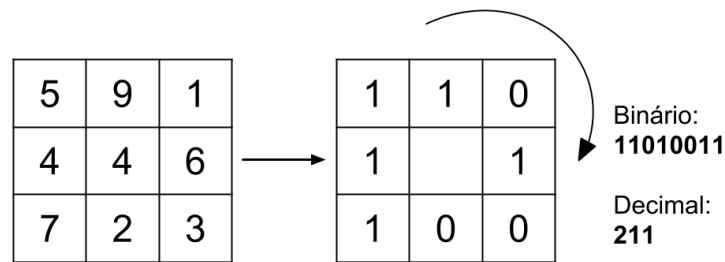
Uma vez que as frequências dos sons foram convertidas para padrões visuais, foi aplicado o método LBP (*Local Binary Pattern*), um descritor visual amplamente utilizado para identificação de texturas (OJALA; PIETIKAINEN; MAENPAA, 2002).

O vetor de características LBP é primeiramente criado a partir da divisão da textura examinada em células. O valor de cada pixel em uma célula é comparado com suas células vizinhas, começando da parte superior esquerda e seguindo uma ordem no sentido horário. Se o valor do pixel é maior que o valor da célula vizinha, escreve-se "0", do contrário, escreve-se "1", resultando em um número binário de 8 dígitos, isto é, o valor LBP calculado para aquela célula (TROIS et al., 2018). Esta operação, descrita na Figura 5.3, é feita com todas as células do espectrograma, criando um vetor de características 256-dimensional que contém informações de textura do espectrograma.

Para criar histogramas de vetores de características LBP, foi usada a biblioteca Python *Scikit-Learn* (v0.19.1)<sup>4</sup>. Assim, os histogramas são usados como entrada para o processo de classificação, descrito no tópico a seguir.

<sup>4</sup>Disponível em: <http://scikit-learn.org>

Figura 5.3 – Exemplo de operação do método LBP.



### 5.2.5 Resultados de Classificação

Considerando os resultados alcançados com a API *Google Cloud Speech* descritos no item 5.2.1, optou-se pela utilização de técnicas mais robustas para classificar dados de fala. Em virtude do dinamismo exigido em sistemas conscientes de situação, técnicas de aprendizado de máquina (*Machine Learning*, ou ML) foram utilizadas, as quais lidam com tarefas preditivas que incluem a definição de quais comportamentos têm a maior tendência de gerar resultados preferidos (NITHYA; ILANGO, 2017).

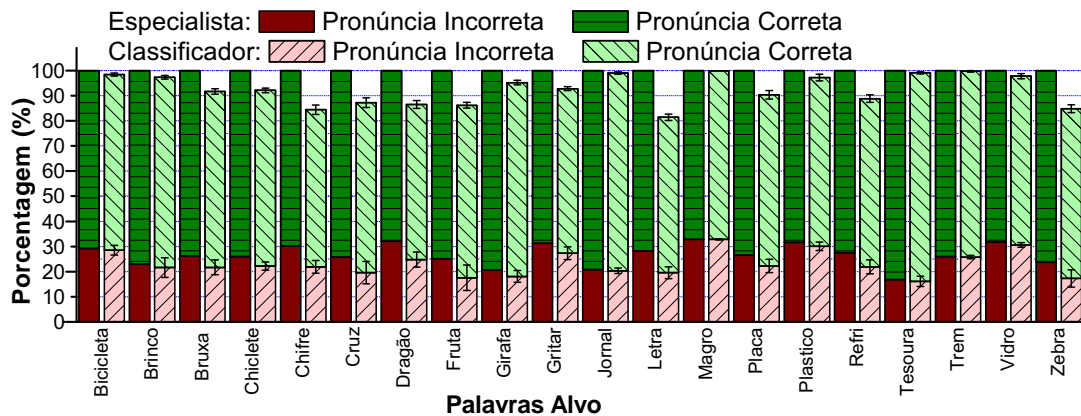
Para realizar a classificação proposta, foram testadas as precisões de quatro métodos: kNN, SVM, rede neural Adams e *Decision Tree* (Árvore de Decisão). Optou-se por utilizar o algoritmo *Decision Tree*, visto que apresentou taxas de precisão mais satisfatórias e estáveis. O método é essencialmente caracterizado pelo fato de que uma amostra desconhecida pode ser classificada em uma classe usando uma ou várias funções de decisão de maneira sucessiva (SWAIN; HAUSKA, 1977), e é aplicado em várias propostas envolvendo reconhecimento de fala, por exemplo, nos trabalhos de Li e Wu (2014) e Yüncü et al. (2014).

O classificador *Decision Tree* foi utilizado em um total de 27.000 amostras do corpus de fala. Para prevenir que a classificação fosse tendenciosa, um total de 10 classificações foram feitas dividindo aleatoriamente o corpus de fala: metade do corpus foi aplicado como conjunto de treinamento do classificador, e a outra metade foi aplicada para teste. Isto significa dizer que, para cada execução, 13.500 entradas foram usadas para treinamento e 13.500 foram usadas para testar o classificador.

O classificador foi treinado para reconhecer a pronúncia de cada palavra como correta ou incorreta, com base no diagnóstico feito pelo fonoaudiólogo. Os resultados de classificação são apresentados na Tabela 5.2 e nos gráficos 5.2 e 5.3. O gráfico 5.2 mostra, no eixo horizontal, todas as palavras alvo escolhidas para realizar a classificação. Para cada palavra há duas colunas: a primeira (de cor mais escura) fornece informação sobre a avaliação do fonoaudiólogo, mostrando a porcentagem de pronúncias corretas e incorretas armazenadas na base de dados. A segunda coluna (de cor mais clara) mostra a precisão média de classificação e o desvio padrão do método *Decision Tree*.

Seguindo a mesma análise usada na avaliação com a API GCS, a predição do

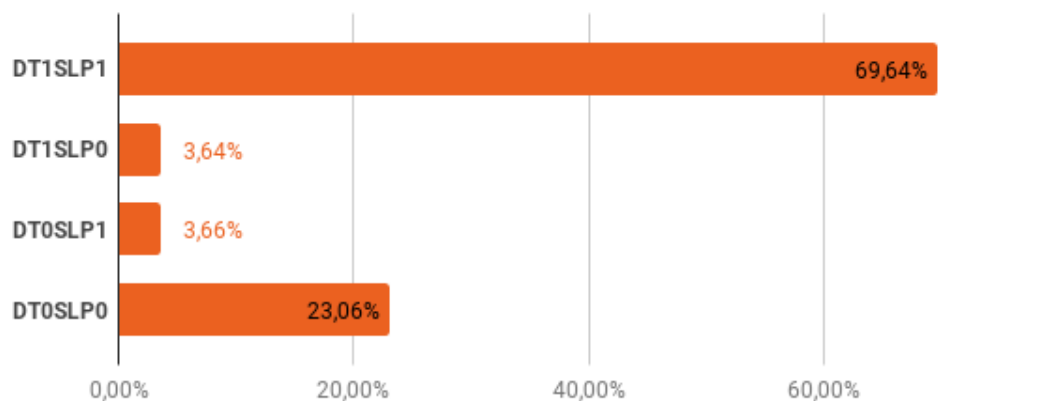
Gráfico 5.2 – Precisão de classificação a partir dos espectrogramas.



Fonte: Da autora.

classificador foi comparada com a avaliação do fonoaudiólogo, considerando uma taxa de concordância (TC) para os casos em que houve concordância entre o método e o especialista. Assim, estabeleceu-se  $TC = DT1SLP1 + DT0SLP0$ , onde DT representa o método Decision Tree. Da mesma forma, foi considerada uma taxa de discordância TD nos casos em que a classificação não esteve de acordo com a avaliação do fonoaudiólogo ( $TD = DT1SLP0 + DT0SLP1$ ). Como demonstrado no gráfico 5.3, do total de 13.500 avaliações feitas pelo especialista humano e classificadas por método de aprendizagem de máquina, houve uma concordância entre o método e o especialista em aproximadamente 12.515 dos casos, ou seja, 92.7% (TC). Por outro lado, houve discordância em aproximadamente 1015 das avaliações, significando uma taxa TD de 7.3%.

Gráfico 5.3 – Porcentagem de concordância (1-1, 0-0) e discordância (0-1, 1-0) entre o método Decision Tree (DT) e o especialista (SLP).



Fonte: Da autora.

