

UNIVERSIDADE FEDERAL DO RIO GRANDE - FURG  
CENTRO DE CIÊNCIAS COMPUTACIONAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

MÁRCIO JOSUÉ RAMOS TORRES

**Interfaces Auditivas para Elementos de  
Workspace Awareness Acessíveis aos Cegos  
em Ferramentas de Modelagem  
Colaborativa de Diagramas**

Dissertação apresentada como requisito parcial  
para a obtenção do grau de Mestre em Engenharia  
da Computação

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Regina Bärwaldt

Rio Grande  
2020

## Ficha Catalográfica

T693i Torres, Márcio Josué Ramos.  
Interfaces auditivas para elementos de *Workspace Awareness* acessíveis aos cegos em ferramentas de modelagem colaborativa de diagramas / Márcio Josué Ramos Torres. – 2020.  
103 f.

Dissertação (mestrado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Computação, Rio Grande/RS, 2020.  
Orientadora: Dra. Regina Bärwaldt.

1. Cegos 2. Acessibilidade 3. Diagramas 4. *Workspace Awareness*  
5. Interfaces Auditivas I. Bärwaldt, Regina II. Título.

CDU 004.5:37-056.262

Catálogo na Fonte: Bibliotecário José Paulo dos Santos CRB 10/2344



MINISTÉRIO DA EDUCAÇÃO  
UNIVERSIDADE FEDERAL DO RIO GRANDE  
CENTRO DE CIÊNCIAS COMPUTACIONAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO  
CURSO DE MESTRADO EM ENGENHARIA DE COMPUTAÇÃO



## DISSERTAÇÃO DE MESTRADO

### **Interfaces Auditivas para Elementos de Workspace Awareness Acessíveis aos Cegos em Ferramentas de Modelagem Colaborativa de Diagramas**

Márcio Josué Ramos Torres

Banca examinadora:

Prof. Dr. José Antonio dos Santos Borges

Prof. Dr. Ricardo Nagel Rodrigues

Prof<sup>a</sup>. Dr<sup>a</sup>. Héli da Salles Santos

Prof<sup>a</sup>. Dr<sup>a</sup>. Regina Bärwaldt  
Orientadora

*“Quando a educação não é libertadora,  
o sonho do oprimido é ser o opressor.”*

— PAULO FREIRE

## AGRADECIMENTOS

Obrigado à minha família, minha mãe que me criou e educou, filha por me lembrar o que é ver o lado bom da vida, e amanceba, por aturar a ansiedade, mau-humor, tensão e todos os estágios que passei.

Obrigado IFRS e serviço público federal por proporcionar esse “empréstimo” de tempo para capacitação, fiz o melhor uso possível do meu afastamento. Também sou grato à educação pública, gratuita e de qualidade oferecida pela FURG. Muito respeito e reverência por essas duas instituições.

Aos meus colegas no IFRS, obrigado pelo incentivo e pela compreensão, pois deixei a equipe desfalcada durante meu afastamento. Obrigado Raquel por herdar a coordenação do curso e tocar o projeto de pesquisa de Tecnologia Assistiva, que serviu de inspiração para esse mestrado. Obrigado Pinta pelas conversas que direcionaram o tema e por colaborar fornecendo sua “voz” para o protótipo.

Obrigado aos professores Adriano Werhli e Paulo Butzen, aprendi muito com ambos. Agradeço à minha orientadora, Regina Bärwaldt, pelas ideias, motivação e acolhimento, nós não concordamos com todos os meios e fins, mas *“qual vai ser teu legado cara?”*

Meu muito obrigado à todos colegas do grupo Informática na Educação, que foram sempre muito solícitos e parceiros nessa empreitada. Especialmente os colegas Douglas, pela revisão do meu texto para qualificação, Tiago, pela revisão do artigo para o FIE, Stefane, por colaborar como modelo de “voz”. Um gratidão especial aos amigos Pinho e Topin, pela parceria, pelas risadas, utilidade e inutilidades no TPM, pelas revisões, dicas, etc, e Pinho, meu amigão, estou certo que te recuperas para tomarmos um chimarrão juntos de novo.

Minha gratidão à Associação Escola Louis Braille de Pelotas por ceder um espaço para o teste com usuários. Fui atendido com muito carinho e atenção por eles. Meu obrigado, também, ao IFSul Pelotas pela solicitude e reserva do espaço.

Finalmente, mas não menos importante, agradeço ao meu pai por anos atrás ter dito *“tu não vai sujar as mãos de graxa”*.

## RESUMO

Existem obstáculos para o aprendizado dos cegos e deficientes visuais, sobretudo nos cursos da área da Tecnologia da Informação e Comunicação, que usam informações diagramáticas ostensivamente em sala de aula, materiais didáticos e ferramentas. Há pesquisas sobre a acessibilidade para a percepção e interação em diagramas, no entanto, os recursos colaborativos acessíveis são menos evidenciados. Esses recursos são importantes, pois o trabalho em grupo apoia a integração social, a troca de saberes e a aprendizagem colaborativa. As funcionalidades básicas, necessárias às aplicações de *groupware* para colaboração, são apresentadas no *framework* de Gutwin e Greenberg, categorizadas como elementos de *workspace awareness*. Entretanto, o *framework* é uma abstração e não providencia detalhes, como a acessibilidade. Assim, a importância educacional e a lacuna na pesquisa sobre *workspace awareness* (WA) acessível motivaram esse estudo, cujo objetivo é projetar uma interface auditiva por meio da sonificação dos elementos de WA de modo que os usuários cegos possam perceber e acompanhar as mudanças no *workspace* em ferramentas de modelagem colaborativa de diagramas. O projeto apoiou-se na metodologia iterativa *Design Science Research* e resultou na divisão deste estudo em duas etapas. A primeira foi dedicada ao projeto da primeira interface auditiva e implementação em um protótipo de editor diagramas. A segunda etapa foi destinada ao reprojeto da interface auditiva, inclusão de duas interfaces baseadas em fala, sintetizada e natural, e atualização do protótipo. As interfaces foram avaliadas por meio de testes com usuários, com os métodos de medição da taxa de sucesso por tarefa e da carga cognitiva subjetiva por meio da ferramenta NASA-TLX. Os resultados obtidos foram satisfatórios na percepção das principais categorias de *awareness* “*quem? fez o quê? e onde?*”, com o desempenho das interfaces auditivas muito próximo das baseadas em fala. Ademais, a carga de trabalho reportada foi menor nas interfaces *sonificadas* que nas *faladas*. Concluiu-se que as interfaces auditivas são uma opção viável para a implementação de um *workspace awareness* acessível às pessoas cegas e que ainda existem questões nessa área para abordar em trabalhos futuros.

**Palavras-chave:** Cegos. Acessibilidade. Diagramas. Workspace Awareness. Interfaces Auditivas.

## **Auditory Interfaces for Workspace Awareness Elements Accessible to the Blind in Diagrams' Collaborative Modeling Tools**

### **ABSTRACT**

There are obstacles for the blind and visually impaired regarding learning issues, especially in the Information and Communication Technology field, where diagrammatic information is ostensibly employed in the classroom, teaching materials, and tools. There are researches on accessibility for perception and interaction to diagrams, however, collaborative resources are less available and evident in literature. These resources are important, as group activities support social integration, exchange of knowledge and collaborative learning. The essential functionalities for collaboration in groupware applications are presented in the Gutwin and Greenberg's framework and categorized as workspace awareness (WA) elements. However, the framework is an abstraction and does not provide details about the accessibility issues. Thus, the educational importance and the gap in the research of accessible awareness motivated this study, whose objective is to design an auditory interface through the sonification of WA elements so that blind users could perceive and follow workspace updates in collaborative diagram modeling tools. This project followed an iterative methodology known as Design Science Research, which resulted in the separation of this study in two stages. The first one was dedicated to the design of the initial auditory interface and the implementation of a diagram modeling tool. The second stage led to redesigning the auditory interface, and including two new speech-based interfaces, one synthesized and other natural, plus the prototype updates. The designed interfaces were evaluated in user testing sessions through the methods of success rate per task and cognitive load measurement using the NASA-TLX tool. The results obtained were satisfactory for the perception of the essential awareness categories "who? is doing what? and where?". The performance of the auditory interfaces was close to the speech-based ones. Besides, auditory interfaces reported a smoother workload compared to the speech-based ones. It was concluded that auditory interfaces are a viable option for the implementation of accessible workspace awareness elements that are suitable to the blind, and that there are still problems in this area to be addressed in future works.

**Keywords:** blind people, accessibility, diagram, workspace awareness, auditory interfaces.

## LISTA DE ABREVIATURAS E SIGLAS

|        |  |
|--------|--|
| AC     | Aprendizagem Colaborativa  |
| API    | Application Programming Interface                                  |
| AVA    | Ambiente Virtual de Aprendizagem                                   |
| CDV    | Cego ou Deficiente Visual  |
| CSCL   | Computer-Supported Collaborative Learning                          |
| CSCW   | Computer-Supported Collaborative Work                              |
| DSR    | Design-Science Research  |
| GUI    | Graphical User Interface   |
| IHC    | Interação Humano-Computador  |
| LDB    | Lei de Diretrizes e Bases  |
| OME    | Outro Mais Experiente  |
| OMS    | Organização Mundial da Saúde                                       |
| ONU    | Organização das Nações Unidas                                      |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| PcD    | Pessoa com Deficiência   |
| RSL    | Revisão Sistemática da Literatura                                  |
| STEM   | Science, Technology, Engineering and Mathematics                   |
| TA     | Tecnologia Assistiva   |
| TIC    | Tecnologia da Informação e Comunicação                             |
| TTS    | Text-To-Speech   |
| UML    | Unified Modeling Language  |
| WA     | Workspace Awareness  |
| ZDP    | Zona de Desenvolvimento Proximal                                   |

## LISTA DE FIGURAS

|             |  |    |
|-------------|--|----|
| Figura 1.1  | <i>Workspace</i> do Google Planilhas .....   | 15 |
| Figura 1.2  | Múltiplos usuários no mesmo <i>workspace</i> .....   | 15 |
| Figura 2.1  | Processos de decodificação de texto e codificação de fala no modelo <i>common-form</i> ..... | 21 |
| Figura 2.2  | Círculo interdisciplinar da <i>sonificação</i> e apresentação auditiva.....                  | 22 |
| Figura 2.3  | Dispositivo de toque <i>HyperBraille Refreshable Display</i> .....                           | 24 |
| Figura 2.4  | Dispositivo cinestético <i>Phantom Omni</i> .....  | 25 |
| Figura 2.5  | Dispositivo cinestético Brimhall's Exoskeleton.....  | 25 |
| Figura 2.6  | Informação pictórica: mapa geográfico.....   | 26 |
| Figura 2.7  | Diagramas a partir de texto com a sintaxe da <i>PlantText</i> .....                          | 29 |
| Figura 2.8  | Ferramenta <i>Model2gether</i> e suas interfaces textual e gráfica.....                      | 30 |
| Figura 2.9  | <i>Cross-modal</i> com ferramentas visuais e hápticas juntas.....                            | 31 |
| Figura 2.10 | Diagrama do processo de filtragem.....   | 35 |
| Figura 2.11 | Abordagens para a percepção de diagramas.....  | 37 |
| Figura 2.12 | Abordagens para a interação em diagramas.....  | 37 |
| Figura 2.13 | A solução considerou funcionalidades para colaboração? .....                                 | 38 |
| Figura 2.14 | O estudo empregou métodos empíricos?.....  | 38 |
| Figura 2.15 | Nuvem de palavras-chave.....   | 40 |
| Figura 2.16 | Matriz Tempo/Espaço para <i>Groupware</i> .....  | 41 |
| Figura 2.17 | Zona de Desenvolvimento Proximal.....  | 44 |
| Figura 3.1  | Modelo de processo do método Design Science Research.....                                    | 51 |
| Figura 3.2  | NASA TLX: subescalas e grade do formulário.....  | 53 |
| Figura 3.3  | <i>LucidChart</i> (à esquerda) e <i>Google Drive Draw</i> (à direita).....                   | 55 |
| Figura 3.4  | <i>Sketch</i> da interface auditiva para o <i>awareness</i> de “quem”.....                   | 56 |
| Figura 3.5  | <i>Sketch</i> da interface auditiva para o <i>awareness</i> de “o quê”.....                  | 57 |
| Figura 3.6  | Direções relativas e atalhos de teclado correspondentes.....                                 | 58 |
| Figura 3.7  | Leiautes bidimensional visual-espacial e direcional-relativo.....                            | 59 |
| Figura 3.8  | <i>Sketch</i> da interface auditiva para o <i>awareness</i> de “onde”.....                   | 60 |
| Figura 3.9  | Diagrama de classes do protótipo projetado.....  | 62 |
| Figura 3.10 | Protótipo da ferramenta para modelagem de diagramas.....                                     | 63 |
| Figura 3.11 | Diagramas resultantes de sessões completas de teste.....                                     | 65 |
| Figura 3.12 | Atualização da <i>sonificação</i> da distância para o <i>awareness</i> de “onde”.....        | 67 |
| Figura 3.13 | Nova técnica de <i>sonificação</i> para a percepção de distância.....                        | 68 |
| Figura 3.14 | <i>Sketch</i> da interface baseada em fala sintetizada.....                                  | 69 |
| Figura 3.15 | <i>Sketch</i> da interface auditiva baseada em fala natural pré-gravada.....                 | 70 |
| Figura 3.16 | Modelo háptico tátil para treinamento dos cegos.....   | 74 |
| Figura 3.17 | Procedimento de teste da segunda etapa.....  | 75 |
| Figura 3.18 | Excerto da base de dados ( <i>log</i> ) das sessões de treinamento.....                      | 75 |
| Figura 3.19 | Tabulação dos registros de sinais executados e percebidos.....                               | 76 |
| Figura 4.1  | Taxas de sucesso da interface projetada na primeira iteração.....                            | 77 |
| Figura 4.2  | Dispersão dos resultados.....  | 78 |
| Figura 4.3  | Percepção aproximada ( $\pm 1$ ) de distância.....   | 78 |
| Figura 4.4  | Média das taxas de sucesso por interface e categoria de <i>awareness</i> .....               | 79 |
| Figura 4.5  | NASA-TLX: média de cada subescala por interface.....   | 80 |
| Figura 4.6  | NASA-TLX: distribuição dos <i>scores</i> finais por interface.....                           | 80 |

|   |    |
|---|----|
| Figura 4.7 <i>Replays</i> por tipo de Interface. ....   | 81 |
| Figura 4.8 <i>Replays</i> por tipo de evento / mudança de elemento de WA. ....  | 81 |
| Figura 4.9 Média das taxas de sucesso das interfaces auditivas projetadas na primeira e segunda etapas. ....                          | 82 |
| Figura 4.10 Taxas de sucesso na percepção de distância em <i>SONIv2</i> , <i>SONIv3</i> e <i>SONIv3-aprox.</i> ....                   | 83 |
| Figura 4.11 Taxas de sucesso na percepção por interface <i>sonificada.</i> ....   | 84 |
| Figura 4.12 Taxas de sucesso na percepção de distância exata nas interfaces <i>SONIv1</i> e <i>SONIv3.</i> ....                       | 84 |
| Figura 4.13 Taxas de sucesso na percepção de distância nas interfaces <i>SONIv2</i> , <i>SONIv1-aprox</i> e <i>SONIv3-aprox.</i> .... | 85 |
| Figura 4.14 Carga de trabalho por subescala das interfaces auditivas <i>SONIv2</i> e <i>SONIv3.</i>                                   | 86 |
| Figura 4.15 Carga de trabalho geral das interfaces auditivas <i>SONIv2</i> e <i>SONIv3.</i>   | 86 |
| Figura 4.16 Taxas de sucesso nas interfaces <i>faladas FALASIN</i> e <i>FALANAT.</i> ....   | 87 |
| Figura 4.17 Carga de trabalho por subescala da NASA-TLX das interfaces <i>faladas FALASIN</i> e <i>FALANAT.</i> ....                  | 87 |
| Figura 4.18 Taxas de sucesso entre as interfaces <i>sonificadas</i> e <i>faladas.</i>   | 88 |
| Figura 4.19 Distribuição das taxas de sucesso entre as interfaces <i>sonificadas</i> e <i>faladas.</i>                                | 88 |
| Figura 4.20 Resultados das subescalas NASA-TLX das interfaces <i>sonificadas</i> e <i>faladas.</i>                                    | 89 |
| Figura 4.21 Taxas de sucesso dos participantes videntes e cegos em todas interfaces. ..   | 92 |
| Figura 4.22 Carga de trabalho por subescala da NASA-TLX comparativo dos participantes videntes e cegos. ....                          | 92 |

## LISTA DE TABELAS

|            |   |    |
|------------|---|----|
| Tabela 2.1 | <i>Strings</i> de busca submetidas às bibliotecas digitais.....                                   | 34 |
| Tabela 2.2 | Artigos atendidos pelos critérios de inclusão. ....   | 34 |
| Tabela 2.3 | As 5 áreas de projeto de CSCW e suas características. ....  | 41 |
| Tabela 2.4 | Elementos de <i>workspace awareness</i> relacionados ao presente. ....                            | 48 |
| Tabela 2.5 | Elementos de <i>workspace awareness</i> relacionados ao passado. ....                             | 48 |
| Tabela 3.1 | Mecanismos para elementos de <i>workspace awareness</i> acessíveis. ....                          | 54 |
| Tabela 3.2 | Questões submetidas aos participantes. ....   | 65 |
| Tabela 3.3 | Últimos quatro casos de teste do 2º participante. ....  | 65 |
| Tabela 3.4 | Comparação das decisões de projeto para cada interface. ....                                      | 71 |
| Tabela 4.1 | Taxas de sucesso por participante e categoria de <i>awareness</i> .....                           | 77 |
| Tabela 4.2 | Significância dos resultados do NASA-TLX: interfaces <i>sonificadas</i> e<br><i>faladas</i> ..... | 89 |