Vale mencionar que o gráfico 5.3 e a tabela 5.2 mostram resultados de uma das execuções, o que quer dizer que as porcentagens podem variar para mais ou para menos. A precisão média considerando todas as execuções foi de 92.5%. Ainda, observa-se que as



Tabela 5.2 – Avaliação de palavras-alvo com Decision Tree

<b>Palavra-Alvo</b>	<b>Avaliações</b>	<b>DT1SLP1</b>	<b>DT1SLP0</b>	<b>DT0SLP1</b>	<b>DT0SLP0</b>
Bicicleta	676	474	3	4	194
Brinco	687	521	9	8	149
Bruxa	682	472	18	31	161
Chiclete	680	478	26	25	151
Chifre	669	421	43	46	159
Cruz	692	474	46	39	133
Dragão	675	426	51	32	166
Fruta	664	461	45	36	122
Girafa	709	545	15	18	131
Gritar	659	435	28	18	178
Jornal	692	546	6	2	138
Letra	658	406	60	66	126
Magro	656	440	1	0	215
Placa	683	461	43	40	139
Plástico	633	423	6	10	194
Refri	651	422	30	50	149
Tesoura	704	586	4	0	114
Trem	689	508	3	2	176
Vidro	664	443	8	10	203
Zebra	678	460	46	57	115
<b>TOTAL</b>	<b>13.500</b>	<b>9.402</b>	<b>491</b>	<b>494</b>	<b>3.113</b>

Fonte: Da autora.

palavras alvo escolhidas para avaliação são diferentes daquelas avaliadas com a API Google Cloud Speech. Isso porque algumas palavras resultavam em uma taxa muito grande de acertos e, portanto, eram inadequadas para treinar o classificador na identificação de erros.

A partir destes resultados, é iniciada a análise do Módulo de Compreensão da arquitetura, abrangendo conceitos da Teoria do Desenvolvimento Cognitivo e da metodologia CBR. O Módulo de Compreensão é abordado detalhadamente a seguir.

### 5.3 MÓDULO DE COMPREENSÃO

No Módulo de Compreensão, as pronúncias de um paciente classificadas como corretas e incorretas no módulo anterior são avaliadas a fim de identificar padrões, os quais serão agrupados para estabelecer tipos de casos fonológicos. A partir destes tipos de casos identificados, novos casos podem ser classificados com base nos similares previamente avaliados. Desta forma, ocorre as etapas de Recuperação e Reuso da metodologia CBR, bem como a estratégia de assimilação da TDC. A seguir, são abordadas etapas de processamento realizadas no Módulo de Compreensão, que referem-se à agrupamento de padrões, recuperação de casos e classificação segundo tipos fonológicos.

#### 5.3.1 Agrupamento de Padrões

No módulo de Classificação, pronúncias de um indivíduo são classificadas como corretas e incorretas. A partir desta classificação, o módulo de Compreensão objetiva identificar ocorrências de padrões de erros nas pronúncias e, assim, estabelecer tipos de casos fonológicos de acordo com tais padrões. Realizar agrupamento de padrões de erro é uma estratégia importante para que, dado um novo caso a ser avaliado pelo sistema de recomendação, o mesmo possa classificá-lo como pertencente a determinado tipo fonológico, com base em padrões de erro previamente identificados.

O processo de agrupamento de padrões do Módulo de Compreensão baseia-se no termo *estratégias de reparo* que, no contexto da fonoaudiologia, refere-se a estratégias empregadas pelas crianças para adequar a realização da língua falada pelos adultos ao seu sistema fonológico (BAESSO, 2009). Em outras palavras, durante a aquisição fonológica, as crianças apresentam substituições para alguns fonemas que, para elas, são de difícil produção, sendo estas substituições denominadas estratégias de reparo (WIETHAN; MOTA, 2012). Estratégias de reparo mais comuns utilizadas por crianças no processo de aquisição da fala são demonstradas na Tabela 5.3, que apresenta as possíveis trocas, omissões ou substituições de fonemas na pronúncia infantil.

Tabela 5.3 – Estratégias de reparo utilizadas no processo de aquisição da fala

Estratégia de Reparo	Classe (possibilidade)	Fonemas
<b>Substituição de líquidas</b> <sup>6</sup>	Líquida <sup>5</sup> (todas)	l, λ, r, R → l, λ, r, R
	Líquida (l, r)	l, r → l, r
<b>Semivocalização de líquida</b> <sup>7</sup>	Líquida (todas)	l, λ, r, R ou outra → j ou y
<b>Dessonorização</b> <sup>9</sup>	Plosiva <sup>8</sup> (b, d, g)	b, d, g, v, z, ʒ → p, t, k, f, s, ʃ
	Fricativa <sup>10</sup> (v, z, ʒ)	
<b>Sonorização</b> <sup>11</sup>	Plosiva (p, t, k)	p, t, k, f, s, ʃ → b, d, g, v, z, ʒ
	Fricativa (f, s, ʃ)	
<b>Anteriorização</b> <sup>12</sup>	Plosiva (k, g)	s, ʒ → s, z
	Fricativa (ʃ, ʒ)	k, g → t, d
<b>Posteriorização</b> <sup>13</sup>	Plosiva (p, b, t, d)	s, z → s, ʒ
	Fricativa (f, v)	t, d → k, g
<b>Plosivização</b> <sup>14</sup>	Fricativa (todas)	f, v, s, z, ʃ, ʒ → p, b, t, d, k, g
	Líquida (todas)	l, r, R, λ → p, b, t, d, k, g
<b>Redução do EC</b> <sup>15</sup>	Apagamento do r ou l do EC	l ou r do EC → ∅
<b>Apagamento de fricativa final</b> <sup>16</sup>	Fricativa (s)	s → ∅
<b>Apagamento de líquida final</b> <sup>17</sup>	Líquida (r, l)	r, l → ∅
<b>Apagamento de líquida intervocálica</b> <sup>18</sup>	Líquida (r, l, λ, R)	r, l, λ, R → ∅
<b>Apagamento de líquida inicial</b> <sup>19</sup>	Líquida (l, R)	l, R → ∅
<b>Metátese</b> <sup>20</sup>	EC (todos)	ex: "vidru" ocorre dr → r/du
<b>Epêntese</b> <sup>21</sup>	EC (todos)	[bi.bili.o.'te.ta], bl por bili

Fonte: Adaptado de (YAVAS; HERNANDORENA; LAMPRECHT, 1991).

O Módulo de Compreensão, por meio do processo de agrupamento de padrões, é responsável por (1) receber a pronúncia da criança, (2) analisar se houve troca, substituição e/ou omissão de fonemas, (3) verificar em que posição da palavra as trocas, substituições e/ou omissões ocorreram e (4) classificar a pronúncia com base nas estratégias de reparo existentes.

Por exemplo, pode-se considerar um caso em que a criança pronunciou incorretamente o fonema /l/. O padrão de erro encontrado em sua pronúncia refere-se a que tipo de caso fonológico ou estratégia de reparo? De uma visão geral, é possível que a criança seja classificada em mais de uma estratégia - substituição de líquidas, semivocalização de líquidas, plosivização ou redução de EC (Encontro Consonantal), por exemplo. Assim, faz-se necessário agrupar padrões encontrados para /l/ a fim de realizar a classificação correta (se /l/ foi substituído por /p/ é um caso de *plosivização*, se /l/ foi omitido no final da palavra é um caso de *apagamento de líquida final*, e assim por diante).

### 5.3.2 Recuperação de Casos

Na metodologia CBR, são recuperados casos que se assemelham a descrição do problema dado como entrada e, a partir desta recuperação, uma solução pode ser reutilizada. Considerando que uma criança, ao produzir uma sentença, pode empregar uma ou mais estratégias de reparo para adequar-se a língua falada em seu meio, estas estratégias de reparo são tidas como casos que, uma vez recuperados, são tomados por base pelo sistema de recomendação para indicar o tipo fonológico de novas pronúncias a serem analisadas.

Para exemplificar, consideremos a estratégia de reparo *Redução do Encontro Consonantal*, onde a criança "apaga" uma das consoantes do encontro consonantal dentro da mesma sílaba. A produção da palavra bruxa como "buxa", da palavra cruz como "kuz", e da palavra fralda como "falda", caracteriza um emprego da estratégia Redução do Encontro Consonantal pela criança. Fatores que constituem uma estratégia de reparo podem ser recuperados toda vez que um caso de entrada apresentar características semelhantes. Assim, dada uma nova avaliação de fala onde a criança troca "bruxa" por "buxa", possivelmente seja um caso de Redução do Encontro Consonantal, com base nas características do caso recuperado.

Entretanto, um caso de entrada pode ser enquadrado em mais de uma estratégia de reparo: se a criança tem dificuldades em pronunciar um ou mais fonemas, quais estratégias de reparo podem ser recuperadas como possibilidades para o seu caso? Para responder a esta questão, cada palavra alvo foi categorizada de acordo com as estratégias de reparo possíveis, a fim de permitir a recuperação de casos prevista no Módulo de Compreensão. A

---

<sup>5</sup>Consoantes líquidas são caracterizadas pela obstrução da corrente expiratória feita com a ponta da língua, e também pela produção de sons vibrantes.

<sup>6</sup>Substituição de consoantes líquidas por outras consoantes líquidas.

<sup>7</sup>Substituição de uma consoante líquida por uma semivogal, por exemplo, "pular" pronunciada "pulai".

<sup>8</sup>Consoantes plosivas são produzidas pelo bloqueio da pressão do ar (através dos lábios unidos, por exemplo) que, a seguir, é desfeito, resultando numa "explosão".

<sup>9</sup>Produção das plosivas ou fricativas sonoras como surdas, por exemplo, "zebra" pronunciada como "sepa".

<sup>10</sup>Consoantes fricativas são produzidas por estreitamentos da passagem de ar (feitos entre os lábios e os dentes, por exemplo).

<sup>11</sup>Produção das consoantes plosivas ou fricativas surdas como sonoras, por exemplo, "tesoura" pronunciada como "desoura".

<sup>12</sup>Substituição de uma consoante palatal ou velar por uma alveolar ou labial, por exemplo, "cachorro" pronunciada como "casorro".

<sup>13</sup>Substituição de uma consoante labiodental, dental ou alveolar por uma palato-alveolar ou velar. Por exemplo, "sol" pronunciado como "xol".

<sup>14</sup>Substituição de uma consoante fricativa por uma plosiva, por exemplo, "vaca" pronunciada como "baca".

<sup>15</sup>Redução de um encontro consonantal dentro da mesma sílaba através do apagamento de uma das consoantes. Por exemplo: "cobra" pronunciada como "coba".

<sup>16</sup>Apagamento do "s" no final da sílaba ou final da palavra.

<sup>17</sup>Apagamento de consoante líquida no final da sílaba ou final de palavra.

<sup>18</sup>Apagamento de consoante líquida que ocorre entre duas vogais.

<sup>19</sup>Apagamento de consoante líquida em início de palavra.

<sup>20</sup>O fonema muda de lugar dentro da mesma palavra.

<sup>21</sup>Inserção de vogal entre duas consoantes.

Tabela 5.4 – Exemplo de categorização de palavras-alvo de acordo com estratégias de reparo

<b>Palavra-alvo</b>	<b>Estratégias de reparo possíveis</b>
Anel	Semivocalização de líquida, apagamento de líquida final
Batom	Dessonorização plosiva, sonorização plosiva, posteriorização plosiva
Bruxa	Substituição de líquida, semivocalização de líquida, desonorização plosiva, sonorização fricativa, anteriorização fricativa, posteriorização plosiva, plosivização fricativa e líquida, redução do EC, epêntese
Cachorro	Substituição de líquida, semivocalização de líquida, sonorização plosiva e fricativa, anteriorização plosiva e fricativa, plosivização fricativa e líquida, apagamento de líquida intervocálica
Chapéu	Sonorização plosiva e fricativa, anteriorização fricativa, posteriorização plosiva, plosivização fricativa
Cobra	Substituição de líquida, semivocalização de líquida, desonorização plosiva, sonorização plosiva, anteriorização plosiva, posteriorização plosiva, plosivização líquida, redução do EC, metátese, epêntese
Dente	Dessonorização plosiva, sonorização plosiva, posteriorização plosiva
Dragão	Semivocalização de líquida, desonorização plosiva, anteriorização plosiva, posteriorização plosiva, plosivização líquida, redução do EC, metátese, epêntese
Flor	Substituição de líquida, semivocalização de líquida, sonorização fricativa, posteriorização fricativa, plosivização fricativa e líquida, redução do EC, apagamento de líquida final, epêntese
Gato	Dessonorização plosiva, sonorização plosiva, anteriorização plosiva, posteriorização plosiva
Igreja	Substituição de líquida, semivocalização de líquida, desonorização plosiva e fricativa, anteriorização plosiva e fricativa, plosivização fricativa e líquida, redução do EC, metátese, epêntese
Lápis	Substituição de líquida, sonorização plosiva, sonorização fricativa, posteriorização plosiva e fricativa, plosivização fricativa e líquida, apagamento de fricativa inicial e final
Língua	Substituição de líquida, desonorização plosiva, anteriorização plosiva, plosivização líquida, apagamento de líquida inicial
Mesa	Dessonorização fricativa, posteriorização fricativa, plosivização fricativa

Fonte: Da autora.

Tabela 5.4 demonstra algumas palavras alvo categorizadas. Esta classificação (contendo todas as 84 palavras alvo) foi utilizada para criação de uma base de dados, onde podem ser recuperadas estratégias de acordo com um caso de entrada específico. Este processo é esclarecido no tópico seguinte.

### 5.3.3 Recomendação de Estratégias de Reparo

Conforme mencionado, uma criança, ao pronunciar palavras ou sentenças, pode empregar uma ou mais estratégia de reparo dependendo de como ela produz determinados fonemas. Por isso, no Módulo de Compreensão da arquitetura, estratégias de reparo podem ser recuperadas para explicar um caso de fala.

Seguindo o modelo de categorização exposto na Tabela 5.4, um banco de dados foi criado para armazenar todas as palavras alvo utilizadas em avaliações de fala, juntamente com as estratégias de reparo possíveis para cada palavra. A Figura 5.4 exemplifica uma das tabelas do banco de dados, contendo palavras alvo e a sequência de siglas que representam as estratégias de reparo. Armazenou-se o valor "1" para as estratégias que podem ocorrer na pronúncia da palavra, e "0" para as que não ocorrem.

Figura 5.4 – Tabela do banco de dados contendo palavras alvo e estratégias de reparo.

palavra	SL_L_L	SL_L_Y	SL_L_R	SL_L_RR	SL_OC_L	SL_OC_R	SV_L_L	SV_L_Y
Anel	0	0	0	0	0	0	1	0
Barriga	0	0	0	1	0	0	0	0
Batom	0	0	0	0	0	0	0	0
Bebê	0	0	0	0	0	0	0	0
Beijo	0	0	0	0	0	0	0	0
Biblioteca	1	0	0	0	1	0	0	0
Bicicleta	1	0	0	0	1	0	0	0
Bolsa	1	0	0	0	0	0	1	0
Brinco	0	0	1	0	0	1	0	0
Bruxa	0	0	1	0	0	1	0	0
Cabelo	1	0	0	0	0	0	1	0
Cachorro	0	0	0	1	0	0	0	0
Caixa	0	0	0	0	0	0	0	0
Calça	1	0	0	0	0	0	1	0
Cama	0	0	0	0	0	0	0	0

Fonte: *PhpMyAdmin* (Captura de tela) - da autora.

Considerando um caso de entrada em que a criança pronuncia incorretamente as palavras *anel*, *jornal*, *fralda* e *pastel*, devem-se recuperar casos (estratégias de reparo) que possam explicar a pronúncia, ordenando-os por maior probabilidade ou utilidade. Assim, na base de dados, executa-se uma consulta que retorne as estratégias desejadas:

```
select  estrategia, count(*) from palavra_estrategia, siglas
where  siglas.sigla = palavra_estrategia.id_estrategia and
palavra in ('Anel','Jornal','Fralda', 'Pastel')
group by  estrategia
order by 2 desc
limit 7;
```

A consulta acima irá retornar o resultado apresentado na Figura 5.5. De acordo com o exemplo, a pronúncia incorreta das palavras *anel*, *jornal*, *fralda* e *pastel* sugere que foi empregada quatro vezes a estratégia "Semivocalização de líquida" referente à consoante líquida "l", três vezes a estratégia "Apagamento de líquida final", e assim por diante, onde o número de ocorrências corresponde ao grau de probabilidade. As estratégias e probabilidades retornadas são recomendadas ao profissional de fonoaudiologia, que deve revisar as sugestões do sistema quanto ao caso de entrada. Esta revisão ocorre no módulo seguinte da arquitetura (Módulo de Assimilação), que deve alterar especificidades do caso, se necessário.

Figura 5.5 – Recuperação de estratégias de reparo com base em um caso de entrada.

estrategia	count(*)
Semivocalizao de liquida (l)	4
Apagamento de liquida final (l)	3
Plosivização Líquida (r)	2
Semivocalizao de liquida (r)	2
Substituicao de liquida (r)	2
Apagamento de liquida final (r)	1
Plosivização Fricativa (f)	1

Fonte: *PhpMyAdmin* (Captura de tela) - da autora.

Além de permitir a recomendação de estratégias de reparo ao especialista, a utilização da matriz de estratégias (tabela contendo palavras alvo e estratégias de reparo) é útil para realizar avaliações de fala personalizadas de acordo com as necessidades do paciente. Em um primeiro momento, por exemplo, a criança pode ser avaliada com todas as palavras alvo. Na ocorrência de erros de pronúncia, recupera-se as estratégias de reparo empregadas por ela. Em um segundo momento, o profissional pode personalizar a avaliação, escolhendo testar a pronúncia do paciente com relação a estratégias de re-

para específicas. Isto configura avaliações mais direcionadas e adaptáveis ao paciente, da mesma forma que dá mais opções ao terapeuta.