## SUMÁRIO

|  |           |
|--|-----------|
| <b>1 INTRODUÇÃO</b> .....  | <b>13</b> |
| <b>1.1 Objetivos</b> .....   | <b>17</b> |
| 1.1.1 Objetivos Geral .....  | 17        |
| 1.1.2 Objetivos Específicos .....                                      | 17        |
| <b>1.2 Motivação</b> .....   | <b>17</b> |
| <b>1.3 Organização do Texto</b> .....                                  | <b>18</b> |
| <b>2 FUNDAMENTOS TEÓRICOS</b> .....                                    | <b>19</b> |
| <b>2.1 Tecnologia Assistiva</b> .....                                  | <b>19</b> |
| <b>2.2 Interfaces Acessíveis aos Cegos</b> .....                       | <b>19</b> |
| 2.2.1 Apresentação Auditiva .....                                      | 20        |
| 2.2.2 Interfaces Hápticas .....  | 24        |
| <b>2.3 Fontes de Informação</b> .....                                  | <b>25</b> |
| 2.3.1 Informações Textuais .....                                       | 26        |
| 2.3.2 Informações Pictóricas.....                                      | 26        |
| 2.3.3 Informações Diagramáticas .....                                  | 27        |
| <b>2.4 Acessibilidade de Diagramas</b> .....                           | <b>27</b> |
| 2.4.1 Interfaces Audíveis .....  | 28        |
| 2.4.2 Interfaces Hápticas .....  | 28        |
| 2.4.3 Soluções Mistas .....  | 29        |
| 2.4.4 Interação em Diagramas .....                                     | 29        |
| 2.4.5 Recursos Colaborativos.....                                      | 30        |
| <b>2.5 Revisão Sistemática</b> .....                                   | <b>31</b> |
| 2.5.1 Método de Revisão Sistemática .....                              | 31        |
| 2.5.2 Questões da Revisão Sistemática.....                             | 32        |
| 2.5.3 <i>Strings</i> de Busca.....                                     | 32        |
| 2.5.4 Bibliotecas Digitais.....  | 33        |
| 2.5.5 Critérios de Inclusão e Exclusão.....                            | 34        |
| 2.5.6 Avaliação da Qualidade .....                                     | 36        |
| 2.5.7 Resultados da Revisão Sistemática .....                          | 36        |
| 2.5.8 Ameaças à Validade .....   | 39        |
| 2.5.9 Considerações .....  | 39        |
| <b>2.6 Aplicações Colaborativas</b> .....                              | <b>40</b> |
| 2.6.1 Groupware.....   | 40        |
| 2.6.2 <i>Computer-Supported Cooperative Work</i> .....                 | 41        |
| 2.6.3 <i>Computer-Supported Collaborative Learning</i> .....           | 41        |
| 2.6.4 Fundamentos Pedagógicos.....                                     | 42        |
| 2.6.5 Mecanismos para Colaboração .....                                | 46        |
| 2.6.6 Workspace Awareness.....   | 47        |
| <b>3 MATERIAIS E MÉTODOS</b> .....                                     | <b>49</b> |
| <b>3.1 Problema de Pesquisa</b> .....                                  | <b>49</b> |
| <b>3.2 Metodologia</b> .....   | <b>50</b> |
| 3.2.1 Design Science Research .....                                    | 50        |
| 3.2.2 Teste de Usabilidade: Taxa de Sucesso.....                       | 51        |
| 3.2.3 Medição de Carga de Trabalho: NASA-RTLX.....                     | 52        |
| <b>3.3 Elementos de Workspace Awareness Acessíveis aos Cegos</b> ..... | <b>53</b> |
| 3.3.1 <i>Awareness</i> de “quem” .....                                 | 54        |
| 3.3.2 <i>Awareness</i> de “o quê” .....                                | 56        |
| 3.3.3 <i>Awareness</i> de “onde” .....                                 | 57        |

|   |           |
|---|-----------|
| <b>3.4 Primeira Iteração: Uma Interface Auditiva para Workspace Awareness.....</b>            | <b>61</b> |
| 3.4.1 Projeto e Implementação.....  | 61        |
| 3.4.2 Avaliação.....  | 62        |
| <b>3.5 Segunda Iteração: Atualização da Interface Auditiva e Opções Baseadas em Fala.....</b> | <b>66</b> |
| 3.5.1 Atualização da Interface Auditiva.....  | 66        |
| 3.5.2 Projeto de Interfaces Baseadas em Fala.....   | 68        |
| 3.5.3 Considerações para Avaliação .....  | 70        |
| 3.5.4 Atualizações do Protótipo .....   | 71        |
| 3.5.5 Avaliação.....  | 72        |
| <b>4 RESULTADOS .....</b>   | <b>77</b> |
| <b>4.1 Primeira Iteração .....</b>  | <b>77</b> |
| <b>4.2 Segunda Iteração.....</b>  | <b>79</b> |
| 4.2.1 Interfaces Auditivas .....  | 82        |
| 4.2.2 Interfaces Baseadas em Fala .....   | 86        |
| 4.2.3 Interfaces <i>Sonificadas e Faladas</i> .....   | 87        |
| <b>4.3 Discussão dos Resultados .....</b>   | <b>89</b> |
| 4.3.1 Limitações.....   | 92        |
| <b>5 CONCLUSÃO .....</b>  | <b>94</b> |
| <b>5.1 Trabalhos Futuros.....</b>   | <b>94</b> |
| <b>REFERÊNCIAS.....</b>   | <b>96</b> |

## 1 INTRODUÇÃO

De acordo com os dados globais sobre deficiência visual, divulgados pela Organização Mundial da Saúde (OMS) em 2010, foi estimado que 285 milhões de pessoas tinham alguma forma de deficiência visual e que 39 milhões eram cegas (PASCOLINI; MARIOTTI, 2012). No mesmo ano, dados do censo realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) contabilizaram 35,7 milhões de Brasileiros com deficiência visual; destes, 6 milhões com deficiência grave e 500 mil cegos (IBGE, 2010). Os números apresentados podem aparentar uma baixa representação, se considerada a população total, mas são bastante significativos se observadas as 13.102 matrículas de estudantes com baixa visão e cegos nos cursos superiores em 2016, conforme dados do Censo da Educação Superior de 2016 liberado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2018b). Cegos, deficientes visuais e outras pessoas com deficiência têm o direito à educação, defendido mundialmente.

A educação de Pessoas com Deficiência (PcD) é abordada na Convenção Sobre os Direitos das Pessoas com Deficiência da Organização das Nações Unidas (ONU). O artigo 24 declara que “*os Estados Partes assegurarão sistema educacional inclusivo em todos os níveis, bem como o aprendizado ao longo de toda a vida*” (UN, 2007). Quanto a legislação nacional, a Lei de Diretrizes e Bases (LDB) (lei nº 9.394/1996) assegura o direito ao currículo, métodos específicos e recursos educacionais necessários para atender suas necessidades (BRASIL, 1996). Em seguida, a lei 10.172/2001 estabeleceu metas para a redução da desigualdade social e regional, assegurando aos estudantes o acesso à educação pública e a permanência nesta. O número de estudantes beneficiados com a reserva de vagas para PcD aumentou seis vezes de 2010 para 2012 (455 para 2698 vagas) (INEP, 2018a; INEP, 2018b). Grande parte, devido às políticas de ações afirmativas estabelecidas pelas leis 12.711/2012 (BRASIL, 2012) e 13.409/2016 (BRASIL, 2016), que determinam cotas para minorias em cursos disponíveis nas universidades e institutos federais, contudo, o ingresso é apenas a primeira barreira para as PcDs.

Existem vários obstáculos para o ensino de estudantes cegos e deficientes visuais que exigem soluções alternativas para apresentar conceitos que seriam tipicamente apenas ilustrados. Sobretudo, nos cursos STEM<sup>1</sup>, que aplicam diagramas para expor dados estruturados e seus relacionamentos (CRYER et al., 2013). Estas barreiras se apre-

---

<sup>1</sup>STEM é um termo usado para agrupar as disciplinas acadêmicas Science, Technology, Engineering and Mathematics (Ciência, Tecnologia, Engenharia e Matemática) no tratamento de políticas educacionais e escolhas curriculares para o desenvolvimento científico e tecnológico

sentam, por exemplo, desde o ensino de gráficos estatísticos (GODFREY; MURRELL; SORGE, 2018) e estruturas moleculares (BROWN; PETTIFER; STEVENS, 2004), até a modelagem UML (LUQUE et al., 2014; MÜLLER, 2012) e projeto de bancos de dados (BROOKSHIRE, 2006). Ainda que a percepção e interação em diagramas sejam assuntos frequentes na pesquisa atual, poucos estudos consideram a modelagem colaborativa. Consequentemente, a necessidade de mais pesquisa e desenvolvimento acerca de recursos para colaboração guiou este estudo na direção da acessibilidade de aplicativos de *groupware* para modelagem de diagramas que, entre outros benefícios, pode apoiar um importante recurso pedagógico conhecido como aprendizagem colaborativa.

A aprendizagem colaborativa (AC) é uma abordagem em que grupos de dois ou mais estudantes trabalham na direção de uma meta comum (GOKHALE, 1995). É uma abordagem muito diferente dos métodos tradicionais de ensino, nos quais o professor detém o conhecimento e é responsável por transmiti-lo aos aprendizes. Não obstante, cegos e deficientes visuais nem sempre possuem meios para se engajarem em atividades colaborativas com outros colegas, sejam estes videntes ou cegos, com destaque às atividades suportadas por computador. Segundo Kohlmann (2012), o trabalho colaborativo *on-line* envolve geralmente os videntes, enquanto a colaboração entre videntes e uma ou mais pessoas cegas é muito raro. Inclusive, ambientes virtuais de aprendizado e ferramentas colaborativas são raramente projetados com recursos de acessibilidade que permitam, por exemplo, a sua operação via Tecnologia Assistiva, como leitores de tela (CLAUDIA; BUZZI; MORI, 2012). Os estudos sobre o suporte aos recursos colaborativos para cegos e deficientes visuais deixam abertas questões sobre como tornar o *workspace awareness* acessível (SÁNCHEZ; BALOIAN, 2006a).

O termo *workspace* é usado para denotar um espaço de trabalho virtual que disponibiliza aos usuários a interface necessária para a conclusão de uma determinada tarefa (Figura 1.1). Tratando-se do trabalho *on-line*, esse *workspace*<sup>2</sup> é compartilhado por mais de um usuário em sessões colaborativas, nas quais alguns usuários podem editar as informações (como textos, diagramas, slides, figuras, etc), enquanto todos podem observar os efeitos e estarem cientes (*being aware*) das mudanças. Dourish e Bellotti (1992) definem *awareness* como “um ‘entendimento das atividades dos outros’, que provê um ‘contexto para suas próprias atividades’”, reconhecido como um pilar para o trabalho em conjunto (COLLAZOS et al., 2002).

---

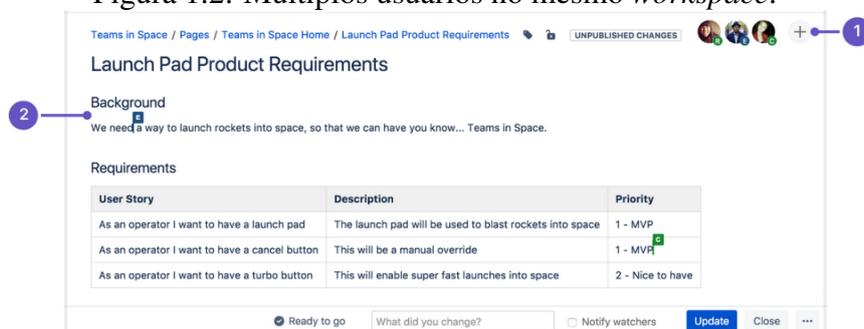
<sup>2</sup>Neste trabalho o termo *workspace* será preferido no lugar de “espaço de trabalho” devido a sua difusão na literatura.

Figura 1.1: *Workspace* do Google Planilhas

|   | A   | B         | C                                 | D        | E         | F  | G          |
|---|---|-----------|-----------------------------------|----------|-----------|--|------------|
| 1 | Workspace   | Estudo    | PERCEPÇÃO                         |          | INTERAÇÃO |  |            |
| 2 |   | Saída     | Abordagem                         | Detalhes | Entrada   | Abordagem                                  | Suportado? |
| 3 | Pölzer, Stephan, and Klaus Miesenberger. "A Tactile Presentation Method of Mind maps in co-located Meetings." The Proceedings of Workshop Tactile/Haptic User Interfaces for Tabletops and Tablets 2014. 2014.                  | tátil     | display<br>braille<br>atualizável |          | toque     | display<br>braille<br>sensível ao<br>toque | sim        |
| 4 | Grillo, Filipe Del Nero, and Renata Pontin de Mattos Fortes. "Accessible modeling on the web: a case study." Procedia Computer Science 27 (2014): 460-470.  | auditória | fala                              |          | teclado   |  | sim        |
| 5 | Godfrey, A. Jonathan R., Paul Murrell, and Volker Sorge. "An Accessible Interaction Model for Data Visualisation in Statistics." International Conference on Computers Helping People with Special Needs. Springer, Cham, 2018. | auditória | fala                              |          | nenhum    |  | não        |
| 6 | Grussenmeyer, William, and Eelke Folmer. "AudioDraw: user preferences in non-visual diagram drawing for..."   |           |                                   |          |           | tablet com<br>tela                         |            |

Fonte: *print* de planilha do autor.

Para projetar o *workspace awareness*, os engenheiros de *software* e de *groupware* podem apoiar-se sobre *frameworks* conhecidos e estabelecidos, que listam e resumizam os elementos necessários para habilitar o *awareness*. Por exemplo, alguns elementos de *awareness* listados no trabalho de Gutwin e Greenberg (2002), como saber “*quem está fazendo o quê e onde*”, são implementados em aplicativos *on-line* como o Google Drive (Fig. 1.1), Microsoft Office 365, Trello, e outros aplicativos colaborativos. Entretanto, esses elementos oferecem pistas visuais e são destinados ao público geral. Por exemplo, o editor Confluence da Atlassian (Figura 1.2) permite a escrita colaborativa e implementa o *awareness* por meio de múltiplos cursores coloridos e identificados por uma letra associada ao avatar dos usuários presentes no *workspace*. No entanto, existem barreiras quanto à acessibilidade desses elementos, na escrita colaborativa (BUZZI et al., 2014; DAS, 2018; SCHOEBERLEIN; WANG, 2012) e outras aplicações.

Figura 1.2: Múltiplos usuários no mesmo *workspace*.

1. **Invite more people:** see who is editing the page and invite others to edit with you.
2. **See what they're doing:** watch others edit the page in real time.

Fonte: adaptado de Atlassian Confluence Support Documentation. Disponível em <<https://confluence.atlassian.com/doc/collaborative-editing-858771779.html>>. Acesso em: 3 dez. 2019.

Uma revisão sistemática (Seção 2.5) foi conduzida para obter o um panorama do

estudo sobre acessibilidade de representações diagramáticas. Obteve-se evidências de que os pesquisadores têm trabalhado para tornar os diagramas acessíveis às pessoas cegas, por meio de interfaces hápticas e recursos audíveis. Entretanto, os resultados também demonstraram uma lacuna na pesquisa sobre funcionalidades para colaboração acessíveis, que tornou-se o tema central deste estudo.

Segundo a literatura, percebe-se que um dos problemas recorrentes, que exige muitas vezes o suporte háptico, é representar a posição dos elementos. Por isso, este trabalho suporta os diagramas que podem ser abstraídos como grafos, nos quais a posição bidimensional dos elementos é irrelevante, tratando a conectividade dos nós (*nodes*) através das ligações (*links*, diagramas do tipo *node-link*). Isto é, as técnicas são genéricas, mas não cobrem todos as classes de diagramas (gráficos, por exemplo). No entanto, a interface projetada pode apoiar os diagramas usados nos cursos STEM, como a representação de diagramas UML, fluxogramas e de entidade-relacionamento — este último usado no protótipo para experimentos.

O projeto da interface foi separado em duas etapas, ou iterações, segundo o método *Design Research Science*. Na primeira, foi projetada uma interface auditiva para os elementos de *workspace awareness* baseada em *earcons*, que foi implementada em um protótipo e avaliada em teste de usabilidade com voluntários. Na segunda, os pontos negativos foram revistos, guiando o reprojeto da interface auditiva proposta inicialmente, a inclusão de uma variante, e a implementação de duas interfaces baseadas em fala, uma sintetizada e outra natural.

As interfaces foram novamente avaliadas, por meio da taxa de sucesso no reconhecimento dos elementos de *workspace awareness* com base apenas nas pistas auditivas. Também foram coletados dados sobre a experiência subjetiva por meio da ferramenta NASA-TLX, que permite mensurar a carga de trabalho percebida. Os resultados obtidos na percepção dos sinais com as interfaces auditivas foram equiparáveis às interfaces baseadas em fala.

A conclusão é de que é viável *sonificar* os elementos *workspace awareness* para construir uma interface auditiva que permita aos cegos acompanhar as ações dos participantes em sessões de modelagem colaborativa de diagramas. Ao longo dessa dissertação são detalhados a definição do escopo, metodologia para projetar uma interface auditiva, sua análise, validação e resultados.

## 1.1 Objetivos

### 1.1.1 Objetivos Geral

O objetivo geral é propor e avaliar interfaces auditivas para tornar os elementos de *workspace awareness* acessíveis aos cegos em ferramentas de *groupware* para a modelagem colaborativa de diagramas. Os elementos são baseados no trabalho de Gutwin e Greenberg (2002) intitulado “*A Descriptive Framework of Workspace Awareness for Real-Time Groupware*”<sup>3</sup>, que classifica as funcionalidades necessárias para que os usuários percebam as ações dos coparticipantes e acompanhem as mudanças no espaço de trabalho.

### 1.1.2 Objetivos Específicos

Este estudo foi conduzido em etapas as quais responderam questões específicas descritas pelos objetivos específicos a seguir:

- Estudar que tipos de interfaces estão sendo pesquisadas, desenvolvidas e aplicadas para a percepção de diagramas, bem como a interação e colaboração;
- Identificar lacunas na pesquisa e possíveis fragilidades acerca da acessibilidade de diagramas para cegos;
- Reunir questões específicas acerca de ferramentas colaborativas acessíveis;
- Projetar interfaces auditivas que permitam a colaboração acessível em ferramentas de modelagem de diagramas;
- Projetar e implementar um protótipo acessível para modelagem colaborativa de diagramas;
- Avaliar a acessibilidade da interface;

## 1.2 Motivação

A motivação deste estudo começou após o ingresso de estudantes cegos no curso Técnico em Informática para Internet no Instituto Federal de Educação, Ciência e Tec-

---

<sup>3</sup>Um Framework Descritivo de Workspace Awareness para Groupware em Tempo Real

nologia do Rio Grande do Sul campus Rio Grande. Neste instituto, o autor leciona em disciplinas como lógica de programação, programação orientada a objetos, engenharia de software e projeto de bancos de dados. A partir dessa experiência, foi percebida a falta de recursos de acessibilidade nas atuais ferramentas para a modelagem de diagramas, sobretudo para o trabalho em grupo. Assim, o presente estudo não só aspira contribuir cientificamente, como também tecnologicamente.

### 1.3 Organização do Texto

O presente trabalho é composto por seis capítulos incluindo esta introdução. A contribuição de cada capítulo do estudo está listada a seguir:

1. *Introdução*: contextualiza e descreve a motivação, objetivos e organização deste estudo;
2. *Fundamentos Teóricos*: aborda os conceitos básicos ligados às interfaces acessíveis e diagramas, revisa o estado-da-arte, delimita o escopo e discorre sobre aplicações colaborativas e assuntos correlatos;
3. *Materiais e Métodos*: declara a hipótese de pesquisa, descreve a metodologia utilizada, aborda os trabalhos similares, apresenta as decisões de projeto para a *sonificação* dos elementos de *workspace awareness*, descreve o projeto do protótipo e método de avaliação;
4. *Resultados*: apresenta e discute os resultados obtidos;
5. *Conclusão*: retoma os objetivos, faz a autocrítica do estudo, e declara as lições e orientações para estudos futuros;
6. *Referências*: lista as referências bibliográficas usadas neste estudo.

## 2 FUNDAMENTOS TEÓRICOS

Este capítulo apresenta os conceitos essenciais para a compreensão do quadro geral, a saber, as principais categorias de informações visuais e suas possíveis representações alternativas, com destaque para as interfaces auditivas e hápticas e suas aplicações com informações diagramáticas.

### 2.1 Tecnologia Assistiva

As pessoas com deficiência (PcD) enfrentam limitações que impedem ou dificultam o acesso a certos recursos, físicos, de *hardware* ou *software*. As soluções para superar essas limitações são chamadas de Tecnologia Assistiva (TA) — também conhecidas como adaptativas ou Ajuda Técnica. Logo, considera-se TA o conjunto de artefatos disponíveis às PcDs, de modo à garantir-lhes independência, qualidade de vida e inclusão social (SONZA et al., 2013). Segundo a Lei. nº 10.098 de 2000 (BRASIL, 2000), considera-se Ajuda Técnica “*qualquer elemento que facilite a autonomia pessoal ou possibilite o acesso e o uso do meio físico*”. Isso inclui bengalas, cadeiras de roda, aparelhos auditivos e outros. Para pessoas com baixa visão, especificamente, alguns exemplos de Tecnologia Assistiva são lupas, opções de alto contraste, fontes diferenciadas, entre outros. Para os cegos, destacam-se o Braille, pisos táteis, sinais sonoros, formas em relevo e sintetizadores de voz. Este estudo, no caso, preocupa-se com a Tecnologia Assistiva para acessar informações gráficas. Portanto, na seção seguinte são apresentadas algumas interfaces acessíveis aos cegos que podem ser empregadas como soluções para Tecnologia Assistiva.

### 2.2 Interfaces Acessíveis aos Cegos

As pessoas cegas tipicamente percebem o mundo por meio de recursos audíveis ou hápticos. Os recursos audíveis podem ser tanto verbais como não-verbais. Isto é, que usam a fala para comunicação, ou outros sons que não são reconhecidos como um discurso. Enquanto recursos hápticos, destacam-se as soluções táteis e cinestéticas. Isto é, enquanto o primeiro trata da sensibilidade nos dedos, mão e pés, o segundo trata da sensação em músculos e tendões. Ambos, o audível e o háptico, são discutidos com mais

detalhes nas seções a seguir.

### 2.2.1 Apresentação Auditiva

As pessoas cegas estão acostumadas à representação auditiva de informações, que é uma das maneiras primárias pelas quais percebem o mundo. Os recursos audíveis são empregados de várias maneiras na construção de interfaces e podem ser classificados, primeiramente, em dois grupos: verbais e não-verbais.

A saída verbal, conhecida como (A) Síntese de Fala ou Text-To-Speech, é amplamente difundida, sobretudo devido aos leitores de tela disponíveis nos sistemas operacionais. Nesse tipo de interface, o conteúdo é falado aos usuários, seja por voz sintética ou gravada. As interfaces não-verbais, em vez, utilizam sons que não são reconhecidos como discurso para a representação de informações ou resposta às interações, também conhecidas como (B) Interfaces Auditivas ou (C) *Sonificação*. Como técnicas para o projeto de interfaces auditivas estão os (D) Ícones Auditivos, (E) *Earcons* e (F) outras. Esses tipos são discutidos nos tópicos a seguir.

#### A) *Síntese de Fala e Text-To-Speech*

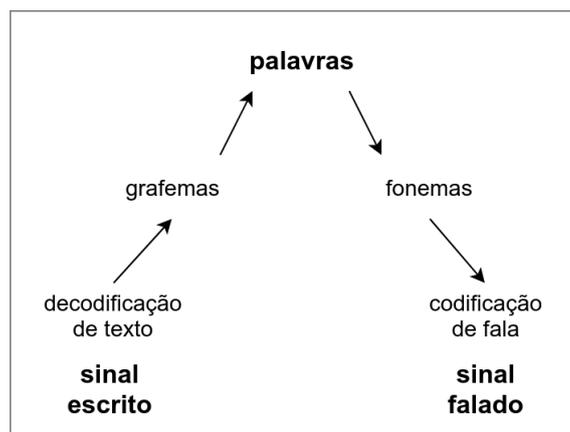
A síntese de fala é a produção artificial da fala humana, isto é, uma saída audível do tipo verbal, normalmente gerada por computador que pode ser implementada tanto em *software* como em *hardware*. Sua função é de converter informações simbólicas em enunciados que são falados aos usuários, tipicamente, informações textuais. O motor do sintetizador de fala analisa o texto, passando-o por várias análises linguísticas, para produzir a fala. O sistema de ponta-a-ponta, desde a análise textual até a produção dos enunciados, é conhecido como Text-To-Speech<sup>1</sup> (TTS).

Existem diversos modelos para TTS, tais como os *signal-to-signal*, *pipelined*, e cada um tem sua organização específica dos componentes. Por exemplo, o modelo *Common Form* (Figura 2.1) baseia-se essencialmente em dois componentes: (a) um sistema de análise de texto que decodifica o texto e constrói uma “forma” e (b) um sistema de síntese de fala que codifica a “forma” como fala. A lógica desse modelo é que tanto a escrita quanto a fala compartilham uma “forma comum”, que serve como ponte entre ambas representações (TAYLOR, 2009).

---

<sup>1</sup>Texto-Para-Fala em português, no entanto esse recurso é mais conhecido pela expressão original em inglês

Figura 2.1: Processos de decodificação de texto e codificação de fala no modelo *common-form*.



Fonte: adaptado de Taylor (2009).

Atualmente, os motores de TTS se destacam como uma das mais importantes tecnologias para acessibilidade de interfaces gráficas com o usuário (GUI - *Graphical User Interfaces*), distribuídos, por exemplo, como utilitários de leitura de tela (*Screen Reader*) disponíveis em computadores e *smartphones* (OLOFSSON, 2018), como JAWS<sup>2</sup>, NVDA<sup>3</sup>, VoiceOver<sup>4</sup> e Orca<sup>5</sup>.

### B) Interface Auditiva

No estudo das interfaces auditivas, segundo Hermann, Hunt e Neuhoff (2011), os pesquisadores examinam como o sistema auditivo pode ser usado como o principal canal de interface para comunicar e transmitir informação por meio de sons não-verbais. O campo é abrangente e inclui sistemas de interação, configurações de *hardware* e soluções técnicas para processamento de dados em formato audível. Segundo Hermann (2008), como um novo campo, ainda não há um consenso acerca das definições dos termos. Porém, sabe-se que a *sonificação* (*sonification*, em Inglês) é o subcampo conhecido por lidar com a renderização de sons em resposta aos dados e interações, por meio de técnicas como ícones auditivos e *earcons*, por exemplo.

<sup>2</sup><<https://www.freedomscientific.com/products/software/jaws/>>

<sup>3</sup><<https://www.nvaccess.org/>>

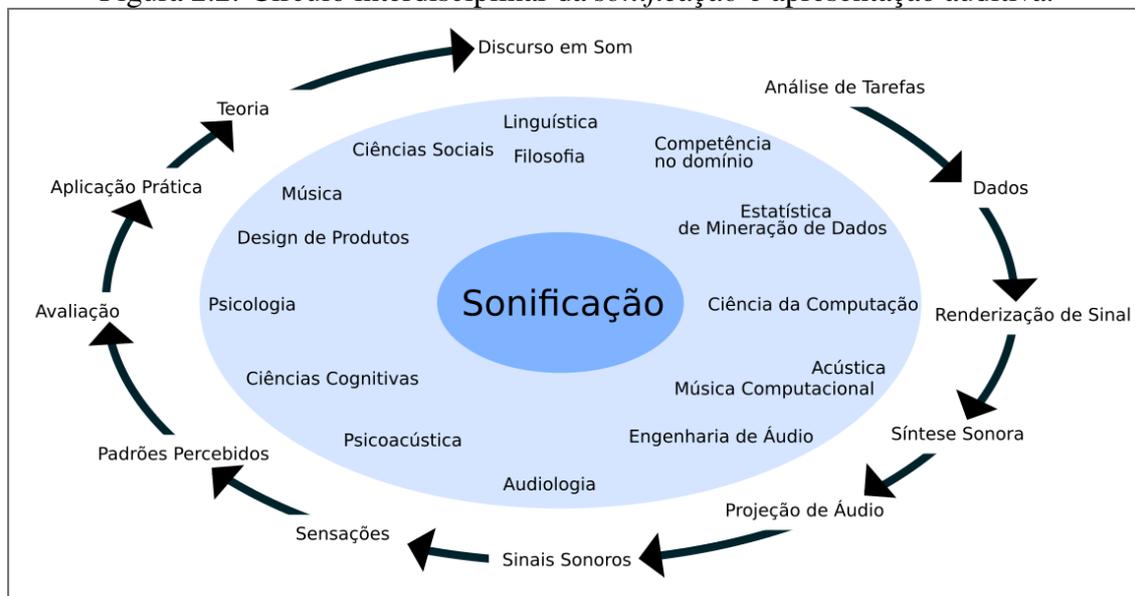
<sup>4</sup><<https://www.apple.com/br/accessibility/iphone/vision/>>

<sup>5</sup><<https://help.gnome.org/users/orca/stable/>>

### C) Sonificação

Conforme Hermann (2008), *sonificação* é a transformação das relações entre dados em um sinal audível, passível de comunicação e interpretação, porém, sem ser reconhecido como um discurso. A área da *sonificação* é multidisciplinar e abrange a linguística, acústica, engenharia de áudio, psico-acústica, etc. Como pode ser visto na Figura 2.2, o círculo interno lista as disciplinas científicas enquanto o círculo externo apresenta as transformações sequenciais da informação. As aplicações da *sonificação* estão em vários elementos do cotidiano, como nos automóveis, equipamentos eletrônicos e outros dispositivos, que usam sons para entregar mensagens aos usuários (HERMANN; HUNT; NEUHOFF, 2011).

Figura 2.2: Círculo interdisciplinar da *sonificação* e apresentação auditiva.



Fonte: adaptado de Hermann, Hunt e Neuhoff (2011).

Existem várias representações que podem ser usadas como técnicas de *sonificação*, isto é, para converter informações em sinais audíveis não-verbais. Nos tópicos a seguir, são abordados os exemplos mais comuns na área, os ícones auditivos (2.2.1) e *earcons* (2.2.1).

### D) Ícones Auditivos

Os ícones auditivos são uma técnica de *sonificação* para o projeto de interfaces auditivas, que têm a intenção de imitar os sons do mundo real e pretendem facilitar a compreensão do usuário por representar mapeamentos diretos. Por exemplo, em um aplicativo de bate-papo, é possível reproduzir o som da datilografia em uma máquina de

escrever para comunicar que os outros usuários estão digitando algo no mesmo instante. Outro exemplo é o som de um papel sendo rasgado para o sinalizar a exclusão de um documento. Como o nome sugere, são sons semelhantes à ideia de ícones, supondo que os usuários capturem o significado da mesma maneira que nos ícones visuais, por meio de uma relação entre significante e significado (GAVER, 1986).

#### E) *Earcons*

Os *earcons* são outro tipo de técnica para *sonificação* que baseia-se na reprodução de sons artificiais abstratos distintos o suficiente para representar eventos específicos ou transmitir informações. São diferentes dos ícones auditivos por não possuírem relação natural entre significante e significado, isto é, os usuários precisam aprender o significado de cada som. Segundo Nees e Walker (2009), os *earcons* permitem uma flexibilidade maior nas representações, pois podem ser atribuídos a quase qualquer objeto, item ou processos nas interfaces. Sem a correspondência direta com o mundo real, se apoiam no projeto da acústica, como a variação do timbre, altura, frequência, intensidade e amplitude do sinal audível, para transmitir e comunicar diferentes partes de informação (BLATTNER; SUMIKAWA; GREENBERG, 1989). Alguns exemplos de *earcons* são sons de alarmes, alertas e avisos, com *beeps* ou melodias.

#### F) *Outras Técnicas de Sonificação*

Existem outras abordagens para *sonificação* apresentadas na literatura, embora a maioria pareça projetada como uma variação das anteriores. No estudo de Csapó e Wersényi (2013), há uma visão geral das representações auditivas comuns. Nesse estudo, além de discutir as previamente abordadas, os autores introduzem as categorias resumidas na lista a seguir:

- *Spearcons*: são sons de fala super-acelerada que, embora não reconhecidas como um discurso, são diferenciáveis o suficiente para comunicar informações;
- *Spindices*: é uma variante dos *spearcons* que usa uma fala super-acelerada da primeira letra das palavras;
- *Auditory Emoticons*: são análogos aos *emoticons* (ou *emojis*) gráficos, representados por sons de risadas, beijos e outros;
- *Musicons*: são pequenos trechos de músicas populares;

- *Spemoticons*: representam estados emocionais e intencionais com sons acelerados, como risadas e tossidas, por exemplo. São semelhantes aos *spearcons*, sem representar um discurso, mas reconhecível;

### 2.2.2 Interfaces Hápticas

Outro modo de transmitir informações é por meio do sensorial háptico que, diferente das interfaces auditivas, está relacionado às sensações percebidas pelo movimento do corpo. Existem duas classes de háptico conhecidas como (A) tátil e (B) cinestético, detalhadas a seguir.

#### A) Interfaces Táteis

Segundo Loomis e Lederman (1986), o toque é uma percepção mediada pelo estímulo cutâneo que acontece geralmente, mas não somente, pelas pontas dos dedos, na sensação da textura, pressão, temperatura e vibração de uma superfície.

A interface tátil mais difundida é o sistema Braille de escrita, publicado em 1829 por Louis Braille, que consiste em uma matriz de seis pontos em alto-relevo. A combinação desses pontos formam letras, números, sinais de pontuação e outros símbolos (SONZA et al., 2013). Para ler, os usuários cegos sentem os símbolos por meio do toque a medida que passam a ponta de seus dedos sobre os pontos em relevo.

Para produção de material gráfico, existem opções de impressão em relevo, impressão tridimensional e *displays* Braille atualizáveis (Fig. 2.3), que permitem representar tanto texto como outras formas.

Figura 2.3: Dispositivo de toque *HyperBraille Refreshable Display*.



Fonte: HyperBraille<sup>6</sup>.

## B) Interfaces Cinestéticas

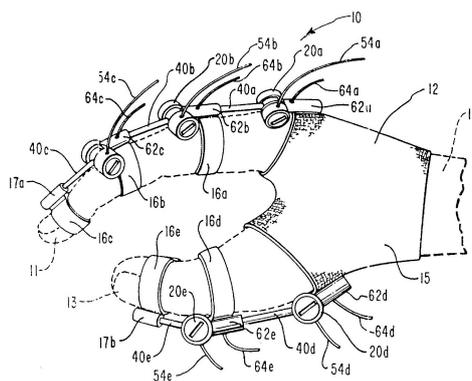
A percepção cinestética é mediada pela sensação de resistência nos músculos e articulações e percebida como peso e resistência ao movimento, por exemplo, de braços, pernas, mão ou corpo inteiro (LOOMIS; LEDERMAN, 1986). Dispositivos cinestéticos tipicamente exigem hardware mais sofisticado e envolvem partes mecânicas ostensivas. Um exemplo de aplicação da cinestética é a tecnologia de *force feedback* disponível em *gamepads* e volantes. Exemplos mais sofisticados são o Phantom Omni Device (Fig. 2.4), que permite a sensação de objetos virtuais em um plano tridimensional, e exoesqueletos, tais como o inventado por Brimhall (1996) para controlar atuadores remotos (Fig. 2.5).

Figura 2.4: Dispositivo cinestético *Phantom Omni*.



Fonte: 3D Systems Inc.<sup>7</sup>.

Figura 2.5: Dispositivo cinestético Brimhall's Exoskeleton.



Fonte: Brimhall (1996).

## 2.3 Fontes de Informação

Apesar de existirem diversas fontes de informações visuais que podem ser categorizadas (MASSIRONI et al., 2001), este estudo aborda três tipos básicos: informação textual, pictórica e diagramática. Todas são brevemente explicadas, ainda que a última, como objeto desta pesquisa, seja mais abrangentemente discutida.

### 2.3.1 Informações Textuais

As informações textuais são tipicamente representadas como informações simbólicas dispostas linearmente e em sequência. É o principal meio de transmitir informação e amplamente usada, disponível em livros, jornais e outras mídias escritas. Existem algumas variações na aparência e semântica, como tipografia, tamanho, peso, cor, parágrafos, colunas e outras questões de leiaute. Normalmente as pessoas cegas acessam o conteúdo textual por meio de tecnologia tátil (sistema Braille, tipicamente) ou síntese de fala nas informações textuais digitais (arquivos .txt, .pdf, .doc, etc).

### 2.3.2 Informações Pictóricas

As informações pictóricas são passadas por meio de representações gráficas como pinturas e desenhos. Por exemplo, incluem afrescos, mapas geográficos (Fig. 2.6), fotografias e plantas baixas arquiteturais. Fontes visuais de informações dessa classe podem ter texto embutido e ainda que apresentem alguma conectividade ela não é tão importante quanto o leiaute, isto é, a posição exata de cada ponto (ou *pixel*) do conteúdo no plano bidimensional. Dada sua natureza, as figuras são mais difíceis de representar alternativamente do que diagramas, isto é, não são facilmente convertíveis para informação textual que poderia ser falada ao usuário por TTS (KING, 2007). Consequentemente, é usual codificar informações pictóricas por meio de recursos hápticos como impressão em relevo, impressão 3D, modelos de madeira, EVA ou outras soluções artesanais, ou por meio de reconhecimento de padrões por visão computacional.

Figura 2.6: Informação pictórica: mapa geográfico.



Fonte: licenciado sob ODbL por OpenStreetMap<sup>8</sup>.

### 2.3.3 Informações Diagramáticas

As informações diagramáticas são compostas de conteúdo simbólico em um espaço bidimensional com ênfase nas conexões, isto é, os diagramas são usados para ilustrar relacionamentos entre entidades de forma concisa em vez de descrevê-los textualmente. As informações diagramáticas estão presentes em vários contextos, por exemplo, em lugares públicos, aeroportos, estações de ônibus e relatórios técnicos. Os diagramas também são usados frequentemente na sala de aula e nos livros didáticos em várias disciplinas, especialmente no campo STEM. As informações diagramáticas mais comumente empregadas estão listadas a seguir:

- *Unified Modeling Language (UML)*: usados para representar classes, métodos, pacotes e outros conceitos e artefatos da programação orientada a objetos;
- *Systems Modeling Language (SysML)*: o mesmo que UML, porém mais abstrato e genérico;
- *Data-flow Diagram (DFD)*: usados para apresentar o fluxo dos dados ao longo de um processo em um sistema;
- *Entity-Relationship Diagram (ERD)*: descrevem entidades inter-relacionadas e suas propriedades para modelar e projetar bancos de dados;
- *Flowcharts*: usados em disciplinas como algoritmos e estruturas de dados e projeto de linguagens de programação;
- *Graphs*: usados na teoria dos grafos, inteligência artificial, linguagens formais, autômatos e computabilidade;

Como as representações diagramáticas estão no cerne deste estudo, os modos comuns para alcançar a percepção, permitir a interação e colaboração, são discutidas na Seção 2.4, a seguir.