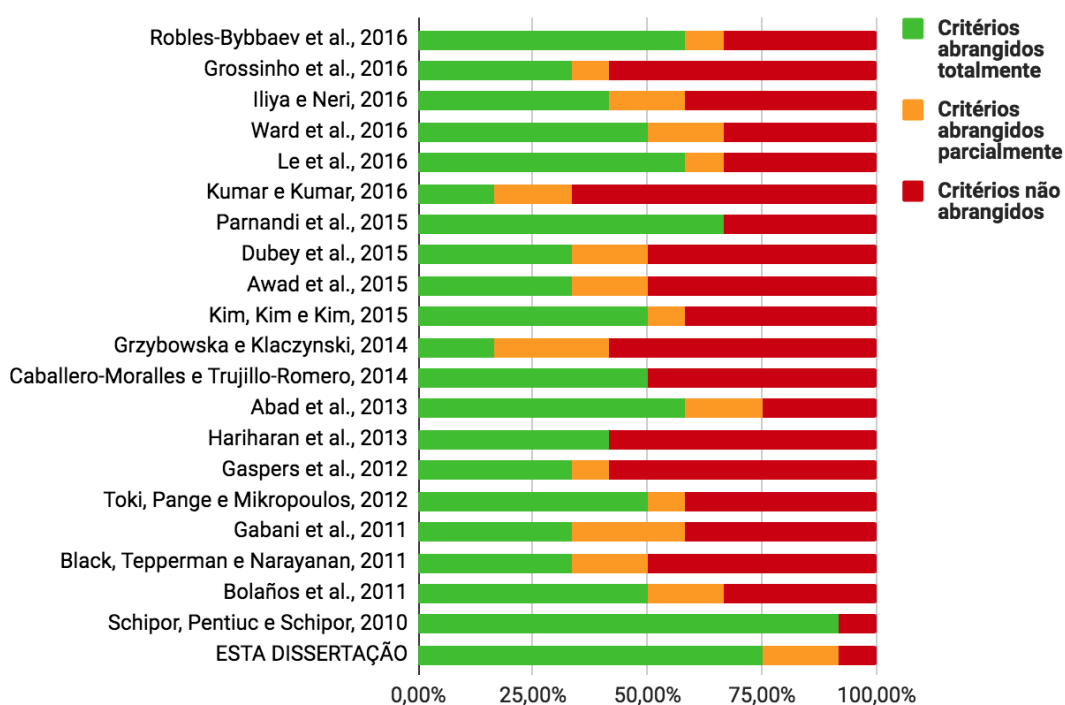
É importante mencionar, ainda, que a matriz de estratégias é altamente configurável, ou seja, pode ser modificada para incluir palavras em outros idiomas. Visto que a arquitetura apresenta-se de forma genérica, estes fatores contribuem para a aplicação da proposta em outros domínios de sistema.



## 6 COMPARAÇÃO COM TRABALHOS RELACIONADOS

Para verificar a utilidade da proposta, esta seção apresenta um comparativo entre os trabalhos relacionados abordados na Seção 3 e a presente dissertação. Tomando como base os critérios de Consciência de Situação propostos por (SALFINGER; RETSCHITZEGGER; SCHWINGER, 2013) e utilizados na revisão sistemática previamente mencionada, o gráfico 6.1 demonstra o grau de abrangência dos critérios nos trabalhos relacionados em comparação com esta dissertação, ao passo que a Tabela 6.1 mostra quais critérios são abordados no presente trabalho.

Gráfico 6.1 – Comparativo entre as propostas na área da fonoaudiologia.



Fonte: Da autora.

Considerando a abrangência de critérios de SA de uma forma geral, pode-se observar que a presente dissertação supera a maioria dos trabalhos relacionados, visto que abrange totalmente 75% dos critérios, incluindo *Modelo de Domínio*, *Suporte à Ação*, *Informação Contextual*, *Incompletude e Inconsistência*, *Adaptação de SA*, *Base de Conhecimento*, *Inteligência Humana*, *Personalização* e *Explicação e Exploração*. Acerca dos critérios abrangidos parcialmente, a proposta apresenta cobertura de 16,66% destes, que envolvem *Captura e Rastreamento de Evolução e Projeção*. Por fim, apenas um critério não encontra-se suportado nesta dissertação, o qual diz respeito ao *Ajuste de Sistema*.

Ao apresentar a arquitetura do sistema de recomendação, juntamente com os módulos componentes, esta dissertação abrange *Modelo de Domínio*, ou seja, uma repre-

Tabela 6.1 – Critérios de SA abrangidos por esta dissertação

Critérios de SA	
Dados de entrada	Heterogêneos
Modelo de Domínio	Sim
Avaliação de Situação	Decision Tree/queries em banco de dados
Suporte à Ação	Sim
Captura e Rastreamento de Evolução	Parcial
Projeção	Parcial
Informação Contextual	Sim
Incompletude e Inconsistência	Sim
Adaptação de SA	Sim
Ajuste de Sistema	Não
Base de Conhecimento	Sim
Inteligência Humana	Sim
Personalização	Sim
Explicação e Exploração	Sim

Fonte: Da autora.

sentação visual do conhecimento manipulado. Como *Suporte à Ação*, a proposta é capaz de sugerir ações de resposta ao fonoaudiólogo, a partir da identificação de estratégias de reparo utilizadas pela criança no momento da pronúncia e com o cálculo de PCC-R, que indica a gravidade da desordem. A *Informação Contextual* é abrangida porque, ao interpretar a situação de fala da criança, considera-se informação contextual sobre ela (como idade, região, escolaridade e gênero). O sistema de recomendação proposto lida com *Incompletude e Inconsistência*, à medida que os dados são pré-processados no Módulo de Classificação da arquitetura.

Ainda, ao mencionar critérios totalmente abrangidos no desenvolvimento desta dissertação, pode-se citar *Adaptação de SA*, a qual envolve adaptação dos métodos sem intervenção do usuário. Este fator evidencia-se na utilização de técnicas de aprendizagem de máquina (por exemplo, o classificador Decision Tree utilizado em amostras de fala no Módulo de Classificação), que adaptam-se de acordo com os modelos previamente aprendidos. A proposta incorpora uma *Base de Conhecimento*, que armazena conhecimento de domínio e dados empregados para interpretação de uma situação, sendo atualizada a cada processamento e, assim, adquirindo capacidades ao longo do tempo. *Incorporação de Inteligência Humana* também ocorre no sistema de recomendação, visto que o processamento computacional integra conhecimento especialista para funcionamento dos módulos da arquitetura. O critério de *Personalização* é alcançado, já que o sistema pode adaptar-se de acordo com as necessidades de cada indivíduo e, por fim, estão presentes *Explicação e Exploração*, nas quais o sistema é capaz de explicar o processamento conduzido em termos de dados ao operador, permitindo-o explorar estes dados.

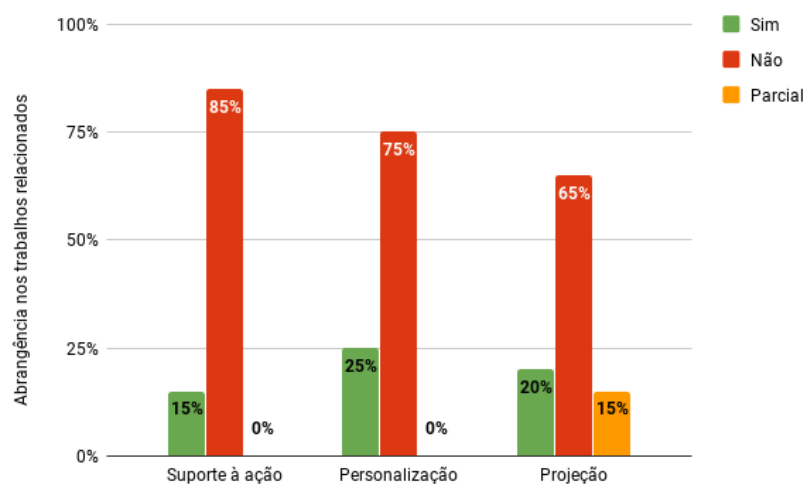
*Captura e Rastreamento de Evolução* e *Projeção* são dois critérios que esta dissertação cobre parcialmente. Com relação ao primeiro, o sistema de recomendação mostra o estado

atual do indivíduo analisado, porém, não apresenta situações em evolução e não analisa se esta evolução necessita de uma reação diferente. Assim, a cobertura do critério dá-se de forma parcial. Uma possível forma de abranger totalmente Captura e Rastreamento de Evolução seria realizar avaliações de fala periódicas com o paciente, armazenando-as e demonstrando variações de desempenho quando solicitado pelo terapeuta. Ao monitorar a situação do paciente, o sistema poderia, ainda, interpretar se há necessidade de adaptar seu comportamento para lidar com as variações observadas.

Com relação à *Projeção*, considera-se abrangência total quando ações são tomadas automaticamente no ambiente observado. Como esta dissertação não objetiva um sistema que faça diagnóstico e sim, que ofereça suporte ao profissional em suas decisões, o critério de *Projeção* é abrangido parcialmente, por meio da previsão de situações. Entretanto, a *Projeção* total poderia ser obtida gerando exercícios terapêuticos automaticamente de acordo com a situação identificada e apresentando-os diretamente ao paciente ou ao seu responsável (familiar ou fonoaudiólogo).

*Ajuste de Sistema* é um critério observado em poucas das propostas analisadas, assim como *Suporte à Ação*, *Personalização* e *Projeção*. O sistema de recomendação abordado nesta dissertação não abrange o critério de *Ajuste de Sistema*, pois não realiza ajuste de tempo de execução a fim de conservar SA mantendo recursos (performance e tempo de resposta, por exemplo). Em contrapartida, esta dissertação apresenta alguns diferenciais com relação aos trabalhos relacionados, em virtude da abrangência de critérios que comumente não são observados nas demais propostas, conforme demonstrado no Gráfico 6.2.

Gráfico 6.2 – Critérios pouco abrangidos por trabalhos relacionados na área da fonoaudiologia.



Fonte: Da autora.

Apenas 15% dos trabalhos revisados apresentam *Suporte à Ação*, o qual é totalmente alcançado pelo sistema de recomendação. No que se refere à *Personalização*,

apenas 25% das propostas apresentam esta característica, customizando seus recursos conforme as preferências ou necessidades do usuário. Nesta dissertação, a Personalização é efetivada pois a avaliação feita com a criança é direcionada de acordo com as características de sua fala. Para exemplificar, em um primeiro momento, avalia-se a pronúncia infantil utilizando todo o conjunto de 84 palavras alvo, selecionadas pelo especialista. Em um segundo momento, a avaliação é feita testando somente as estratégias de reparo que podem se enquadrar ao paciente, de acordo com os padrões de erro existentes em sua fala. Neste sentido, avaliações são realizadas de forma personalizada, visto que abrangem somente as estratégias de reparo necessárias e adequadas a um paciente específico.

O critério de Projeção é abrangido de forma parcial por esta dissertação, ao passo que 65% dos trabalhos revisados não o incluem. Isto deve-se ao fato de que o sistema proposto, conforme já mencionado, não tem o objetivo de tomar o lugar do profissional fonoaudiólogo, tomando decisões e gerando terapias de forma automatizada. Por outro lado, o sistema visa fornecer suporte à decisão clínica, mostrando situações de pacientes que possam orientar o especialista no seu trabalho cotidiano.

No geral, ao comparar esta dissertação com os trabalhos relacionados, os resultados obtidos demonstram que a presente proposta é eficaz na integração de consciência de situação, mostrando-se superior a 19 dos 20 trabalhos analisados. Cabe ressaltar, ainda, que as taxas alcançadas consideram o sistema apresentado neste estudo e, portanto, limitam-se aos módulos validados. Assim, a validação dos módulos restantes da arquitetura deve interferir positivamente na integração de consciência de situação na proposta, de forma a cobrir os demais critérios atualmente não abordados, ou abordados de maneira parcial.

A partir destas considerações, na próxima seção são apresentadas as conclusões da pesquisa, as contribuições trazidas para a área e as possibilidades a serem exploradas em trabalhos futuros.

## 7 CONSIDERAÇÕES FINAIS

No contexto da Computação Ubíqua e Pervasiva, a compreensão de situações representa um fator decisivo para a tomada de decisões em uma grande variedade de domínios, destacando-se a área da saúde, onde é necessário realizar o monitoramento do paciente. Na área da fonoaudiologia, especificamente, poucas propostas utilizam modelagem de conhecimento e interpretação de situações para melhorar as tarefas desempenhadas pelo terapeuta, e há pouca disponibilidade de profissionais para realizar o monitoramento adequado do paciente. Estas atividades podem ser melhoradas através de um sistema automatizado que, consciente da situação que está ocorrendo ao seu redor, é capaz de oferecer suporte ao fonoaudiólogo.

Por esta razão, esta dissertação apresentou uma arquitetura de sistema de recomendação voltado à área da terapia da fala, cujo objetivo principal é o auxílio na triagem de desordens dos sons da fala e no planejamento terapêutico. A arquitetura é composta de cinco módulos, que fazem uso conjunto de três teorias: consciência de situação, raciocínio baseado em casos e teoria do desenvolvimento cognitivo.

No Módulo de Percepção da arquitetura, foram coletados dados de fala de 1.362 crianças de 3 a 9 anos realizando tarefas de nomeação. No total, 133.143 avaliações fonoaudiológicas foram realizadas a partir de 84 palavras alvo. As avaliações e informações contextuais relacionadas foram armazenadas em banco de dados e utilizadas como dados de entrada para o processamento conduzido no Módulo de Classificação.

O Módulo de Classificação objetiva, principalmente, classificar a pronúncia da criança como correta ou incorreta. Para tal, os dados coletados no primeiro módulo da arquitetura foram pré-processados e transformados em espectrogramas, a fim de extrair características para alimentar o método de classificação. Os resultados obtidos neste módulo mostraram uma precisão de classificação acima de 92% para as pronúncias infantis, em concordância com as avaliações feitas pelo fonoaudiólogo.

Já no Módulo de Compreensão da arquitetura proposta, foram agrupados padrões de erro nas pronúncias infantis, de forma a verificar quais estratégias são empregadas pelas crianças para aproximarem-se da língua falada em seu meio. Estas estratégias (chamadas *estratégias de reparo* referem-se a trocas, omissões ou substituições de fonemas ocorrentes no momento da pronúncia. Foram identificadas quais estratégias de reparo são possíveis de ocorrer para cada palavra alvo, assim, dado um caso de entrada, estratégias de reparo são recuperadas do banco de dados como sugestão ao terapeuta.

De forma geral, com o sistema de recomendação proposto, foram atingidos bons níveis de classificação de fala, além de grande abrangência de critérios que caracterizam integração de consciência de situação, evidenciando diferenciais do presente trabalho com relação aos trabalhos relacionados. Com os resultados obtidos, confirma-se a hipótese

de que a união das teorias abordadas favorece a área da fonoaudiologia. Além disso, os resultados podem ser expandidos, à medida que forem testados e aperfeiçoados os módulos da arquitetura.

## 7.1 CONTRIBUIÇÕES DA PESQUISA

Como contribuições, esta dissertação apresenta, em primeiro lugar, uma revisão sistemática da literatura, trazendo trabalhos da área da fonoaudiologia que incorporam características de consciência de situação em suas metodologias. As análises apresentadas na revisão sistemática fornecem compreensão sobre os métodos mais utilizados no domínio e quais aspectos de SA permanecem um desafio na fonoaudiologia, fator que sustenta o propósito deste trabalho. Conhecer o estado da arte da consciência de situação na fonoaudiologia auxilia pesquisas futuras no desenvolvimento de propostas na área e dá orientação aos pesquisadores sobre aspectos que precisam ser explorados.

Como principal contribuição desta pesquisa, é apresentado um sistema de suporte à decisão clínica na fonoaudiologia, que visa ajudar o profissional de fonoaudiologia em tarefas envolvendo triagem de distúrbios e planejamento terapêutico. Tal suporte clínico é sustentado, principalmente, com a identificação de estratégias de reparo e classificação da fala, a qual é feita através da utilização de métodos de aprendizagem de máquina. Os métodos empregados na arquitetura aprendem características na fala infantil e classificam novas situações com base nos modelos aprendidos.

A classificação tomou como entrada características extraídas a partir de espectrogramas (representações visuais das frequências no áudio coletado). O uso de espectrogramas para extração de características representa uma nova abordagem na área da fonoaudiologia, visto que a literatura relacionada comumente utiliza outros métodos já consolidados para pré-processamento e reconhecimento de fala. A utilização de espectrogramas para distinção de padrões de pronúncia demonstrou resultados promissores, reafirmando-se, assim, outra contribuição da pesquisa.

Por fim, como contribuição final deste trabalho, construiu-se uma arquitetura que faz uso conjunto de três teorias para criar um sistema computacional: Raciocínio Baseado em Casos, Consciência de Situação e Teoria do Desenvolvimento Cognitivo. O uso de cada teoria não vem para substituir características de outra, mas sim para complementar seus recursos. Desta forma, o sistema mantém uma base de conhecimento constantemente atualizada, à medida que é capaz de identificar situações, oferecer suporte à decisão através de recomendações, e incluir recursos cognitivos ao seu processamento. O fato de unir três teorias para funcionamento do sistema proposto, faz com que a arquitetura seja genérica, ou seja, ela pode ser adaptada de forma multidisciplinar e oferecer contribuições não somente no domínio da Computação, mas também em outros domínios.