## 2.4 Acessibilidade de Diagramas

Antes de tudo, as informações disponíveis sob a forma de diagramas precisam ser percebidas pelos usuários cegos. Estar apto a extrair e assimilar as informações de um diagrama pode ser considerado o nível de entrada na acessibilidade desse tipo de representação visual. Assim, para serem perceptíveis aos cegos, as informações diagramáticas

são tipicamente representadas com mídias alternativas, tais como áudio ou háptico ou ambos (multimodal).

#### 2.4.1 Interfaces Audíveis

Nas interfaces audíveis, pode-se usar as ferramentas de TTS disponíveis em várias bibliotecas padrão nas plataformas de desenvolvimento, que também podem ser obtidas como um componente *off-the-shelf*. Por exemplo, a *Web Audio API* é uma biblioteca disponível nos principais navegadores modernos que permite aos desenvolvedores projetarem aplicativos com saída de fala programada (MACHADO; CONFORTO; SANTAROSA, 2017). O mesmo é possível com bibliotecas externas, por exemplo, como a *FreeTTS*<sup>9</sup> usada por Brown, Pettifer e Stevens (2004) para o desenvolvimento de um navegador não-visual de diagramas moleculares. Desse modo, as ferramentas de síntese de fala são usadas para narrar as informações representadas pelo diagrama. Contudo, além de fala, existem outros modos de comunicar informações por meio de recursos audíveis como, por exemplo, ícones auditivos e *earcons*.

#### 2.4.2 Interfaces Hápticas

As interfaces hápticas são usadas para a acessibilidade de diagramas por meio de métodos manuais, tais como modelos construídos com madeira ou EVA, que são conhecidos como métodos *low-tech* (ou baixa tecnologia). Esses modelos são relativamente simples e baratos de construir se comparados com as opções hápticas *hi-tech*, que incluem impressoras Braille, de relevo ou 3D e dispositivos cinestéticos como o *Phantom Omni Device* (Fig. 2.4) e *displays* Braille com recursos de *force feedback*. Com a exceção dos dispositivos Braille atualizáveis, a limitação mais proeminente das soluções hápticas é a representação estática (BENNETT, 1999; BALIK, 2014), isto é, não é possível reagir às mudanças no diagrama sem exigir reimpressões ou a reestruturação dos modelos manuais.

---

<sup>9</sup>Disponível para a plataforma Java.

### 2.4.3 Soluções Mistas

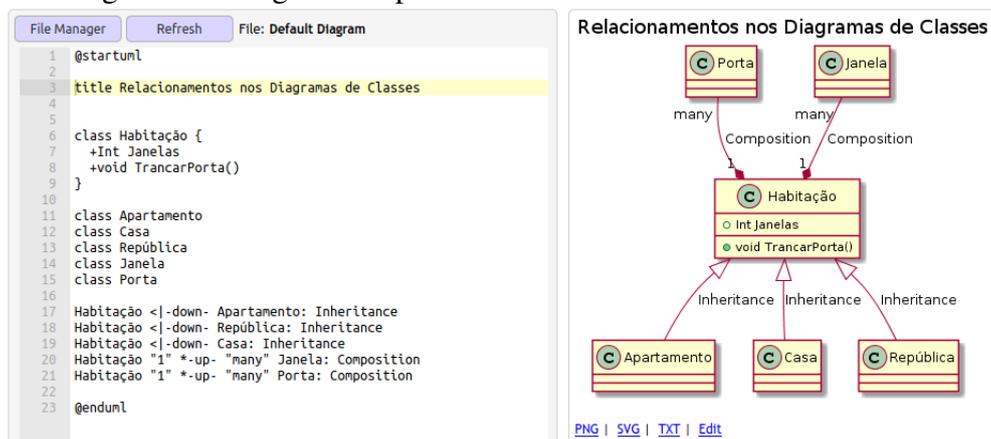
Existem estudos que combinam tecnologias audíveis e hápticas para criar soluções multimodais. Por exemplo, pode-se complementar a informação audível com a posição bidimensional dos elementos por meio de recursos hápticos. A informação da posição no espaço é um dos pontos fracos das soluções puramente baseadas em áudio (METATLA et al., 2012; MOLL; HUANG; SALLNäs, 2010).

### 2.4.4 Interação em Diagramas

Com respeito à interação em diagramas, as soluções para navegação podem ser tanto diretas ou indiretas. Por indireta, primeiro, significa que a representação diagramática é convertida para outro formato melhor suportado, tipicamente textual (SILVA; PANSANATO; FABRI, 2010; LUQUE et al., 2014; LUQUE; BRANDÃO; BRANDÃO, 2016). Desse modo, as ferramentas de TTS podem ser usadas junto com um teclado, para navegar e alterar o conteúdo textual de acordo com a sintaxe, que depois pode ser convertida de volta à forma diagramática gráfica (LUQUE et al., 2014; LOITSCH et al., 2018). No entanto, algumas ferramentas não convertem a representação inicial, permitindo a navegação e edição diretamente sobre a representação diagramática (PANSANATO et al., 2012; SANTOS et al., 2012).

Existem ferramentas para diagramas diretamente baseadas em texto, como a *Plant-Text* que produz diagramas *PlantUML* (Fig. 2.7). No estudo de Luque et al. (2014) há uma lista de ferramentas *web* que geram diagramas UML a partir de descrições textuais.

Figura 2.7: Diagramas a partir de texto com a sintaxe da *PlantText*.



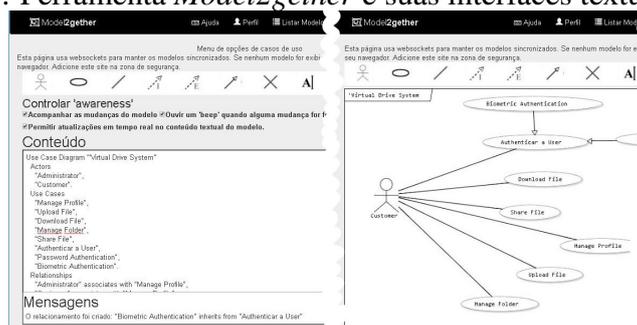
Fonte: autor via PlantText disponível em <<https://www.planttext.com/>>.

## 2.4.5 Recursos Colaborativos

Além da percepção e interação, outra questão importante na acessibilidade de diagramas são os recursos colaborativos, isto é, permitir que pessoas cegas e videntes possam trabalhar juntas nas atividades de modelagem. Esse problema é particularmente importante na área da educação, de modo que os estudantes cegos possam colaborar com outros colegas, também cegos ou videntes, em atividades práticas. Existem estudos que abordam essa questão com duas representações, uma intuitiva para usuários videntes (que arrastam os elementos e os conectam visualmente) e outra acessível aos usuários cegos. Outros estudos procuram encontrar uma representação comum que pode ser compartilhada e utilizável para ambos cegos e videntes.

As ferramentas para modelagem de diagramas que oferecem recursos para colaboração com duas representações, tipicamente o fazem com duas interfaces, uma gráfica e outra textual. Essa é a solução empregada por Luque, Brandão e Brandão (2016) para um editor de diagramas UML baseado na *web* (Fig. 2.8) que permite às pessoas cegas e videntes trabalharem no mesmo diagrama ao mesmo tempo. Grillo (2014) introduz a mesma abordagem, também para o desenho de diagramas UML, por meio de interfaces gráfica e textual, chamadas de “*duas visões*” em seu estudo, também com o objetivo permitir que pessoas com e sem deficiência visual trabalhem juntas.

Figura 2.8: Ferramenta *Model2gether* e suas interfaces textual e gráfica.



Fonte: Adaptado de Luque, Brandão e Brandão (2016)

Em outros estudos, a colaboração é tratada com interfaces multimodais, como as que combinam entradas e saídas hápticas e auditivas. Por exemplo, Bornschein e Prescher (2016) providenciaram interfaces visuais e não-visuais para construir gráficos combinando recursos auditivos e de retorno tátil. Um abordagem semelhante foi usada por Metatla et al. (2012), por meio de uma estação de trabalho equipada com visão gráfica em conjunto com o dispositivo cinestético Phantom Omni (Fig. 2.9) para permitir a colaboração entre usuários cegos e videntes.

Figura 2.9: *Cross-modal* com ferramentas visuais e hápticas juntas.



Fonte: Metatla et al. (2012)

## 2.5 Revisão Sistemática

Esta seção apresenta uma revisão sistemática da literatura com a meta de obter um panorama dos estudos primários conduzidos e revelar as abordagens recentes sendo empregadas para tornar diagramas acessíveis às pessoas cegas e verificando, adicionalmente, se recursos para interação e colaboração são considerados.

### 2.5.1 Método de Revisão Sistemática

Nos primeiros estágios deste estudo, embora houvesse um tópico de interesse, não havia um problema específico sobre acessibilidade de diagramas a explorar. Logo, o primeiro passo foi delinear um escopo para a estudo, reunir os avanços recentes e procurar lacunas na pesquisa. Para isso, foi empregado um método disciplinado para coletar e analisar os dados conhecido como revisão sistemática.

As Revisões Sistemática da Literatura (RSL) são instrumentos para identificar, avaliar e interpretar a pesquisa relevante acerca de questões específicas. As pesquisas coletadas são um conjunto de estudos primários, enquanto as próprias revisões são consideradas estudos secundários (KITCHENHAM, 2004).

RSLs são diferentes das revisões tradicionais pelo modo que são conduzidas, respondendo à questões bem definidas por meio de processos sistemáticos de seleção, filtro e síntese, submetendo os estudos aos critérios de qualidade e elegibilidade pré-definidos com o objetivo de eliminar ou pelo menos mitigar vieses (KYSH, 2013).

A revisão sistemática conduzida nessa fase, especificamente, adere às diretrizes propostas por Kitchenham e Charters (2007), separadas nas fases de planejamento, condução e reportagem. Portanto, começou por definir as questões da pesquisa, os termos para a busca, critérios de inclusão e exclusão e as bibliotecas digitais as quais seriam sub-

metidas as buscas. Durante a revisão, os estudos foram coletados, filtrados segundo os critérios, sintetizados e categorizados. No fim, as descobertas são reportadas, com destaque às respostas para as questões definidas, as lacunas e tópicos que poderiam precisar de investigação adicional.

### 2.5.2 Questões da Revisão Sistemática

As questões propostas nessa revisão foram projetadas para reunir os estudos que usam ou propõem abordagens específicas para adaptar as informações contidas nas representações diagramáticas para pessoas cegas considerando três dimensões: (1) como os cegos podem perceber as informações, (2) interagir com o diagrama (isto é, navegar e/ou editar) e (3) colaborar com outras pessoas. Portanto, resultou nas seguintes Questões da Revisão (QR):

- *QR1*: que abordagem para o cego *perceber* o diagrama?
- *QR2*: que abordagem para o cego *interagir* com o diagrama?
- *QR3*: que abordagem para o cego *colaborar* com outros?

A primeira questão (QR1) pretende reunir as abordagens desenvolvidas ou usadas por pesquisadores para tornar as representações diagramáticas acessíveis, isto é, permitir que pessoas cegas possam percebê-las e compreender a informação codificada que o diagrama pretende comunicar. A segunda (QR2) pretende reunir as abordagens usadas para a acessibilidade dos recursos de interação, isto é, que permitem ao cego navegar pelos componentes do diagrama e até mesmo alterá-lo. A terceira e última questão (QR3) busca relevar se as pesquisas recentes têm considerado as questões e recursos de colaboração ou cooperação.

### 2.5.3 Strings de Busca

As *strings* de busca foram escolhidas de acordo com sua aderência às questões da revisão e projetadas para serem amplas, destacando-se as palavras “cego” e “diagrama”<sup>10</sup>. Nesse protocolo e com base nesses termos de busca elementares, foram definidos os sinônimos e termos auxiliares que podem melhorar os resultados. Além disso, foi verificado

---

<sup>10</sup>Submetidas como “*blind*” e “*diagram*” às bases de dados internacionais

se os motores de busca das bibliotecas digitais suportavam o recurso de *stemming*, que permitiria a entrada apenas da palavra “diagrama” sem suas flexões (ex: o seu plural), que por sua vez gerou as *strings* de busca iniciais.

Após as buscas preliminares, percebeu-se que o termo “cego” (“*blind*” em inglês) recuperava estudos irrelevantes, devido aos artigos mencionarem esta palavra enquanto método de pesquisa, como “método cego” ou “duplo-cego”. Por este motivo, os sufixos “usuário” (“*user*”), “pessoa” (“*person*”) e “pessoas” (“*people*”) foram acrescentados ao termo principal. Além desses, também foi incluída a expressão “deficiente visual” (“*visually impaired*”), pois alguns estudos usam essa expressão como referência às pessoas cegas (ORTIZ; ARISTIZABAL; CARABALLO, 2016).

As *strings* de busca foram aplicadas ao título, resumo e palavras-chave, sempre que permitido pelo motor de busca, para evitar buscas desnecessárias no texto inteiro<sup>11</sup> que poderiam aumentar o número de falsos positivos. Ademais, embora as *strings* tenham sido projetadas para serem amplas, que resultou em uma quantidade considerável de resultados, ainda havia a possibilidade de alguns estudos relevantes não terem sido contemplados, dadas algumas especificidades que são discutidas na seção de Ameaças à Validade (2.5.8).

#### 2.5.4 Bibliotecas Digitais

As bibliotecas digitais, ou bases dados, foram selecionadas com prioridade àquelas que são conhecidas por terem publicações nas áreas de Tecnologia Assistiva e Educação Especial, como seguem:

- *ACM Digital Library* (ACM)
- *Science Direct* (Elsevier)
- *IEEE Xplore* (IEEE)
- *Scopus*
- *Springer Link* (Springer)
- *Web of Science* (WoS)

As *strings* de busca adaptadas para essas bibliotecas digitais e consolidadas podem ser vistas na Tabela 2.1.

---

<sup>11</sup>*full-text search*

Tabela 2.1: *Strings* de busca submetidas às bibliotecas digitais.

| <b>Bib.Digital</b> | <b>String de busca</b>  |
|--------------------|---|
| <i>ACM</i>         | recordAbstract:(+diagram* +"blind people")<br>recordAbstract:(+diagram* +"blind person")<br>recordAbstract:(+diagram* +"blind user")<br>recordAbstract:(+diagram* +"visually impaired") |
| <i>Elsevier</i>    | Title, abstract, keywords: ("diagram") AND<br>("visually impaired"OR "blind people"OR "blind<br>person"OR "blind user")   |
| <i>IEEE</i>        | (("Abstract": "blind people"OR "Abstract": "blind<br>person"OR "Abstract": "blind user"OR<br>"Abstract": "visually impaired") AND<br>"Abstract": diagram*)                              |
| <i>Scopus</i>      | TITLE-ABS-KEY(("blind people"OR "blind person"OR<br>"blind user"OR "visually impaired") AND "diagram")  |
| <i>Springer</i>    | ("blind+people"OR "blind user"OR "blind person"OR<br>"visually impaired") AND ("diagram*")  |
| <i>WoS</i>         | TS(("blind people"OR "blind person"OR "blind<br>user"OR "visually impaired") AND diagram*)  |

### 2.5.5 Critérios de Inclusão e Exclusão

Os estudos foram filtrados em três fases de acordo com critérios de inclusão e exclusão preestabelecidos. Na primeira, os resumos foram submetidos às regras de inclusão listadas a seguir:

- Estudos publicados entre 2013 e 2018;
- Estudos os quais constassem os termos de busca no título, palavras-chave ou resumo.

O objetivo foi coletar um número considerável de artigos que poderiam ser filtrados nas fases seguintes. As quantidades exatas de resumos encontrados por biblioteca digital e que atendiam os critérios de inclusão podem ser vistas na Tabela 2.2.

Tabela 2.2: Artigos atendidos pelos critérios de inclusão.

| <b>Biblioteca Digital</b> | <b>Resultados</b> |
|---------------------------|-------------------|
| <i>ACM DL</i>             | 13                |
| <i>Science Direct</i>     | 1                 |
| <i>IEEE Xplore</i>        | 12                |
| <i>Scopus</i>             | 69                |
| <i>Springer</i>           | 239               |
| <i>Web of Science</i>     | 35                |

Na segunda fase, os resumos dos estudos selecionados foram lidos para, então, submeter aos seguintes critérios de exclusão:

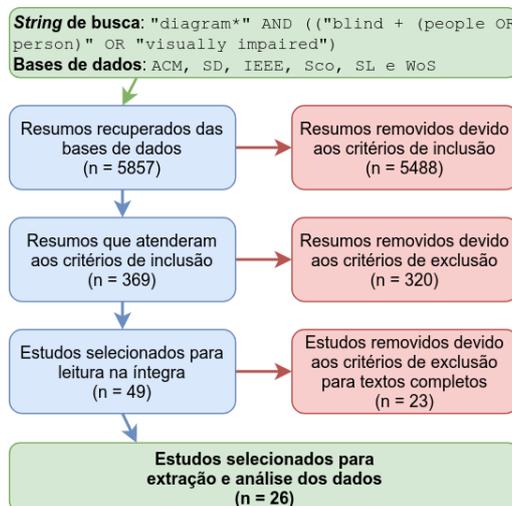
1. Estudos duplicados;
2. Artigos resumidos;
3. Estudos sem relação com cegos ou deficientes visuais;
4. Estudos sem relação com a acessibilidade de representações diagramáticas;
5. Estudos que propunham métodos manuais;
6. Literatura cinzenta;
7. Estudos secundários;

Após a aplicação dos critérios acima enumerados, a lista foi estreitada para 49 estudos, que foram lidos na íntegra e submetidos aos últimos dois critérios de elegibilidade listados a seguir:

8. Estudos que não propõem ou avaliam uma tecnologia assistiva destinada às pessoas cegas ou deficientes visuais enquanto usuários finais;
9. Estudos redundantes que, mesmo em artigos diferentes, discorrem exatamente sobre a mesma prática ou solução.

As três fases de filtragem encurtaram a lista para 26 estudos, que foram usados, portanto, para responder às questões previamente definidas nessa RSL. Os resultados são detalhados e discutidos ao fim desta seção. O processo completo de filtragem pode ser visto na Figura 2.10.

Figura 2.10: Diagrama do processo de filtragem.