## 7.2 LIMITAÇÕES DA PESQUISA

A principal limitação referente a esta dissertação diz respeito à qualidade dos áudios coletados para análise e processamento. Visto que as avaliações de fala foram realizadas em ambiente pouco silencioso e sem utilização dos dispositivos adequados, os áudios frequentemente continham ruído de fundo e alguns destes eram praticamente inaudíveis. Este fator contribuiu para o atraso no processamento dos dados (que exigiram etapa de filtragem, corte e remoção de ruído), além de possivelmente afetar os resultados de classificação.

Outra limitação observada foi a utilização da API *Google Cloud Speech* nos dados coletados. Além de apresentar um desempenho pouco satisfatório ao reconhecer a pronúncia (provavelmente influenciado pela qualidade do som), a ferramenta foi testada somente com palavras em português brasileiro. Em caso de testes da API com outros idiomas, seria viabilizada a obtenção de indicativos sobre fatores que influenciam o reconhecimento de fala.

## 7.3 TRABALHOS FUTUROS

Como trabalho futuro desta dissertação objetiva-se, primeiramente, cobrir todos os critérios de Consciência de Situação apresentados nas seções anteriores. Foi demonstrado que, com esta pesquisa, 75% dos critérios de SA são abrangidos totalmente com os módulos componentes da arquitetura proposta. Assim, um dos trabalhos futuros diz respeito ao alcance dos critérios restantes, que compreendem Captura e Rastreamento de Situações em Evolução, Projeção e Ajuste de Sistema.

Na sequência, espera-se a validação de todos os módulos da arquitetura, ou seja, os Módulos de Assimilação e Acomodação, que visam adaptação inteligente e gerenciamento da base de conhecimento do sistema, respectivamente. Com a validação completa da arquitetura, é possível testar todos os aspectos da união de teorias proposta nesta dissertação e como cada teoria se comporta nas etapas de processamento distintas.

À medida que forem validados os módulos componentes da arquitetura, objetiva-se, ademais, testar o sistema de recomendação utilizando novos dados, atentando para a qualidade da coleta realizada. É possível que, com a aquisição de dados de fala mais robustos e consistentes, os resultados de classificação do sistema sofram alterações positivas, consequentemente melhorando o desempenho da proposta, de forma geral. Após coleta de dados na Unidade Móvel da arquitetura, um possível trabalho futuro diz respeito à identificação do usuário, por meio de uma assinatura de voz que permita facilitar o gerenciamento das avaliações feitas com um mesmo indivíduo e o monitoramento de sua fala.

No que se refere ao processo de classificação, é desejável que a pronúncia seja

classificada utilizando uma graduação, ou seja, níveis de indicação da presença de uma desordem da fala (por exemplo: porcentagem de classificações corretas abaixo de 50% representam alta indicação de desordem de fala no paciente). Esta graduação deve ser realizada com base em estratégias utilizadas na fonoaudiologia para categorizar a gravidade do caso avaliado.

Por fim, para validar os módulos restantes e incrementar os testes com o sistema de recomendação, objetivam-se experimentos com outros algoritmos de aprendizagem de máquina, a fim de avaliar qual é o mais adequado para o atual problema de pesquisa.

#### 7.4 PUBLICAÇÕES

A seguir, apresentam-se os artigos produzidos, referentes a esta dissertação:

1. Franciscatto, M. H., Lima, J. C. D., Moro, A., Maran, V., Augustin, I., Soares, M. K. e da Rocha, C. C. (2018) *A Case-Based System Architecture based on Situation-Awareness for Speech Therapy*. 20th International Conference on Enterprise Information Systems (ICEIS), páginas 461-468. (PUBLICADO)
  - Qualis: B2.
2. Franciscatto, M. H., Augustin, I., Lima, J. C. D., Maran, V. *Situation Awareness in the Speech Therapy Domain: A Systematic Mapping Study*. Computer Speech & Language - Special Issue: Speech and Language Processing for Behavioral and Mental Health Applications. (ACEITO PARA PUBLICAÇÃO)
  - Qualis: A2.
3. Franciscatto, M. H., Trois, C., Lima, J. C. D., Augustin, I. *Blending Situation Awareness with Machine Learning to Identify Children's Speech Disorders*. 12th International Conference on Application of Information and Communication Technologies (AICT). (SUBMETIDO)
  - Qualis: B2.



## REFERÊNCIAS BIBLIOGRÁFICAS

AAMODT, A.; PLAZA, E. Case-based reasoning: Foundational issues, methodological variations, and system approaches. **AI communications**, IOS press, v. 7, n. 1, p. 39–59, 1994.

ABAD, A. et al. Automatic word naming recognition for an on-line aphasia treatment system. **Computer Speech & Language**, Elsevier, v. 27, n. 6, p. 1235–1248, 2013.

AGGARWAL, C. C. An introduction to recommender systems. In: **Recommender Systems**. [S.l.]: Springer, 2016. p. 1–28.

AHMED, M. U.; BEGUM, S.; FUNK, P. Case studies on the clinical applications using case-based reasoning. In: IEEE. **Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on**. [S.l.], 2012. p. 3–10.

AIELLO, V. et al. Next-generation technologies for preventing accidental death of children trapped in parked vehicles. In: IEEE. **Information Reuse and Integration (IRI), 2014 IEEE 15th International Conference on**. [S.l.], 2014. p. 508–513.

ARAÚJO, R. B. de. Computação ubíqua: Princípios, tecnologias e desafios. In: **XXI Simpósio Brasileiro de Redes de Computadores**. [S.l.: s.n.], 2003. v. 8, p. 11–13.

ARI, N.; USTAZHANOV, M. Matplotlib in python. In: IEEE. **11th International Conference on Electronics, Computer and Computation (ICECCO)**. [S.l.], 2014. p. 1–6.

AWAD, S. S. et al. Android-based real-time signal processing to treat speech-language pathologies. In: IEEE. **Computer Engineering Conference (ICENCO), 2015 11th International**. [S.l.], 2015. p. 68–75.

BAESSO, J. S. **O uso de estratégias de reparo nos constituintes coda e onset complexo por crianças com aquisição fonológica normal e desviante**. 2009. 155 p. Dissertação (Mestrado em Distúrbios da Comunicação Humana) — Universidade Federal de Santa Maria, Santa Maria, 2009.

BALLINGER, B. et al. On-demand language model interpolation for mobile speech input. In: **Eleventh Annual Conference of the International Speech Communication Association**. [S.l.: s.n.], 2010.

BETTINI, C. et al. A survey of context modelling and reasoning techniques. **Pervasive and Mobile Computing**, Elsevier, v. 6, n. 2, p. 161–180, 2010.

BISWAS, S. K. et al. A hybrid cbr classification model by integrating ann into cbr. **International Journal of Services Technology and Management**, Inderscience Publishers (IEL), v. 21, n. 4-6, p. 272–293, 2015.

BLACK, M. P.; TEPPERMAN, J.; NARAYANAN, S. S. Automatic prediction of children's reading ability for high-level literacy assessment. **IEEE Transactions on Audio, Speech, and Language Processing**, IEEE, v. 19, n. 4, p. 1015–1028, 2011.

BOLAÑOS, D. et al. Flora: Fluent oral reading assessment of children's speech. **ACM Transactions on Speech and Language Processing (TSLP)**, ACM, v. 7, n. 4, p. 16, 2011.

CABALLERO-MORALES, S.-O.; TRUJILLO-ROMERO, F. Evolutionary approach for integration of multiple pronunciation patterns for enhancement of dysarthric speech recognition. **Expert Systems with Applications**, Elsevier, v. 41, n. 3, p. 841–852, 2014.

CAMERON, M. A. et al. Emergency situation awareness from twitter for crisis management. In: ACM. **Proceedings of the 21st International Conference on World Wide Web**. [S.l.], 2012. p. 695–698.

CHEU, E.-Y.; NG, S.-K.; QUEK, H.-C. An interval type-2 neural fuzzy inference system based on piaget's action-cognitive paradigm. In: IEEE. **Evolutionary Computation, 2009. CEC'09. IEEE Congress on**. [S.l.], 2009. p. 925–932.

CHUCHUCA-MÉNDEZ, F. et al. An educative environment based on ontologies and e-learning for training on design of speech-language therapy plans for children with disabilities and communication disorders. In: **Ciencias de la Informática y Desarrollos de Investigación (CACIDI), IEEE Congreso Argentino de**. [S.l.: s.n.], 2016. p. 1–6.

CULMONE, R. et al. Aal domain ontology for event-based human activity recognition. In: IEEE. **Mechatronic and Embedded Systems and Applications (MESA), 2014 IEEE/ASME 10th International Conference on**. [S.l.], 2014. p. 1–6.

DEY, A. K. Understanding and using context. **Personal and ubiquitous computing**, Springer-Verlag, v. 5, n. 1, p. 4–7, 2001.

DIAS, F. O desenvolvimento cognitivo no processo de aquisição de linguagem. **Letrônica**, v. 3, n. 2, p. 107–119, 2010.

DUBEY, H. et al. Echowear: smartwatch technology for voice and speech treatments of patients with parkinson's disease. In: **Proceedings of the conference on Wireless Health**. [S.l.: s.n.], 2015.

ENDSLEY, M. R. Toward a theory of situation awareness in dynamic systems. **Human factors**, SAGE Publications Sage CA: Los Angeles, CA, v. 37, n. 1, p. 32–64, 1995.

ENDSLEY, M. R.; GARLAND, D. Theoretical underpinnings of situation awareness: A critical review. **Situation awareness analysis and measurement**, Lawrence Erlbaum Associates, Mahwah, NJ, p. 3–32, 2000.

FARIAS, G. P. et al. A bdi-fuzzy agent model for exchanges of non-economic services based on the social exchange theory. In: IEEE. **Intelligent Systems (BRACIS), 2013 Brazilian Conference on**. [S.l.], 2013. p. 26–32.

FENG, Y.-H.; TENG, T.-H.; TAN, A.-H. Modelling situation awareness for context-aware decision support. **Expert Systems with Applications**, Elsevier, v. 36, n. 1, p. 455–463, 2009.

FILHO, J. D. P. R. et al. Mhars: A mobile system for human activity recognition and inference of health situations in ambient assisted living. **Journal of Applied Computing Research**, v. 5, n. 1, p. 44–58, 2016.

FROST, M.; GABRIELLI, S. Supporting situational awareness through a patient overview screen for bipolar disorder treatment. In: **Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on**. [S.l.: s.n.], 2013. p. 298–301.

GABANI, K. et al. Exploring a corpus-based approach for detecting language impairment in monolingual english-speaking children. **Artificial Intelligence in Medicine**, Elsevier, v. 53, n. 3, p. 161–170, 2011.

GARCIA, S. M. D. S. A construção do conhecimento segundo jean piaget. **Ensino em Re-vista**, 1997.

GASPERS, J. et al. An evaluation of measures to dissociate language and communication disorders from healthy controls using machine learning techniques. In: **ACM. Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium**. [S.l.], 2012. p. 209–218.

GELDERBLOM, H.; KOTZÉ, P. Designing technology for young children: what we can learn from theories of cognitive development. In: **ACM. Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology**. [S.l.], 2008. p. 66–75.

GILSON, R. D. Situation awareness - special issue preface. **Human factors**, v. 37, n. 1, 1995.

GOTO, J. et al. Activity recognition system for watching over infant children. In: **IEEE. ROMAN, 2013 IEEE**. [S.l.], 2013. p. 473–477.

GROSSINHO, A. et al. Robust phoneme recognition for a speech therapy environment. In: **Serious Games and Applications for Health (SeGAH), 2016 IEEE International Conference on**. [S.l.: s.n.], 2016. p. 1–7.

GRZYBOWSKA, J.; KLACZYNSKI, M. Computer-assisted hfcc-based learning system for people with speech sound disorders. In: **Pacific Voice Conference (PVC), 2014 XXII Annual**. [S.l.: s.n.], 2014. p. 1–5.

GUILLÉN, R.; USREY, R. Case based reasoning using speech data for clinical assessment. **Innovations in Applied Artificial Intelligence**, Springer, p. 323–354, 2005.

HARIHARAN, M. et al. A hybrid expert system approach for telemonitoring of vocal fold pathology. **Applied Soft Computing**, Elsevier, v. 13, n. 10, p. 4148–4161, 2013.

HASEGAWA-JOHNSON, M. et al. Hmm-based and svm-based recognition of the speech of talkers with spastic dysarthria. In: **IEEE. Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on**. [S.l.], 2006. v. 3, p. III–III.

HUSAIN, W.; PHENG, L. T. The development of personalized wellness therapy recommender system using hybrid case-based reasoning. In: **IEEE. Computer Technology and Development (ICCTD), 2010 2nd International Conference on**. [S.l.], 2010. p. 85–89.

ILIYA, S.; NERI, F. Towards artificial speech therapy: A neural system for impaired speech segmentation. **International journal of neural systems**, World Scientific, v. 26, n. 6, p. 1–16, 2016.

ISINKAYE, F.; FOLAJIMI, Y.; OJOKOH, B. Recommendation systems: Principles, methods and evaluation. **Egyptian Informatics Journal**, Elsevier, v. 16, n. 3, p. 261–273, 2015.

KIM, M. J.; KIM, Y.; KIM, H. Automatic intelligibility assessment of dysarthric speech using phonologically-structured sparse linear model. **IEEE Transactions on Audio, Speech, and Language Processing**, IEEE Press, v. 23, n. 4, p. 694–704, 2015.

KIRSTE, T. A reference model for situation-aware assistance. In: **BAD HONNEFF, GERMANY. Proceedings of the Mensch und Computer**. [S.l.], 2001.

KOKAR, M. M.; ENDSLEY, M. R. Situation awareness and cognitive modeling. **IEEE Intelligent Systems**, IEEE, v. 27, n. 3, p. 91–96, 2012.

KUMAR, S. A.; KUMAR, C. S. Improving the intelligibility of dysarthric speech towards enhancing the effectiveness of speech therapy. In: IEEE. **Advances in Computing, Communications and Informatics (ICACCI), 2016 International Conference on**. [S.l.], 2016. p. 1000–1005.

KUO, J. Y. et al. Multiagent cooperative learning strategies for pursuit-evasion games. **Mathematical Problems in Engineering**, Hindawi Publishing Corporation, v. 2015, 2015.

KURMIS, M. et al. Development of the real time situation identification model for adaptive service support in vehicular communication networks domain. **Advances in Electrical and Electronic Engineering**, Faculty of Electrical Engineering and Computer Science VSB-Technical University of Ostrava, v. 11, n. 5, p. 342, 2013.

LE, D. et al. Automatic assessment of speech intelligibility for individuals with aphasia. **IEEE/ACM Transactions on Audio, Speech, and Language Processing**, IEEE, v. 24, n. 11, p. 2187–2199, 2016.

LEE, H. J.; KIM, H. S. ehealth recommendation service system using ontology and case-based reasoning. In: IEEE. **Smart City/SocialCom/SustainCom (SmartCity), 2015 IEEE International Conference on**. [S.l.], 2015. p. 1108–1113.

LI, X.; WU, X. Decision tree based state tying for speech recognition using dnn derived embeddings. In: IEEE. **Chinese Spoken Language Processing (ISCSLP), 2014 9th International Symposium on**. [S.l.], 2014. p. 123–127.

LIU, W.; LI, X.; HUANG, D. A survey on context awareness. In: IEEE. **Computer Science and Service System (CSSS), 2011 International Conference on**. [S.l.], 2011. p. 144–147.

LOPES, J. et al. A situation-aware pervasive approach for assessing therapeutic goals in healthcare environment. In: ACM. **Proceedings of the 31st Annual ACM Symposium on Applied Computing**. [S.l.], 2016. p. 125–130.

LYYTINEN, K.; YOO, Y. Ubiquitous computing. **Communications of the ACM**, v. 45, n. 12, p. 63–96, 2002.

MARTÍN-RUIZ, M. L.; DUBOY, M. A. V.; CRUZ, I. Pau de la. Deployment and validation of a smart system for screening of language disorders in primary care. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 13, n. 6, p. 7522–7545, 2013.

MATHEW, L. R.; ANSELAM, A. S.; PILLAI, S. S. Analysis of Id-celp coder output with sound exchange and praat software. In: IEEE. **Advanced Communication Control and Computing Technologies (ICACCCT), 2014 International Conference on**. [S.l.], 2014. p. 1281–1285.

MEGALINGAM, R. K. et al. Medisuit: Wearable health monitoring system for elders and bed-ridden patients. In: **Intelligent Systems and Control (ISCO), 2016 10th International Conference on**. [S.l.: s.n.], 2016. p. 1–6.

MOHAMED, S. A. E.; HASSANIN, A. S.; OTHMAN, M. T. B. Educational system for the holy quran and its sciences for blind and handicapped people based on google speech api. **Journal of Software Engineering and Applications**, Scientific Research Publishing, v. 7, n. 03, p. 150, 2014.

NETO, A. D. F. **MODELO PARA PREDIÇÃO DE AÇÕES E INFERÊNCIA DE SITUAÇÕES DE RISCO EM AMBIENTES SENSÍVEIS AO CONTEXTO**. 2015. 74 p. Dissertação (Mestrado em Computação) — Universidade Federal de Santa Maria, Santa Maria, 2015.

NITHYA, B.; ILANGO, V. Predictive analytics in health care using machine learning tools and techniques. In: IEEE. **Intelligent Computing and Control Systems (ICICCS), 2017 International Conference on**. [S.l.], 2017. p. 492–499.