### 2.5.6 Avaliação da Qualidade

Os estudos foram submetidos às seguintes questões para avaliar sua qualidade:

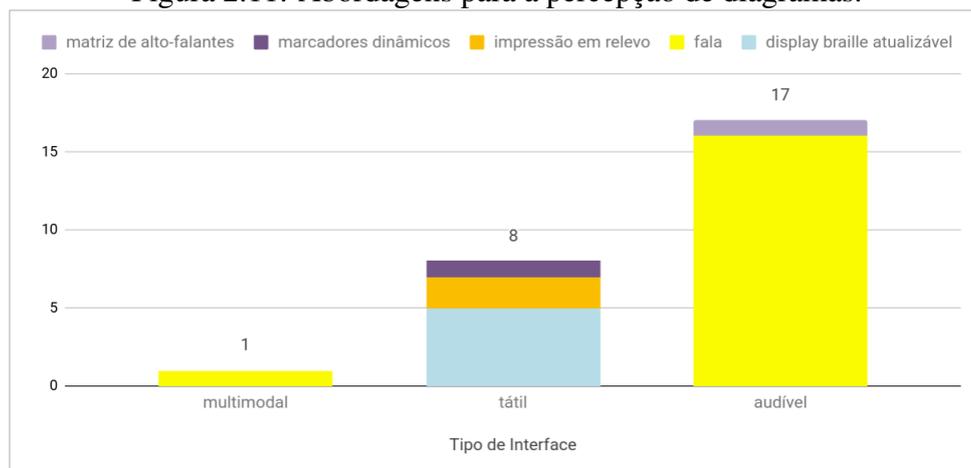
- O estudo utilizou métodos empíricos? Isto é, houveram usuários envolvidos na avaliação? Se sim, foi experimental ou observacional?
- Houve alguma evidência? Se sim, foi quantitativa, qualitativa ou anedotal?

Os métodos empíricos, principalmente os experimentais, são considerados as melhores abordagens para a eficácia de uma intervenção, inclusive para a pesquisa e desenvolvimento de Tecnologia Assistiva (OKOLO; BOUCK, 2007). No entanto, nesta RSL, optou-se por não *rankear* ou excluir os estudos que não se baseiam em métodos empíricos. Boa parte dos estudos são do campo da Educação Especial, não se encaixam nas “*hard sciences*”, e exibem variabilidade de participantes, baixa prevalência e representatividade (ODOM et al., 2005). Isso exige, por exemplo, uma leniência acerca da avaliação dos métodos de pesquisa, para não ignorar nenhuma ideia em potencial. Portanto, os dados extraídos sobre a qualidade foram usados apenas para traçar uma visão geral dos métodos de pesquisa em uso nos estudos recentes.

### 2.5.7 Resultados da Revisão Sistemática

Quanto a percepção de diagramas, primeira questão desta revisão sistemática (QR1), a maioria dos estudos, 17 de 26 (65%), utilizou ou projetou uma interface audível com o usuário. Desses 17, 16 eram baseados em fala e ofereciam uma saída narrativa por meio de TTS, isto é, 94% dos baseados em áudio e 61% do total. Apenas um foi projetado sobre uma interface não-verbal — 5% dos baseados em áudio. Os *displays* Braille atualizáveis foram a segunda abordagem mais usada, com cinco ocorrências (19% do total). Ainda haviam outras quatro soluções hápticas, com base na impressão em relevo e outros mecanismos táteis personalizados. A Figura 2.11 apresenta a síntese dessa questão.

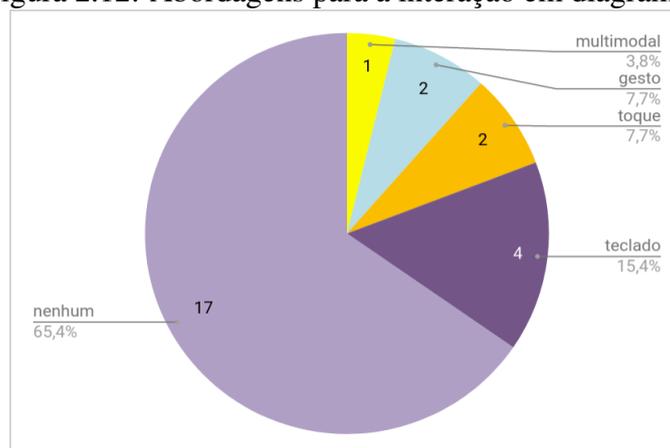
Figura 2.11: Abordagens para a percepção de diagramas.



Fonte: autor

Sobre a interação em diagramas e respondendo à segunda questão (QR2), 17 dos 26 estudos não implementaram nenhum recurso para a interação e preocuparam-se apenas com a percepção das representações diagramáticas (Fig. 2.12). Nos casos em que a interação foi considerada, em cinco de nove estudos o teclado foi o recurso de entrada principal. A entrada por gestos foi a segunda interface mais comum, capturados via tátil em dois e por captura e processamento de imagem em outros dois estudos, no total de quatro artigos. É importante destacar que síntese de fala e teclado são as interfaces mais frequentes pelas quais os cegos operam computadores.

Figura 2.12: Abordagens para a interação em diagramas.

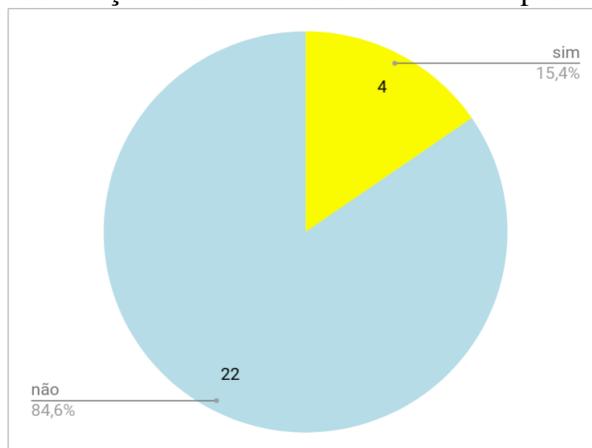


Fonte: autor

Quanto a terceira questão (QR3), apenas quatro estudos apresentaram recursos colaborativos, dois com base no mesmo tempo e lugar, também conhecido como face-a-face ou local, e outros dois ao mesmo tempo em lugares diferentes, também conhecido como interação síncrona distribuída (JOHANSEN, 1988). Portanto, como pode ser visto

na Figura 2.13, 22 dos 26 estudos (84%, ou 5 de 6) não levaram em consideração qualquer recurso colaborativo.

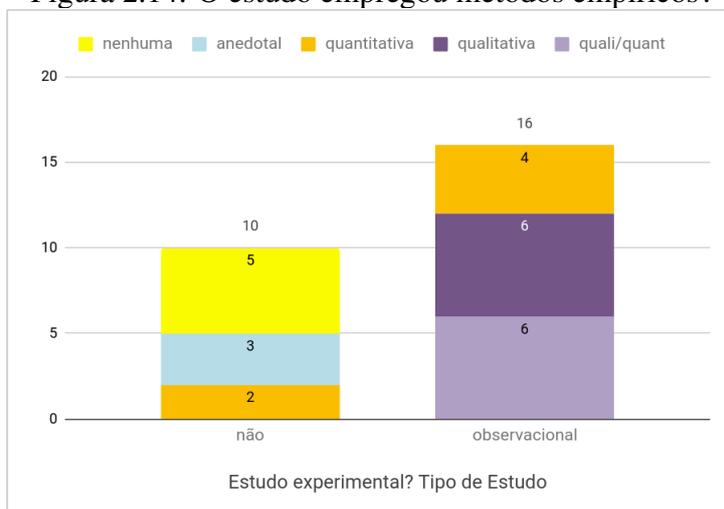
Figura 2.13: A solução considerou funcionalidades para colaboração?



Fonte: autor

Quanto a qualidade dos estudos, como pode ser visto na Figura 2.14, os resultados mostraram que a maioria utilizou métodos empíricos (16 dos 26) conhecidos como estudos observacionais, e demonstraram um equilíbrio entre métodos quantitativos e qualitativos, com 10 e 12 ocorrências, respectivamente. Nenhum estudo foi categorizado como pesquisa experimental, isto é, não utilizou um grupo de controle ou realizou atribuições aleatórias. Daqueles que não foram empíricos (10 dos 26), cinco não apresentaram evidências e três descreveram os resultados informalmente e foram classificados como evidência anedotal.

Figura 2.14: O estudo empregou métodos empíricos?



Fonte: autor

### 2.5.8 Ameaças à Validade

Foram identificadas duas ameaças à validade desta RSL, uma em relação à acurácia do critério de inclusão e outra a respeito da extração e avaliação dos dados.

A primeira é quanto aos resultados obtidos após a RSL em comparação com resultados anteriores. Nota-se que alguns estudos podem não ter sido incluídos por não mencionarem as palavras “diagrama” ou “cego” no título, resumo ou palavras-chave do artigo. Por exemplo, um estudo que mencione “*UML accessibility*” e não “*UML diagrams accessibility*” não seria encontrado com base nos termos e *strings* de busca definidos na preparação desta RSL.

A segunda é sobre a análise, pois Kitchenham (2004) recomenda que outros pesquisadores e especialistas avaliem e discutam os resultados com o objetivo de evitar ou mitigar vieses. Essa recomendação não foi cumprida na condução desta RSL.

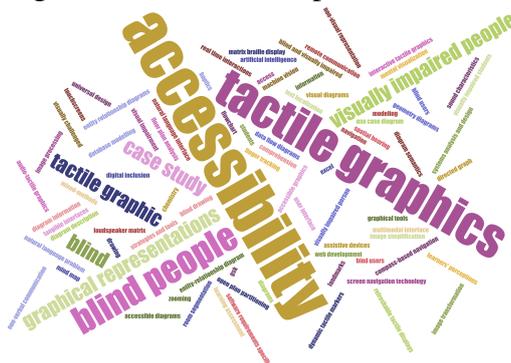
### 2.5.9 Considerações

Segundo os resultados obtidos, predomina o uso de interface baseadas em áudio, principalmente a síntese de fala. A descrição textual de conteúdo diagramático é simples para servir de entrada aos leitores de tela, inerentemente acessível e amplamente empregado (OWEN; COBURN; CASTOR, 2014). A maioria dos estudos encontrados nesta RSL preocupavam-se com um domínio específico ou restrito, isto é, não propunham uma solução genérica e adaptável para qualquer área do conhecimento.

Apesar da praticidade das interfaces audíveis, alguns apresentam alternativas, por exemplo, com o destaque da palavra-chave “*tactile graphics*” (Figura 2.15), que pode ser uma abordagem promissora para a acessibilidade de gráficos. As interfaces auditivas, que estão razoavelmente estabelecidas, não destacam-se nas palavras-chave, embora a maioria dos artigos a tenham utilizado.

Por fim, observou-se a escassez de estudos sobre recursos colaborativos acessíveis, tema discutido na seção seguinte.

Figura 2.15: Nuvem de palavras-chave.



Fonte: autor.

## 2.6 Aplicações Colaborativas

Esta seção aborda o trabalho e aprendizado colaborativo apoiado por sistemas computadorizados, os recursos para habilitá-los e as implicações pedagógicas.

### 2.6.1 Groupware

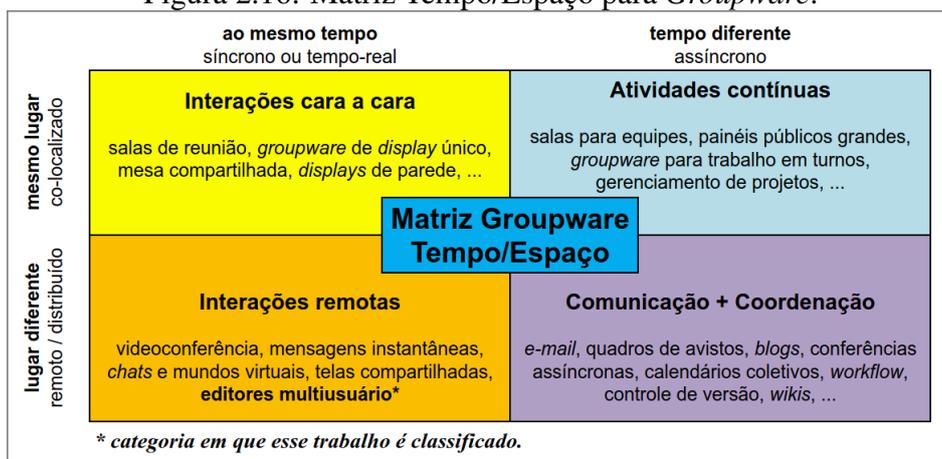
*Groupware* é definido por Ellis, Gibbs e Rein (1991, p. 40) como “*sistemas baseados em computadores que apoiam grupos de pessoas engajadas numa tarefa (ou meta) comum e que fornece uma interface para um ambiente compartilhado*”<sup>12</sup>, também conhecido como Software Colaborativo, tem o objetivo de apoiar grupos na comunicação, colaboração e coordenação de atividades.

As aplicações de *Groupware* podem ser classificadas de acordo com a matriz de Tempo/Espaço proposta por Johansen (1988) e que pode ser vista na Figura 2.16. A dimensão do tempo também é conhecida como colaboração síncrona ou assíncrona, respectivamente, enquanto a dimensão do espaço (ou lugar) também é conhecida como co-localizada ou remota<sup>13</sup> (RAMA; BISHOP, 2006).

Os fundamentos, definições e outras questões de *Groupware* se sobrepõem com a área de *Computer-Supported Cooperative Work (CSCW)*, que estuda os efeitos psicológicos, sociais e organizacionais além das ferramentas (SCHMIDT; BANNON, 1992) e é abordado na seção a seguir.

<sup>12</sup>Traduzido pelo autor. Texto original: “*computer-based systems that support groups of people engaged in a common task (or goal) and that provide an interface to a shared environment*”

<sup>13</sup>Outros autores, como Gutwin e Greenberg (1998), chamam o “*Groupware síncrono remoto*” de “*Groupware distribuído em tempo real*”

Figura 2.16: Matriz Tempo/Espaço para *Groupware*.

Fonte: adaptado de Johansen (1988).

### 2.6.2 Computer-Supported Cooperative Work

O campo do *Computer-Supported Cooperative Work*<sup>14</sup> (CSCW) emergiu do esforço conjunto de economistas, psicólogos sociais, antropologistas e educadores (GRUDIN, 1994). Os estudos de CSCW podem ou ser centrados na tecnologia ou no trabalho em si. Segundo Mills (2003), as pesquisas voltadas ao projeto de funcionalidades se concentram em cinco áreas: comunicação, configuração, coordenação, acesso à informação e interação, as quais possuem subcategorias que podem ser observadas na Tabela 2.3.

Tabela 2.3: As 5 áreas de projeto de CSCW e suas características.

| Área de Projeto     | Características-chave   |
|---------------------|---|
| Comunicação         | Assíncrono, Áudio, Dados, Privado, Compartilhado, Estruturado, Síncrono, Texto, Desestruturado, Vídeo                                       |
| Configuração        | Adaptação, Composição, Evolução, Extensão   |
| Coordenação         | Controle de Acesso, Concorrência, Consistência, Delegação, Agendamento, Versionamento   |
| Acesso à Informação | Distribuição, Filtragem, Recuperação, Estrutura   |
| <b>Interação</b>    | <b>Gerenciamento da Atenção, Consciência (<i>Awareness</i>), Gerenciamento de Contexto, Estabelecimento e Manutenção de Relacionamentos</b> |

Fonte: adaptado de Mills (2003). Grifo do autor.

### 2.6.3 Computer-Supported Collaborative Learning

Segundo Stahl, Koschmann e Suthers (2006), *Computer-Supported Collaborative Learning*<sup>15</sup> (CSCL) é um campo de estudo da criação de significado no contexto

<sup>14</sup>Trabalho Colaborativo Apoiado por Computador em português

<sup>15</sup>Aprendizagem Colaborativa Apoiada por Computador em Português

de atividades conjuntas e os modos como essas práticas são mediadas por sistemas computadorizados. A área beneficia-se dos desenvolvimentos do campo do CSCW (LIPONEN, 2002) e tem influência nas atividades de sala de aula, físicas e virtuais, facilmente mesclando-se com ideias de *E-Learning* (educação a distância) (STAHL; KOSCHMANN; SUTHERS, 2006).

#### **2.6.4 Fundamentos Pedagógicos**

De acordo com Springer, Stanne e Donovan (1999), não há uma raiz teórica única para a (A) aprendizagem colaborativa ou cooperativa, sendo que as atividades de aprendizagem em grupo emergiram de vários campos, como na filosofia da educação e psicologia cognitiva e social. Para Davidson e Major (2014), os teóricos Dewey, Piaget e Vygotsky influenciaram ambos métodos de aprendizagem.

Oxford (1997a) declara que, nos desenvolvimentos iniciais, a visão de Dewey era de que os estudantes não aprendiam isoladamente; em vez disso, o indivíduo aprende por estar envolto em uma comunidade e no mundo como um todo. E Johnson, Johnson e Smith (1998) declara que, na visão de Piaget, ocorrem (saudáveis) conflitos sócio-cognitivos quando os indivíduos cooperam em um ambiente, e que isso causa um desequilíbrio que estimula a construção de novas estruturas conceituais, entendimento e conhecimento.

Contudo, foi Lev Vygotsky que contribuiu significativamente para a epistemologia social interacionista, reconheceu que as ideias têm origens sociais, que são construídas por meio da comunicação com outros e que o sistema cognitivo de um indivíduo é o resultado da comunicação em grupos e não pode ser separado da vida social, apresentou o conceito de (B) Zona de Desenvolvimento Proximal e conduziu vários estudos na área de (C) defectologia (VYGOTSKY, 1980).

Os assuntos enumerados nos parágrafos anteriores são detalhados nos tópicos a seguir.

##### *A) Aprendizagem Colaborativa*

De acordo com McInnerney e Roberts (2004), as instituições educacionais lideram um movimento na direção de aulas dinâmicas e com participação e engajamento dos estudantes. Nesse sentido, os professores estão para facilitadores ou guias, que conduzem

os estudantes através de uma experiência de aprendizado, em vez de mestres ou sábios que detêm o conhecimento e são responsáveis por transmiti-los aos alunos (STOLL, 1996). Segundo os autores, em vez de os estudantes ouvirem as informações passivamente, eles se destacam pelo protagonismo nos assuntos de sala de aula, considerando o professor mais como um recurso do que uma fonte.

Existem diferentes visões sobre a definição e como se conduz uma aprendizagem colaborativa. Nas palavras de Smith e MacGregor (1992, p.1), “‘*Aprendizagem Colaborativa*’ é um termo guarda-chuva para variedade de abordagens educacionais envolvendo o esforço intelectual conjunto de estudantes, ou estudantes e professores.”<sup>16</sup>. Para Dillenbourg (1999) e Laal e Laal (2012), há uma falta de consenso sobre o termo devido sua menção de acordo com cada campo acadêmico. AC pode significar, por exemplo, o material do curso e compartilhamento de atividades, assim como pode também significar o esforço conjunto para resolver um problema.