NWIABU, N. et al. Situation awareness in context-aware case-based decision support. In: IEEE. **Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2011 IEEE First International Multi-Disciplinary Conference on**. [S.l.], 2011. p. 9–16.

\_\_\_\_\_. Case-based situation awareness. In: IEEE. **Cognitive Methods in Situation Awareness and Decision Support (CogSIMA), 2012 IEEE International Multi-Disciplinary Conference on**. [S.l.], 2012. p. 22–29.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 24, n. 7, p. 971–987, 2002.

OOSTHUIZEN, R.; PRETORIUS, L. System dynamics modelling of situation awareness. In: **Military Communications and Information Systems Conference (MilCIS), 2015**. [S.l.: s.n.], 2015. p. 1–6.

PARNANDI, A. et al. Development of a remote therapy tool for childhood apraxia of speech. **ACM Transactions on Accessible Computing (TACCESS)**, ACM, v. 7, n. 3, p. 10, 2015.

PERERA, C. et al. Context aware computing for the internet of things: A survey. **IEEE Communications Surveys & Tutorials**, IEEE, v. 16, n. 1, p. 414–454, 2014.

PIAGET, J. **A Construção do Real na Criança**. Rio de Janeiro: Zahar: Trad. Álvaro Cabral, 1970.

\_\_\_\_\_. **Epistemologia Genética**. [S.l.]: Petrópolis: Vozes, 1970.

\_\_\_\_\_. **O Nascimento da Inteligência na Criança**. Rio de Janeiro: Zahar: Trad. Álvaro Cabral, 1975.

PRASOMPHAN, S. Detecting human emotion via speech recognition by using speech spectrogram. In: IEEE. **Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on**. [S.l.], 2015. p. 1–10.

RADOVIĆ, D. et al. Internet radio player implementation using ffmpeg software support. In: **Smart Systems and Technologies (SST), 2017 International Conference on**. [S.l.: s.n.], 2017. p. 259–262.

RAMAN, V. et al. Performance based cbr mass detection in mammograms. In: IEEE. **Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference on**. [S.l.], 2010. p. 565–568.

RASHID, A. M. et al. Getting to know you: learning new user preferences in recommender systems. In: ACM. **Proceedings of the 7th international conference on Intelligent user interfaces**. [S.l.], 2002. p. 127–134.

RIESBECK, C. K.; SCHANK, R. C. **Inside case-based reasoning**. [S.l.]: Erlbaum, Northvale, NJ, 1989.

ROBLES-BYKBAEV, V. E. et al. An ontology-based expert system to generate therapy plans for children with disabilities and communication disorders. In: **Ecuador Technical Chapters Meeting (ETCM), IEEE**. [S.l.: s.n.], 2016. v. 1, p. 6.

SALFINGER, A.; RETSCHITZEGGER, W.; SCHWINGER, W. Maintaining situation awareness over time. p. 1–8, 2013.

SCHEUNEMANN, D. A. **Ciência de situação na IOT: uma arquitetura explorando processamento híbrido de contexto**. 2016. 81 p. Dissertação (Mestrado em Engenharia Eletrônica e Computação) — Universidade Católica de Pelotas, Pelotas, 2016.

SCHILIT, B.; ADAMS, N.; WANT, R. Context-aware computing applications. In: IEEE. **Mobile Computing Systems and Applications, 1994. WMCSA 1994. First Workshop on**. [S.l.], 1994. p. 85–90.

SCHILIT, B. N.; THEIMER, M. M. Disseminating active map information to mobile hosts. **IEEE network**, IEEE, v. 8, n. 5, p. 22–32, 1994.

SCHIPOR, O. A.; PENTIUC, S. G.; SCHIPOR, M. D. Improving computer based speech therapy using a fuzzy expert system. **Computing and Informatics**, v. 29, n. 2, p. 303–318, 2010.

SEZGIN, E.; OZKAN, S. A systematic literature review on health recommender systems. In: IEEE. **E-Health and Bioengineering Conference (EHB), 2013**. [S.l.], 2013. p. 1–4.

SIMATWA, E. M. Piaget's theory of intellectual development and its implication for instructional management at pre-secondary school level. **Educational Research and Reviews**, Academic Journals, v. 5, n. 7, p. 366, 2010.

SIVAPALAN, S. et al. Recommender systems in e-commerce. In: IEEE. **World Automation Congress (WAC), 2014**. [S.l.], 2014. p. 179–184.

STANTON, N. A.; CHAMBERS, P. R.; PIGGOTT, J. Situational awareness and safety. **Safety science**, Elsevier, v. 39, n. 3, p. 189–204, 2001.

STAPLES, M.; NIAZI, M. Experiences using systematic review guidelines. **Journal of Systems and Software**, Elsevier, v. 80, n. 9, p. 1425–1437, 2007.

SWAIN, P. H.; HAUSKA, H. The decision tree classifier: Design and potential. **IEEE Transactions on Geoscience Electronics**, IEEE, v. 15, n. 3, p. 142–147, 1977.

TOKI, E. I.; PANGE, J.; MIKROPOULOS, T. A. An online expert system for diagnostic assessment procedures on young children's oral speech and language. **Procedia Computer Science**, Elsevier, v. 14, p. 428–437, 2012.

TROIS, C. et al. Exploring textures in traffic matrices to classify data center communications. In: **32nd International Conference on Advanced Information Networking and Applications (AINA)**. [S.l.: s.n.], 2018.

TSENG, J. C. et al. An interactive healthcare system with personalized diet and exercise guideline recommendation. In: IEEE. **Technologies and Applications of Artificial Intelligence (TAAI), 2015 Conference on**. [S.l.], 2015. p. 525–532.

WANGENHEIM, A. V.; WANGENHEIM, C. G. V.; RATEKE, T. **Raciocínio Baseado em Casos**. 2ª ed. Revisada e Atualizada: Bookess, 2013.

WARD, L. et al. Automated screening of speech development issues in children by identifying phonological error patterns. In: **INTER\_SPEECH**. [S.l.: s.n.], 2016. p. 2661–2665.

WATSON, I. Case-based reasoning is a methodology not a technology. **Knowledge-based systems**, Elsevier, v. 12, n. 5, p. 303–308, 1999.

WEISER, M. Some computer science issues in ubiquitous computing. **Communications of the ACM**, ACM, v. 36, n. 7, p. 75–84, 1993.

\_\_\_\_\_. The computer for the 21st century. **Mobile Computing and Communications Review**, v. 3, n. 3, p. 3–11, 1999.

WIETHAN, F. M.; MOTA, H. B. Emprego de estratégias de reparo para os fonemas fricativos no desvio fonológico. **Revista da Sociedade Brasileira de Fonoaudiologia**, SciELO Brasil, v. 17, n. 1, 2012.

WOERN DL, W. et al. A model for proactivity in mobile, context-aware recommender systems. In: **Proceedings of the fifth ACM conference on Recommender systems**. [S.l.: s.n.], 2011. p. 273–276.

YAU, S. S. et al. Situation-aware contract specification language for middleware for ubiquitous computing. In: IEEE. **Distributed Computing Systems, 2003. FTDCS 2003. Proceedings. The Ninth IEEE Workshop on Future Trends of**. [S.l.], 2003. p. 93–99.

YAU, S. S.; WANG, Y.; KARIM, F. Development of situation-aware application software for ubiquitous computing environments. In: IEEE. **Computer Software and Applications Conference, 2002. COMPSAC 2002. Proceedings. 26th Annual International**. [S.l.], 2002. p. 233–238.

YAVAS, M.; HERNANDORENA, C.; LAMPRECHT, R. **Avaliação Fonológica da Criança (Phonological Assessment of Child Speech)**. [S.l.]: Porto Alegre, Brazil: Artes Médicas, 1991.

YE, J.; DOBSON, S.; MCKEEVER, S. Situation identification techniques in pervasive computing: A review. **Pervasive and mobile computing**, Elsevier, v. 8, n. 1, p. 36–66, 2012.

YORDANOVA, K. et al. Towards a situation model for assessing challenging behaviour of people with dementia. In: **Proceedings of the 3rd International Workshop on Sensor-based Activity Recognition and Interaction**. [S.l.: s.n.], 2016. p. 7.

YÜNCÜ, E.; HACIHABIBOGLU, H.; BOZSAHIN, C. Automatic speech emotion recognition using auditory models with binary decision tree and svm. In: IEEE. **Pattern Recognition (ICPR), 2014 22nd International Conference on**. [S.l.], 2014. p. 773–778.

ZANKER, M. et al. Comparing recommendation strategies in a commercial context. **IEEE Intelligent Systems**, IEEE, v. 22, n. 3, 2007.

ZHANG, Y. et al. A multi-disciplinary medical treatment decision support system with intelligent treatment recommendation. In: IEEE. **Computer and Communications (ICCC), 2016 2nd IEEE International Conference on**. [S.l.], 2016. p. 838–842.