Laal e Ghodsi (2012) apresentam alguns benefícios da aprendizagem colaborativa em sua revisão sistemática, tais como benefícios sociais, psicológicos, avaliativos e educacionais, bem como a promoção da diversidade, o aumento da autoestima e a promoção de habilidade de pensamento crítico. Gokhale (1995) também destaca o pensamento crítico e reporta um aumento significativo das notas daqueles que estudavam colaborativamente comparados àqueles que estudavam sozinhos. Springer, Stanne e Donovan (1999), em sua meta-análise, reportaram resultados positivamente significativos, como mais conquistas acadêmicas, aumento da permanência e atitudes mais favoráveis.

De acordo com Oxford (1997b), a aprendizagem cooperativa é considerada mais estruturada e prescritiva, o professor mantém uma posição de autoridade e é responsável por gerenciar as interações mais de perto, enquanto a aprendizagem colaborativa é mais centrada no aprendiz e, de acordo com Panitz (1999), adepta às filosofias sócio construtivista e interacionista. Além disso, os grupos colaborativos são auto-gerenciados, menos estruturados, recebem menos instruções e são avaliados qualitativamente pelos professores. Enquanto na aprendizagem cooperativa o grupo pode ou não chegar à uma resposta aceitável, os grupos colaborativos podem chegar à uma solução conflitante com a dos outros grupos e até mesmo com as expectativas do professor.

Os métodos de aprendizagem cooperativa e colaborativa tiveram raízes diferentes. Enquanto a primeira emergiu da prática na sala de aula de escolas de educação básica a segunda foi aplicada ao pós-secundário, ensino superior e ensino de adultos, nas quais

<sup>16</sup>Tradução do autor para “‘*Collaborative Learning*’ is an umbrella term for a variety of educational approaches involving a joint intellectual effort by students, or students and teachers together”

era preciso menos estrutura e direcionamento do que no ensino de jovens e crianças (OXFORD, 1997b). Segundo Bruffee (1995), ambos métodos possuem a mesma meta a longo prazo e, além disso, o método colaborativo usado nas faculdades e universidades pode ser visto como um complemento do método cooperativo empregado nas escolas de ensino básico<sup>17</sup>. A importância e teorias pedagógicas que apoiam a aprendizagem em grupo são discutidas a seguir.

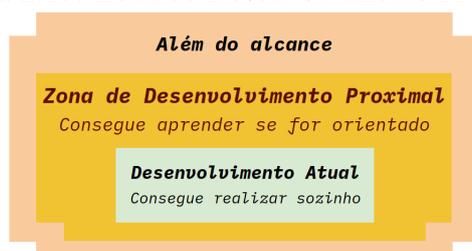
### B) Zona de Desenvolvimento Proximal de Vygotsky

Existem dois princípios que suportam a teoria de desenvolvimento cognitivo de Vygotsky: o (1) Outro Mais Experiente (OME) e a (2) Zona de Desenvolvimento Proximal (ZDP). Ambos estabelecem suporte pedagógico para as práticas da aprendizagem colaborativa (VYGOTSKY, 1980). Além dos professores, qualquer indivíduo com mais conhecimento ou habilidade pode ser considerado um OME, como os pais, colegas de aula, colegas de trabalho e outras pessoas, os quais representam uma fonte em potencial para a troca de experiências de aprendizado, e permitem o nivelamento do conhecimento e das habilidades entre pares desde que elas estejam em suas ZDPs.

Vygotsky propôs que as pessoas possuem um nível de desenvolvimento real e que há um potencial nível imediato de maior desenvolvimento, que é alcançável com a orientação e encorajamento de um OME, isto é, auxiliado por um parceiro mais habilidoso. Esse passo único na direção do incremento de conhecimento ou habilidade é conhecido como Zona de Desenvolvimento Proximal (Fig. 2.17). ZDP é definida com as palavras de Vygotsky como:

a distância entre o nível real de desenvolvimento, determinado pela capacidade de resolver independentemente um problema e o nível de desenvolvimento potencial, determinado através da resolução de um problema sob a orientação de um adulto ou em colaboração com outro companheiro mais capaz (VYGOTSKY, 1980, p.86).

Figura 2.17: Zona de Desenvolvimento Proximal.



Fonte: autor

<sup>17</sup>Bruffee (1995, p.12) coloca as seguintes questões: “Existe realmente uma diferença entre os dois? Se sim, qual seria, e ela realmente importa?”

As teorias de Vygotsky são frequentemente discutidas nos estudos de tanto aprendizagem colaborativa como cooperativa. Por exemplo, Doolittle (1995) discutiu a relação entre a ZDP e o aprendizado cooperativo, que é uma estratégia efetiva de educação formal para apresentar experiências sociais e culturais que apoiam, assim, a ZDP. Oxford (1997a), nos seus estudos sobre aquisição de segunda língua, declara que a aprendizagem colaborativa tem uma base filosófica sócio-construtivista e interacionista que promove comunidades de aprendizagem de modo que os estudantes negociem suas ZDPs para alcançar a proficiência.

### *C) Estudos de Defectologia de Vygotsky*

Vygotsky conduziu vários estudos na área da defectologia, definida como um campo de conhecimento teórico e um trabalho científico-prático sobre as barreiras para os desenvolvimento de crianças com deficiência, em que essas não seriam menos desenvolvida comparada às “normais”, mas sim desenvolvidas de outra forma (VYGOTSKI, 2015). Vygotsky propôs que as pessoas “normais” não deveriam ser usadas como referências para as pessoas com deficiências, da mesma maneira que adultos não são uma referência para o desenvolvimento de crianças (NUERNBERG, 2008).

Quanto a deficiência visual e cegueira, Vygotsky propunha a compensação social, baseada na linguagem e comunicação, sobre a falta do sentido visual. Sua teoria se opunha à noção de compensação biológica, isto é, a amplificação dos demais canais sensoriais como toque e audição. Logo, haviam apenas problemas de mobilidade e orientação, enquanto as funções psicológicas estariam totalmente preservadas (NUERNBERG, 2008). No entanto, a cegueira põe obstáculos à interação social, que resulta em desvios e, desse modo, a colaboração é essencial para desenvolver as funções intelectuais dos sujeitos e um dos fatores centrais para o desenvolvimento cultural dos indivíduos (VYGOTSKI, 2015).

A educação tradicional foi criticada por Vygotski (2015) por impor uma pedagogia individualista baseada em mestre/aluno, pois a educação especial deveria assegurar a colaboração entre as pessoas cegas e videntes, com a metáfora de que o papel dos olhos é como o de uma ferramenta, como um microscópio ou telescópio, Vygotsky disse:

la experiencia de otra persona, la utilización de los ojo ajenos, la colaboración con un vidente — nos encontramos de inmediato en un terreno esencialmente nuevo y el ciego adquiere su microscopio y su telescopio, que amplían inconmensurablemente su experiencia y lo insertan estrechamente en la trama común del mundo (VYGOTSKI, 2015, p.84).

## 2.6.5 Mecanismos para Colaboração

Ao projetar sistemas de *software* ou aplicativos é preciso considerar as implicações sociais, já que os seres humanos são inerentemente sociais. Os mecanismos para isso e suas principais categorias, segundo Preece et al. (1994), são listadas a seguir:

- *Mecanismos conversacionais*: permite e facilita o fluxo de conversas;
- *Mecanismos para coordenação*: permite que as pessoas trabalhem em equipe e interajam umas com as outras;
- *Mecanismos para conscientização (awareness)*: usado para permitir que os indivíduos estejam conscientes do que está acontecendo.

Enquanto mecanismos conversacionais, considera-se os meios de comunicação mediados por computador, entre usuários, que pode ser textual, verbal ou simbólico, ou entre usuários e o sistema, que pode oferecer *feedback* por meio de imagens, sons, vibração e quaisquer outros tipos de interfaces de saída. Entre os exemplos mais conhecidos, na área de tecnologias colaborativas para comunicação, estão o *e-mail*, videoconferência e salas de bate-papo. A comunicação também pode ocorrer em ambientes virtuais, como jogos e editores *on-line*.

Os mecanismos de coordenação são necessários para permitir que os coparticipantes interajam entre eles por meio de regras e agendamento de recursos. Por exemplo, considere um calendário compartilhado de modo que todos os usuários possam adicionar entradas e planejar eventos. Neste caso, o aplicativo colaborativo precisa restringir algumas ações para evitar conflitos, isto é, gerenciar as questões de concorrência.

Dourish e Bellotti (1992, p.1) definem *awareness*<sup>18</sup> como “*um ‘entendimento das atividades dos outros’, que provê um ‘contexto para suas próprias atividades’*”<sup>19</sup>. Segundo Gutwin, Penner e Schneider (2004), o *awareness* está garantido para equipes co-localizadas (ou locais), já que os indivíduos podem se comunicar e observar as mudanças no ambiente e estarem cientes das ações dos colegas. Entretanto, o cenário não é o mesmo na colaboração remota, pois os recursos de comunicação podem ser insuficientes para tratar a coordenação das tarefas. Portanto, análogo a um espaço de trabalho (*workspace*<sup>20</sup>)

<sup>18</sup>O termo *awareness* será preferido neste estudo, em vez de suas contrapartidas em português “consciência” ou “ciência”, por ser menos ambíguo.

<sup>19</sup>Tradução do autor para *an understanding of the activities of others, which provides a context for your own activity*

<sup>20</sup>O termo *workspace* será preferido neste estudo, em vez de suas contrapartidas em português “espaço de trabalho” ou “local de trabalho”.

físico compartilhado, onde as pessoas estão cientes da localização dos outros, o que eles estão desempenhando e quais são suas intenções, o mesmo é necessário para *workspaces* virtuais, para permitir que os usuários acompanhem as atividades da equipe e trabalhem mais eficientemente.

### 2.6.6 Workspace Awareness

A expressão *Workspace Awareness*<sup>21</sup> (WA) foi cunhada por Gutwin e Greenberg (1996, p.208) como “a coleção minuto-a-minuto do conhecimento que uma pessoa dispõe sobre o estado das interações dos outros no workspace”<sup>22</sup>.

O WA é conhecido como uma questão-chave para ajudar as pessoas à trabalharem juntas no campo da CSCW (SÁNCHEZ; BALOIAN, 2006; GUTWIN; GREENBERG, 1996; COLLAZOS et al., 2002; BORGES; PINO, 1999). Segundo Gutwin e Greenberg (1996), WA auxilia a coordenação de tarefas e recursos quando os indivíduos estão trabalhando em atividades compartilhadas, acontece naturalmente nas situações face-a-face, porém não é tão simples nos sistemas de *groupware* em tempo-real em que, por exemplo, as pessoas podem não ter a mesma visão do espaço e não ter a mesma riqueza de comunicação.

Há dois problemas típicos ao projetar o suporte ao WA: (1) quais informações sobre a interação dos outros devem ser coletadas? (2) como apresentar essas informações para os outros participantes? Essas duas questões foram tratadas por Gutwin e Greenberg (1996), que construíram um *framework* para abordar a primeira questão e discutiram mecanismos comuns para abordar a segunda.

Gutwin e Greenberg (2002) desenvolveram um *framework* conceitual para lidar com os desafios comuns no projeto de *groupware* (aplicável também às áreas de CSCW e CSCL), relacionados às informações que devem ser coletadas e apresentadas em sistemas distribuídos em tempo-real. Nesse *framework*, as informações são categorizadas como elementos, cujo conjunto básico responderia às questões nas seguintes categorias: “quem?”, “o quê?”, “onde?”, “quando” e “como?”. Essas categorias abordam elementos específicos de conscientização, que foram divididos temporalmente como parte do presente ou passado e podem ser vistos nas Tabelas 2.4 e 2.5.

<sup>21</sup>“conscientização do espaço de trabalho” em Português, no entanto a expressão em Inglês será preferida no restante deste estudo

<sup>22</sup>Tradução do autor para “the collection of up-to-the-minute knowledge a person holds about the state of another’s interaction with the workspace.”

Tabela 2.4: Elementos de *workspace awareness* relacionados ao presente.

| <b>Categoria</b> | <b>Elemento</b> | <b>Questões específicas</b>            |
|------------------|-----------------|--|
| <i>Quem</i>      | Presença        | Há alguém no <i>workspace</i> ?        |
|                  | Identidade      | Quem está participando? Quem é esse?   |
|                  | Autoria         | Quem está fazendo isso?                |
| <i>O quê</i>     | Ação            | O que eles estão fazendo?              |
|                  | Intenção        | Essa ação pertence à qual objetivo?    |
|                  | Artefato        | Em qual objeto eles estão trabalhando? |
| <i>Onde</i>      | Localização     | Onde eles estão trabalhando?           |
|                  | Olhar           | Onde eles estão olhando?               |
|                  | Visão           | Até onde eles podem ver?               |
|                  | Alcance         | Até onde eles podem alcançar?          |

Fonte: adaptado de Gutwin e Greenberg (2002)

Tabela 2.5: Elementos de *workspace awareness* relacionados ao passado.

| <b>Categoria</b> | <b>Elementos</b>         | <b>Questões específicas</b>             |
|------------------|--------------------------|---|
| Como             | Histórico de ações       | Como essa operação aconteceu?           |
|                  | Histórico de artefatos   | Como esse artefato acabou nesse estado? |
| Quando           | Histórico de eventos     | Quando esse evento ocorreu?             |
| Quem/passado     | Histórico de presença    | Quem esteve aqui e quando?              |
| Onde/passado     | Histórico de localização | Onde uma pessoa esteve?                 |
| Qual/passado     | Histórico de ações       | O quê uma pessoa esteve fazendo?        |

Fonte: adaptado de Gutwin e Greenberg (2002)

Para a questão de como apresentar essa informação aos outros participantes, Gutwin e Greenberg (1996) reuniram da literatura mecanismos comuns de *workspace awareness* e os organizaram como pode ser visto na listagem a seguir:

- *Comunicação direta*: os usuários comunicam suas ações explicitamente;
- *Produções indiretas*: a comunicação acontece por meio das ações, expressões ou fala indireta;
- *Comunicação consequencial*: os usuários observam ou ouvem os efeitos das ações dos outros usuários;
- *Feedthrough*: os usuários observam os efeitos nos artefatos como sendo o resultado da ação de outros usuários;
- *Feedback ambiental*: observa-se os efeitos indiretos das ações dos outros no *workspace* como um todo;

Esses mecanismos providenciam um mapeamento natural, o qual os projetistas de Groupware/CSCW podem considerar na criação de técnicas e *widgets* para transmitir informações de *awareness* em um *workspace* virtual. Além disso, de acordo com Gutwin e Greenberg (1996), seu *framework* conceitual pode prover um vocabulário útil que pode ser usado como ponto de partida para projetar aplicativos colaborativos.

### 3 MATERIAIS E MÉTODOS

Este capítulo apresenta as etapas da pesquisa na direção de um *workspace awareness* acessível, com a intenção de habilitar recursos de colaboração em ferramentas de modelagem de diagramas, permitindo que pessoas cegas interajam e aprendam com outras. Primeiramente, o problema de pesquisa é explicado detalhadamente, seguido pela metodologia de pesquisa e solução proposta. Após, o projeto da interface auditiva é apresentado, abordando a sonificação projetada para cada categoria de *awareness*.

#### 3.1 Problema de Pesquisa

**Considerando um aplicativo de modelagem colaborativa de diagramas, é possível usar uma interface auditiva que, por meio da *sonificação* dos elementos de *workspace awareness*, torne os recursos colaborativos acessíveis aos usuários cegos e permita que eles percebam as ações dos coparticipantes e as mudanças no espaço de trabalho?**

Os sistemas colaborativos tipicamente apresentam informações de *awareness* por meio de interfaces gráficas, isto é, com *widgets* e outras pistas visuais para sinalizar mudanças no *workspace*. Por exemplo, aplicativos de escrita colaborativa *on-line*, como o Google Docs ou Microsoft Word, representam e diferenciam os participantes simultâneos com cursores coloridos que se movimentam no documento e permitem que os membros estejam cientes do que os outros estão editando e em que local do texto.

Existem estudos que usaram, com sucesso, sons para apresentar informações de *awareness* para pessoas videntes (METATLA, 2016; RAJAN et al., 2013); contudo, pistas audíveis para videntes representam mais um complemento do que uma substituição às pistas visuais.

Naturalmente, pessoas com deficiência visual contam com as notificações de atualização auditivas tanto quanto as visuais, enquanto os cegos dependem totalmente do áudio para a percepção, tanto dos artefatos e usuários presentes no *workspace*, como das ações e mudanças no ambiente compartilhado.

Existem vários estudos acerca da acessibilidade para percepção e autoria de diagramas, que reportaram soluções viáveis para permitir que cegos trabalhem com representações diagramáticas, como pode ser visto na Seção 2.5. Entretanto, sem os recursos colaborativos, as pessoas cegas ou com deficiência visual desenvolveriam apenas habili-

dades individuais (SÁNCHEZ; BALOIAN, 2006). Adicionalmente, estudantes cegos não poderiam beneficiar-se do trabalho em grupo que demanda decisões coletivas a respeito do domínio do problema, como as que acontecem na atividade de modelagem coletiva. Naturalmente, a ausência de atividades em grupo prejudica as interações sociais e, logo, o processo de aprendizagem.

Apesar deste estudo abordar teorias pedagógicas para apoiar e justificar a necessidades de recursos colaborativos acessíveis no ferramental de *groupware*, está fora do escopo desta pesquisa avaliar os métodos pedagógicos.

## 3.2 Metodologia

Este estudo foi conduzido durante dois anos. O primeiro foi dedicado à revisão da literatura e identificação do problema, apresentado na revisão sistemática da Seção 2.5 e o segundo foi dedicado ao projeto e desenvolvimento das interfaces auditivas e do protótipo. O último período (2º ano) foi dividido em duas partes (dois semestres) segundo de um método iterativo orientado a resolução de problemas de engenharia chamado *Design Science Research* (Seção 3.2.1). Logo, o protótipo e as interfaces auditivas foram testados em dois momentos. Na primeira iteração (semestre) foi conduzido um teste de usabilidade baseado em taxa de sucesso (Seção 3.2.2), para verificar se as tarefas de percepção do grupo eram objetivamente bem-sucedidas com a interface projetada. Na segunda iteração foram conduzidos novos testes de usabilidade, mais a medição subjetiva da carga de trabalho (Seção 3.2.3) por meio da ferramenta *NASA Task Load Index*. Esses métodos são detalhados nas seções a seguir.

### 3.2.1 Design Science Research

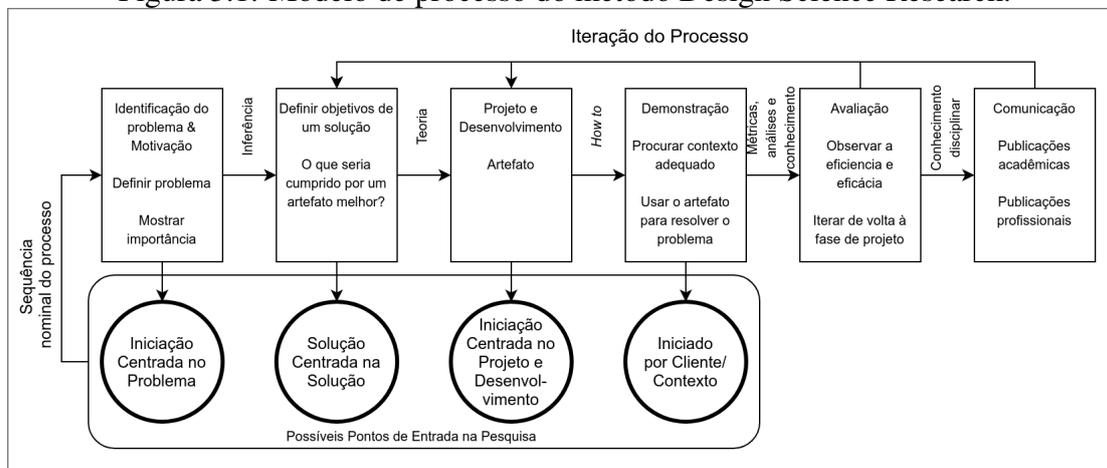
O projeto, desenvolvimento e refinamento das interfaces auditivas propostas é baseado no método *Design Science Research* (DSR), especificado por Peffers et al. (2007). Segundo os autores, DSR é orientado ao projeto e desenvolvimento de artefatos bem-sucedidos e incorpora princípios, práticas e procedimentos. Para Hevner e Chatterjee (2010), o método engloba a ideia de que fazer *design* inovador pode providenciar contribuições evidentes à base de conhecimento, que inclui construtos, modelos, métodos e instâncias.

De acordo com Peffers et al. (2007), DSR é um processo de resolução de problemas orientado à engenharia, de maneira que as disciplinas de engenharia a aceitam como uma metodologia válida por valorizar soluções incrementáveis que sejam efetivas e aplicadas.

O processo de DSR inclui seis passos principais: (1) identificar o problema e a motivação; (2) definir requisitos para uma solução; (3) projetar e desenvolver um artefato que atenda aos requisitos da solução; (4) demonstrar a solução; (5) avaliar a solução; (6) comunicar os resultados (Fig. 3.1). Este estudo, no caso, foi separado em duas iterações descritas a seguir:

- *1ª iteração / projeto da interface auditiva e protótipo*: abrange as diretrizes da interface, o protótipo e sua avaliação, cujos desenvolvimentos podem ser vistos na Seção 3.4 e os resultados na Seção 4.1
- *2ª iteração / refinamento e atualizações*: abrange o reprojeto da interface, o projeto de opções baseadas em fala e as atualizações do protótipo, cujos processos podem ser vistos na Seção 3.5 e os resultados na Seção 4.2.

Figura 3.1: Modelo de processo do método Design Science Research.



Fonte: adaptado de Peffers et al. (2007)

### 3.2.2 Teste de Usabilidade: Taxa de Sucesso

Os testes de usabilidade (também conhecidos como testes com usuários) são métodos empíricos usados para avaliar uma solução na área de Interação Humano-Computador (IHC), conduzido por um facilitador (ou moderador) que solicita aos participantes que realizem determinadas tarefas, enquanto observa e coleta informações. Podem ser quali-

tativos ou quantitativos, enquanto o primeiro foca na revelação dos comportamentos dos usuários enquanto usam o produto, o segundo objetiva coletar métricas que descrevem a experiência do usuário, tais como a taxa de sucesso e o tempo despendido na tarefa.

Segundo Nielsen (2001), a medição da taxa de sucesso do usuário no cumprimento de tarefas é uma métrica quantitativa simples. O teste não fornece dados sobre o motivo das falhas ou dados sobre o desempenho. Contudo, as informações podem ser coletadas sem dificuldade, isto é, sem uma preparação (*setup*) muito complexa do testes, e geram dados estatísticos bastante reveladores. O número de usuários recrutados depende de orçamento, tempo e disponibilidade de sujeitos. Segundo Nielsen (1994), os métodos empíricos de testagem na área de IHC são frequentemente intimidadores dada sua complexidade. Por exemplo, nesta pesquisa, especificamente, notou-se os efeitos da baixa representação populacional durante o recrutamento de sujeitos para o teste. Esse problema também foi observado em outros estudos, como em Thapa, Ferati e Giannoumis (2017) e Freire et al. (2010). Contudo, testes com cinco usuários podem revelar boa parte dos problemas de usabilidade, sendo preferível testes iterativos com pequenos grupos do que o investimento em um único extensivo teste de usabilidade (TURNER; LEWIS; NIELSEN, 2006; NIELSEN, 2012).

### 3.2.3 Medição de Carga de Trabalho: NASA-RTLX

Neste estudo é usada a ferramenta NASA Task Load Index (NASA-TLX) para mensurar a carga de trabalho (*workload*, em inglês) que representa, segundo Hart (1986), o custo para um usuário completar as exigências de uma tarefa e busca quantificar o esforço realizado por um indivíduo para identificar fadiga, *stress* e outros fatores que podem mitigar o desempenho.

NASA-TLX é uma ferramenta para a medição da carga de trabalho subjetiva multidimensional que possui duas partes: (1) separação em seis dimensões, ou subescalas; (2) definição de pesos individuais para cada subescala. Quando a segunda parte é suprimida, o exame é conhecido como *unweighted* (sem pesos), ou NASA Raw Task Load Index (NASA-RTLX). As subescalas, e uma breve explicação de cada, são listadas a seguir:

- *Demanda Mental (baixa/alta)*: atividade mental e perceptual necessária (ex.: pensar, decidir, calcular, lembrar, etc);
- *Demanda Física (baixa/alta)*: atividade física exigida (ex.: empurrar, puxar, pres-



usuários em um sistema de *groupware*. De acordo com os autores, as informações podem estar relacionadas tanto ao instante presente quanto ao passado, no total de 17 elementos (Tabelas 2.4 e 2.5) para uma experiência de *workspace awareness* (mais informações na Seção 2.6.6). Para implementar e avaliar as interfaces auditivas para o *workspace awareness*, foi realizado um recorte dos 17 elementos. A meta é recuperar informações básicas sobre o presente e cobrir a comunicação de três categorias fundamentais (quem? o quê? onde?) por meio de técnicas de *sonificação* em vez de recursos de fala.

A Tabela 3.1 mostra os seis elementos selecionados e as técnicas de *sonificação* para torná-los acessíveis. Cada técnica é construída sobre um mecanismo, que foram selecionados da lista disponível no trabalho de Gutwin e Greenberg (1996), o qual apresenta os mapeamentos naturais considerados neste estudo para habilitar uma colaboração acessível aos cegos. As técnicas para a *sonificação* desses elementos são discutidas nas seções a seguir.

Tabela 3.1: Mecanismos para elementos de *workspace awareness* acessíveis.

| <b>Categoria</b> | <b>Elemento</b> | <b>Mecanismo/Técnica</b>   |
|------------------|-----------------|--|
| Quem             | Presença        | Contemplado por fala indireta e <i>earcons</i>                               |
|                  | Identidade      | Contemplado por fala indireta e comunicação consequencial via <i>earcons</i> |
|                  | Autoria         | Contemplado por comunicação consequencial via <i>earcons</i>                 |
| Qual             | Ação            | Contemplado por comunicação consequencial via <i>earcons</i>                 |
|                  | Intenção        | Não contemplado  |
|                  | Artefato        | Não contemplado  |
| Onde             | Localização     | Contemplado por comunicação consequencial via <i>earcons</i>                 |
|                  | Olhar           | Não contemplado  |
|                  | Visão           | Não contemplado  |
|                  | Alcance         | Não contemplado  |

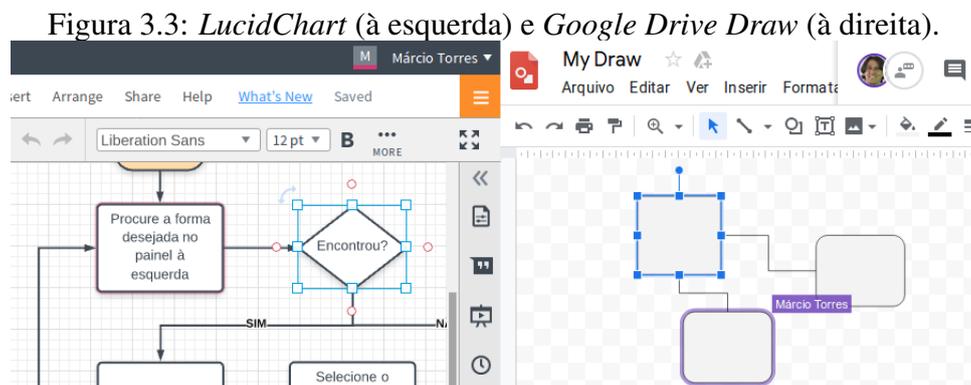
Fonte: autor, adaptado de Gutwin e Greenberg (2002)

### 3.3.1 Awareness de “quem”

A interface auditiva para tornar acessível a ciência de “quem” (“*who*” awareness) foi projetada sobre os mecanismos de fala indireta e *sonificada* com *earcons* (símbolos auditivos). O desafio consiste em como mapear um *earcon* à um usuário específico ou, em outras palavras, como um som não-verbal particular poderia representar um agente específico no *workspace*.

A solução para essa questão foi inspirada por outras ferramentas colaborativas,

tais como o Google Drive e LucidChart, que usam cores para representar os usuários junto com seus *avatares* (Figure 3.3). Naturalmente, um usuário cego não percebe essa informação visual<sup>1</sup>; portanto, foi necessário projetar um substituto para os *avatares* e cores, de forma semelhante, porém cabível para uma interface auditiva. Logo, foi projetada a *sonificação* das cores por meio da **cor do tom**.



Fonte: *print* das interfaces.

A cor do tom também é conhecida como qualidade do tom ou, no campo da acústica, de timbre. Diferentes tipos de instrumentos se distinguem pelo timbre, sejam de percussão, cordas ou sopro. É possível reproduzir a mesma nota musical com instrumentos diferentes e, logo, com “cores” diferentes. Por exemplo, a nota *Dó* produzida por um piano ou um violão seria um *Dó* de distintas tonalidades.

Desse modo, foram atribuídos diferentes instrumentos para representar diferentes usuários. Por exemplo, considere dois coparticipantes no *workspace*, além do usuário local, neste caso, cada um teria um instrumento atribuído. Se estão presentes Maria e João, enquanto um violão representaria Maria, um piano poderia representar João. Assim, as pessoas podem ser representadas tanto pelos seus nomes de usuário (reproduzidos por TTS) quanto por instrumentos (reproduzidos por *earcons*), ou ambos, como projetado neste trabalho.

Existem três elementos de *workspace awareness* que abordam o conhecimento de “quem”, no sentido de quais usuários estão compartilhando uma sessão de modelagem simultaneamente com o usuário local. Na lista a seguir, os elementos de *workspace awareness* da categoria “quem?” são revisitados com a interface acessível projetada neste estudo:

- *Quem/Presença*: sempre que um usuário realiza o *login* e *logout* o sistema emite o som de um instrumento e o nome do usuário junto com as expressões “entrou” ou

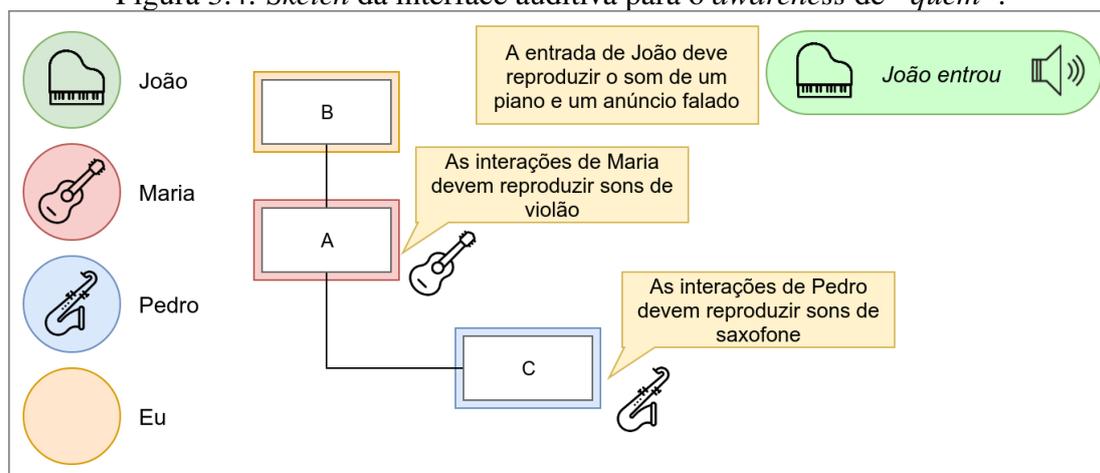
<sup>1</sup> Mesmo um usuário com deficiência visual conhecida como daltonismo poderia encontrar dificuldades em operar um aplicativo que use cores para representar os outros usuários.

“saiu”;

- *Quem/Identidade*: para listar todos os usuários presentes no *workspace* o usuário local pode pressionar uma tecla de atalho para o sistema reproduzir os nomes de usuário e seus instrumentos;
- *Quem/Autoria*: cada usuário pode criar, editar e excluir artefatos e suas propriedades, logo essas ações devem possuir sons para representá-las, reproduzidos com o instrumento do usuário para sinalizar ação e o agente;

Os usuários videntes perceberiam as mudanças no *workspace* tradicionalmente, isto é, via interface gráfica, tal como os editores colaborativos de diagramas atuais oferecem. Além disso, se um usuário cego está presente no *workspace*, os outros usuários precisam saber identificar os instrumentos que os representam e podem, opcionalmente, seguir as pistas auditivas. O *workspace* deve parecer-se com o *sketch* apresentado na Figura 3.4.

Figura 3.4: *Sketch* da interface auditiva para o *awareness* de “quem”.



Fonte: autor

### 3.3.2 Awareness de “o quê”

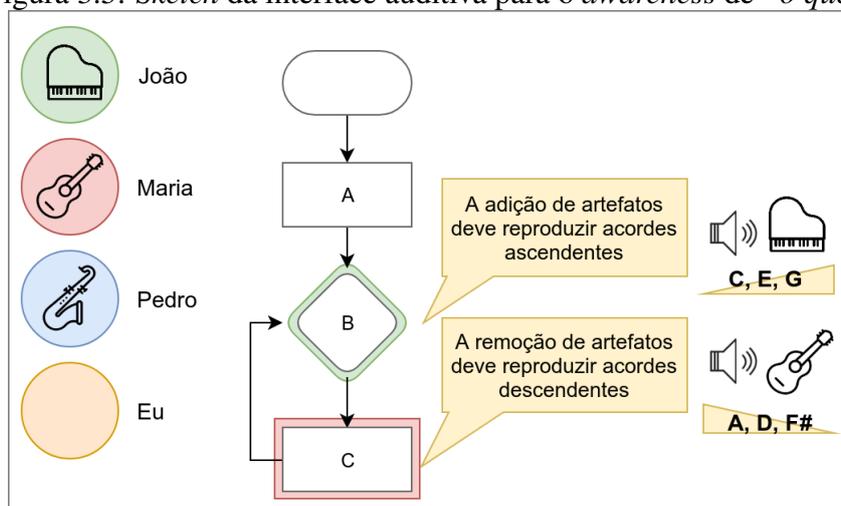
Um artefato pode ser alvo de várias ações, tais como ser criado, renomeado, excluído ou ter propriedades alteradas. Essas ações poderiam ser faladas aos usuários, no entanto, neste estudo elas são *sonificadas*. Para tal, as ações precisaram ser simplificadas para alcançar um projeto genérico aplicável independente do domínio sendo modelado, logo, três ações genéricas foram consideradas: inclusão, exclusão e alteração.

A *sonificação* projetada para essa questão foi variar a direção dos acordes, isto

é, diferenciando-os por tocá-los de forma contínua (sem variação de altura), ascendente (das notas baixas às altas) ou descendente (das notas altas às baixas), para representar respectivamente às ações de alteração, inclusão e exclusão por meio do efeito sonoro de distanciamento ou aproximação.

Por exemplo, considere o ato de incluir um artefato ordinário, que pode ser representado pela reprodução do acorde “*Dó maior*” de modo que as notas *Dó-Mi-Só* são reproduzidas sequencialmente em vez de simultaneamente (Fig. 3.5). O inverso, isto é, a reprodução de *Só-Mi-Dó* sequencialmente, representaria a exclusão de um artefato.

Figura 3.5: *Sketch* da interface auditiva para o *awareness* de “o quê”.



Fonte: autor

### 3.3.3 *Awareness* de “onde”

Para projetar uma interface auditiva que represente o *awareness* de “onde” foi necessário lidar com uma questão anterior: o leiaute espacial. Alguns autores (KENNEL, 1996; BROWN; PETTIFER; STEVENS, 2004) consideram que a posição bidimensional dos objetos pode ser relevante, cuja informação de localização é normalmente encontrada em soluções que usam métodos táteis ou cinestéticos (BROWN; STEVENS; PETTIFER, 2004). As propostas baseadas apenas nos recursos audíveis tipicamente usam técnicas de espacialização, sejam por fontes múltiplas (MIURA et al., 2016) ou por emulação em canais estéreo, conhecidos como áudio espacializado ou binaural (BALAN; MOLDOVEANU; MOLDOVEANU, 2015). Ademais, existem soluções multimodais, que utilizam recursos audíveis e hápticos para transmitir informações detalhadas sobre os nós do diagrama simultaneamente com a localização espacial (SYAL; CHATTERJI; SARDANA, 2016; COHEN; MEACHAM; SKAFF, 2006).

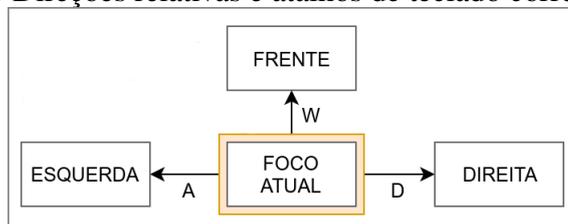
Entretanto, o leiaute espacial foi desconsiderado na interface auditiva projetada

neste trabalho, que foi baseado sobre a abstração dos diagramas como *node/link*, isto é, com ênfase nos relacionamentos (ver a Seção 2.3). Outros estudos usaram a mesma abordagem, isto é, não transmitiam a posição dos artefatos (MILLER, 2009). Por exemplo, Bennett (2002) conduziu experimentos com grupo de usuários cegos, em que uns recebiam a informação de posição e outros não; suas descobertas mostraram que a informação de posicionamento não gera uma diferença relevante na compreensão e aprendizado de diagramas.

O modelo projetado para a navegação é baseado em direções relativas ao artefato com o foco, de onde o usuário navega lateralmente ou para a frente. Foram planejadas três direções, apresentados na Figura 3.6, que também são usadas para *sonificar* o *awareness* de “onde” por meio de efeitos estereofônicos relativos ao posicionamento do usuário local em relação aos coparticipantes. O efeito estéreo permite usar fones de ouvido ordinários (estereofônicos) para sinalizar até 180°. A *sonificação* projetada para sinalizar as mudanças no *workspace* e opções para navegação dependem das direções relativas listadas a seguir:

- *esquerda*: o sinal auditivo é reproduzido apenas no canal esquerdo; pressionando a tecla A o usuário navega para o artefato à esquerda;
- *frente*: o sinal auditivo é reproduzido em ambos canais; pressionando a tecla W o usuário navega para o artefato à frente;
- *direita*: o sinal auditivo é reproduzido apenas no canal direito; pressionando a tecla D o usuário navega para o artefato à direita.

Figura 3.6: Direções relativas e atalhos de teclado correspondentes.

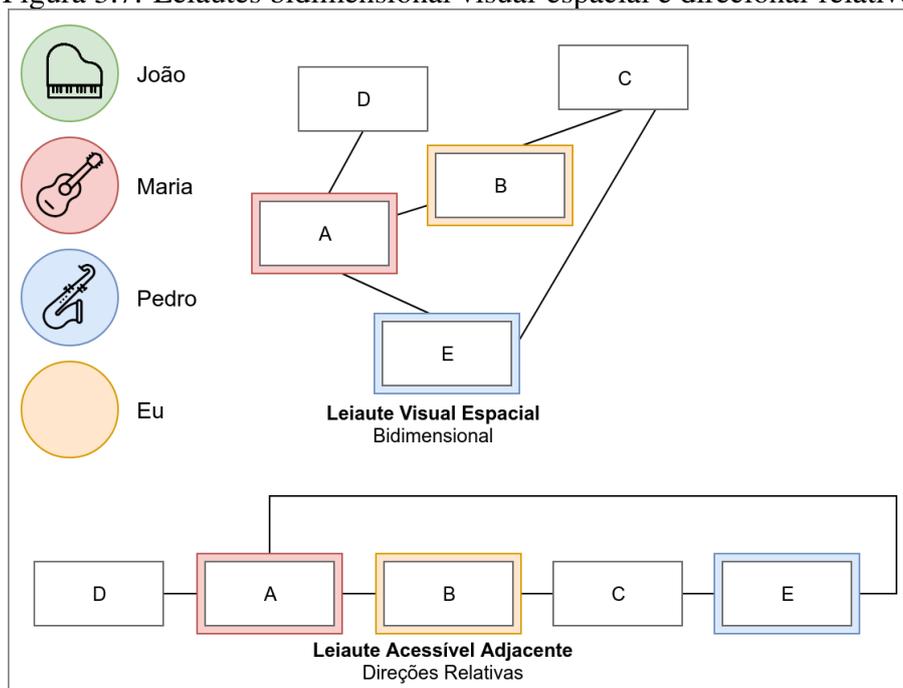


Fonte: autor.

Neste modelo alternativo acessível, o posicionamento bidimensional exato torna-se irrelevante para os usuários cegos, ou seja, um artefato desenhado logo abaixo do artefato com o foco pode ser representado como se estivesse à frente segundo o modelo auditivo. O algoritmo assegura que novos artefatos sejam conectados primeiro às extremidades esquerda e direita para priorizar a navegação lateral. As diferenças entre o leiaute visual-espacial e o direcional-relativo projetado podem ser vistas na Figura 3.7. Como

este estudo objetiva permitir o trabalho colaborativo de cegos e videntes, ambos leiautes são necessários, isto é, enquanto o visual-espacial bidimensional é usado pelos videntes, o direcional-relativo pretende auxiliar os cegos à perceber e navegar pelo *workspace* por meio da interface auditiva e técnica de *sonificação* usada para sinalizar cada elemento de *workspace awareness*.

Figura 3.7: Leiautes bidimensional visual-espacial e direcional-relativo.



Fonte: autor.

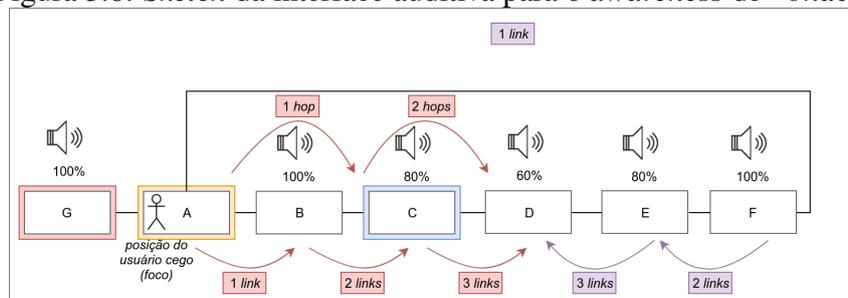
A navegação funciona por pressionar teclas de atalhos específicas que permitem a movimentação do foco de um artefato para outro artefato em uma dada direção. Por exemplo, segundo a Figura 3.7, considere que o usuário cego está com o foco no artefato *B*, então, mover para a esquerda o leva ao artefato *A*, do mesmo modo que mover para a direita três vezes<sup>2</sup>.

A *sonificação* projetada para um *awareness* de localização acessível foi inspirada nas interações reais de pessoas no mesmo espaço de trabalho físico com a seguinte pergunta: “*de que forma o cego identifica uma ação num espaço compartilhado?*”. Um resposta possível seria “*ouvindo os sons resultantes das ações*”. Neste caso, sempre que outros usuários presentes no *workspace* fizerem alterações, um som deve ser reproduzido de acordo com a direção relativa (Fig. 3.6). O usuário cego pode saber mais detalhes sobre a ação realizada se navegar na direção da origem do sinal sonoro.

<sup>2</sup>Sabe-se que essa é uma solução que necessita mais estudo e testes, contudo as questões de navegação não fazem parte da meta primária deste estudo, sendo que diversas pesquisas foram realizadas nesta categoria (ver a Seção 2.5)

A *sonificação* foi projetada para parecer distante ou próxima do usuário local (artefato com o foco) para comunicar a distância, por meio da variação da intensidade sonora (volume do som). Essa técnica causa um efeito de desvanecimento (*fading* em Inglês) a medida que as ações ocorram em objetos mais distantes do objeto focado pelo usuário local. Isto é, os sons das alterações adjacentes são reproduzidos com mais intensidade, enquanto as alterações distantes, a partir de um *hop*<sup>3</sup>, são reproduzidos com menor intensidade, gradualmente, à medida que os *hops* aumentam. A Figura 3.8 ilustra o processo de contagem de *links* e *hops*, e o desvanecimento pretendido da intensidade, projetado para reduzir 20% do volume por cada *hop*. Observe que o som deve parecer tão próximo quanto o caminho mínimo para chegar ao artefato alterado, isto é, com o menor número de *links* necessários para navegar até a origem do som.

Figura 3.8: *Sketch* da interface auditiva para o *awareness* de “onde”.



Fonte: autor.

As técnicas de sonificação projetadas para as categorias *awareness* e desenvolvimento da interface auditiva são descritas na lista a seguir:

- **Interface mista verbal/não-verbal para informar “quem está compartilhando o workspace comigo?”**: os nomes de usuário são falados ao usuário local via TTS seguido da reprodução de algumas notas com o instrumento musical associado aos usuários;
- **Interface auditiva para informar “o que meus colegas estão fazendo?”**: as ações dos usuários são *sonificadas* via *earcons* de acordo com a ação executada e reproduzidas com o instrumento associado ao usuário que está agindo;
- **Interface auditiva para informar “onde meus colegas estão trabalhando?”**: os *earcons* são reproduzidos de acordo com a direção relativa ao usuário local por meio de efeitos estereofônicos e com intensidade proporcional à distância do artefato alvo da ação.

<sup>3</sup>*Hop* é um termo da área de redes usado para denotar um seguimento entre uma origem e um destino, por meio da contagem de quantos nós foram “pulados”.

### 3.4 Primeira Iteração: Uma Interface Auditiva para Workspace Awareness

Esta seção apresenta o processo de projeto, implementação e avaliação do protótipo de uma ferramenta acessível para modelagem colaborativa de diagramas. O protótipo concretiza a interface auditiva projetada na Seção 3 e implementa as *sonificações* propostas para as categorias de *awareness*.

#### 3.4.1 Projeto e Implementação

A solução foi projetada na forma de uma ferramenta de modelagem *web*, acessível pela Internet por meio de um navegador para computadores *desktop*. Isto é, essa versão não considera o acesso por dispositivos móveis. A Tecnologia Assistiva foi implementada com base nas seguintes tecnologias:

- HTML5, CSS3 e JavaScript/ES6;
- SVG: formato aberto para gráficos vetoriais escalonáveis (*Scalable Vector Graphics*);
- Vue.js<sup>4</sup>: *framework* progressivo para construir interfaces de usuário reativas;
- SoundJS: biblioteca para reprodução de áudio<sup>5</sup>;
- Tonejs-instruments<sup>6</sup>: amostras de instrumentos comuns para uso em quaisquer bibliotecas para reprodução de áudio.

As classes e componentes podem ser vistos na Figura 3.9 e suas responsabilidades são descritas a seguir:

- *Workspace*: composto por artefatos e responsável por renderizá-los;
- *Artifact*: objeto genérico para representar um *node*, tabela, classe ou outro artefato de um diagrama;
- *Property*: representa os atributos de um artefato ou conexão particular;
- *Link*: objetos que agregam e unem dois artefatos, isto é, representam uma aresta entre os vértices;
- *Session*: guarda os dados de sessão, como os usuários logados e outras propriedades;
- *EventHub*: objeto responsável por rotear e propagar os eventos após a execução das

<sup>4</sup>Mais sobre Vue.js em <<https://br.vuejs.org/v2/guide/index.html>>

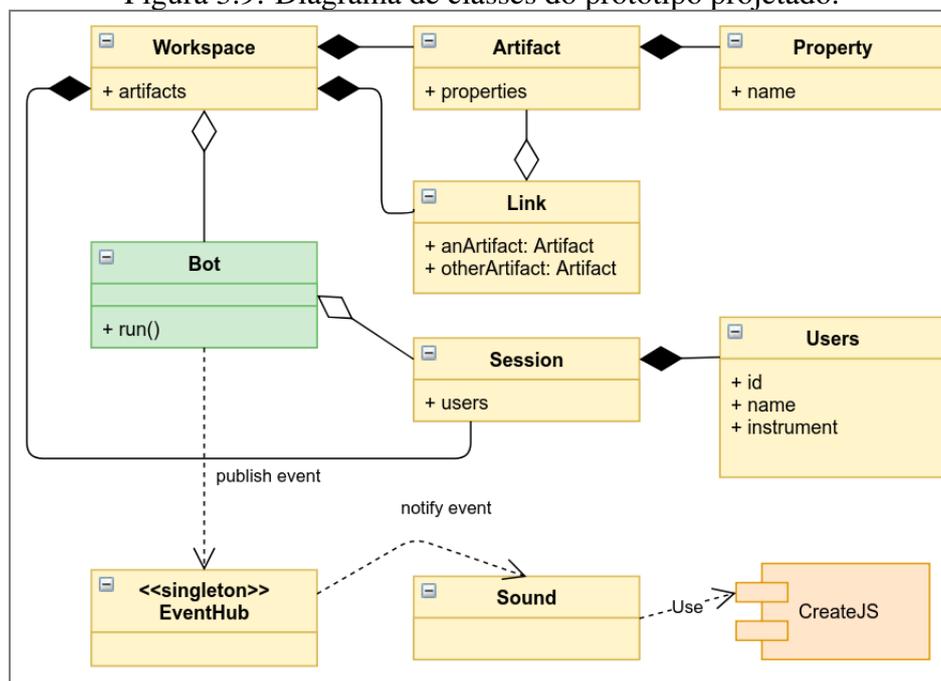
<sup>5</sup>Mais sobre CreateJS e SoundJS em <<https://createjs.com/>>

<sup>6</sup>Disponível para *download* em <<https://github.com/nbrosowsky/tonejs-instruments>>

ações;

- *Sound*: responsável por reproduzir amostras específicas de determinados instrumentos musicais;
- *Bot*: executor de testes o qual simula ações de usuários e mantém um histórico.

Figura 3.9: Diagrama de classes do protótipo projetado.



Fonte: autor.

Como um protótipo<sup>7</sup>, a ferramenta foi desenvolvida com as funcionalidades mínimas necessárias para a *sonificação* de elementos de *workspace awareness* nas três categorias básicas quanto ao instante presente.

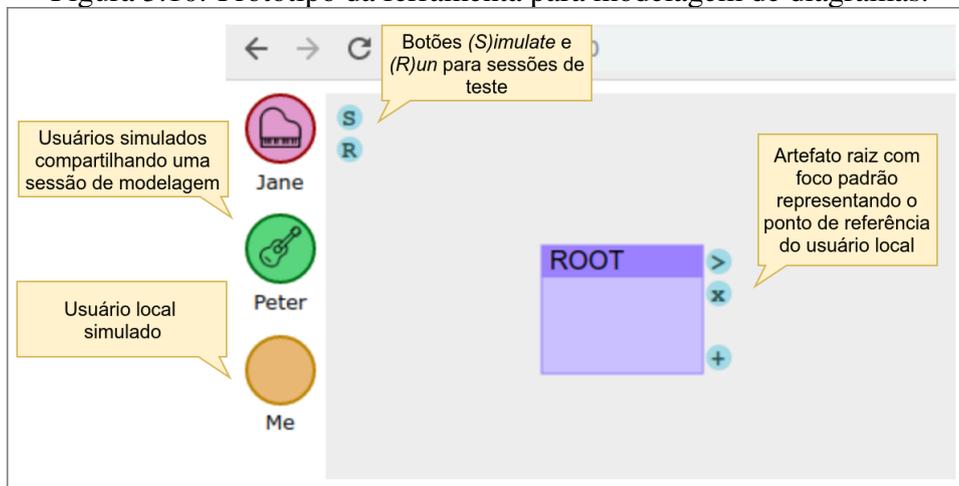
### 3.4.2 Avaliação

Para simular atividades de múltiplos usuários, a solução considera três usuários previamente registrados e logados: (a) “eu”, que representa o usuário local cego, (b) *Jane*, uma usuária vidente e remota representada por um piano como instrumento e (c) *Peter*, assim como Jane porém representado por um violão. O protótipo também considera que o usuário local (“eu”) está com o foco em um artefato preexistente e fixo chamado de “*root*” (Fig. 3.10). Desse modo, o desdobramento das ações de inclusão ou remoção de artefatos no *workspace* são ouvidas pelo usuário local relativamente ao artefato raiz com

<sup>7</sup>O protótipo desenvolvido é *open-source*, licenciado pela GNU General Public License v3.0 e disponível em <<https://github.com/marciojrtores/awareness>>.

o foco, por meio da reprodução do acorde *Dó maior*, com o intervalo de 100ms entre as notas, nas três direções (esquerda, direita e frente) que o diagrama pode expandir.

Figura 3.10: Protótipo da ferramenta para modelagem de diagramas.



Fonte: autor.

Dadas essas configurações, foram recrutados participantes para um teste de usabilidade que avaliou a interface auditiva projetada. Para tal, foi verificado se as *sonificações* projetadas para os elementos de *workspace awareness* eram percebidas corretamente e distinguíveis por meio da taxa de sucesso por tarefa (ver Seção 3.2.2). Nos tópicos a seguir são detalhados os (A) participantes, a (B) preparação do ambiente para avaliação, os (C) procedimentos e (D) coleta de dados.

#### A) Participantes

Nesta primeira iteração e na ausência de usuários cegos com conhecimento sobre diagramas, foram recrutados cinco usuários videntes para este teste de usabilidade<sup>8</sup> com experiência em diagramas e questões de usabilidade e acessibilidade.

#### B) Preparação do Ambiente

Para a preparação do ambiente foram usados como equipamentos um *laptop* Dell com processador Core i7 de 2.4Ghz 5ª geração e 8GB de memória RAM, um fone de ouvido com resposta às frequências de 20Hz a 20Khz. O *software* instalado foi o sistema operacional Ubuntu 16.04.2 LTS com os *drivers* mais recentes. Como navegador *web* foi usado o Google Chrome v75.0.3770.142.

<sup>8</sup>Estudantes do curso de Mestrado em Engenharia de Computação do Programa de Pós-graduação em Computação (PPGCOMP) e membros do grupo de pesquisa de Informática na Educação

O ambiente de teste foi instalado nas dependências do Centro de Ciências Computacionais/FURG. O *laptop* foi posicionado de modo que os participantes não tinham acesso à interface gráfica ou aos recursos de entrada (teclado e *mouse*). A única interface era o retorno dos sinais auditivos pelo fone de ouvido.

O volume do áudio foi ajustado para 50%, para não ultrapassar o limite de 50dB que, segundo a OMS (2015), representa o fim do conforto, enquanto 55dB é o início do estresse do aparelho auditivo.

### *C) Procedimentos*

As sessões de teste foram compostas de um tempo para treinamento e outro para condução, que somados compreendiam aproximadamente 20 minutos por participante. Durante o período de treinamento, cada participante teve a oportunidade de usar a ferramenta para perceber os sinais auditivos e todas as dúvidas foram respondidas, como o modelo direcional-relativo, instrumentos, efeito de desvanecimento (variação da intensidade/volume do áudio) e efeito estereofônico (alternância dos canais estéreo). Como os usuários eram videntes e membros do grupo de pesquisa, tiveram acesso ao modelo visual durante a etapa de treinamento, que durou aproximadamente 10 minutos. A sessão de teste só foi conduzida após o participante declarar que estava pronto.

Durante a sessão de teste os participantes foram expostos à 20 tarefas, aqui referenciadas como Casos de Teste (CT). Em cada CT, um usuário simulado aleatório (Jane ou Peter – com a mesma probabilidade), executa uma ação aleatória (inclusão – 60% prob. – ou exclusão – 40% prob. – de um artefato) em uma direção aleatória (esquerda, direita ou frente – com a mesma probabilidade). A probabilidade da ação favoreceu a inclusão para sempre gerar diagramas com artefatos — as ações se anulariam se a probabilidade de inclusão e exclusão fossem equilibrada. Cada iteração disparava um sinal auditivo que carregava quatro informações: (1) o usuário na forma de um instrumento (piano ou violão); (2) a ação na forma de um acorde ascendente ou descendente; (3) a direção reproduzida no canal esquerdo, direito ou ambos e (4) a distância na forma de decréscimo gradativo da intensidade do som.

Após a execução da atividade simulada, o CT foi concluído com a submissão das questões aos participantes, conforme Tabela 3.2. A sessão de teste, então, era pausada e retomada apenas após o participante responder o que percebeu<sup>9</sup>.

---

<sup>9</sup>Um exemplo sem pausas foi gravado para fins de demonstração e está disponível em <[https://drive.google.com/file/d/11FJJ0WXskjpTtV\\_su\\_th9fpN-Ua5dVmm/view](https://drive.google.com/file/d/11FJJ0WXskjpTtV_su_th9fpN-Ua5dVmm/view)>

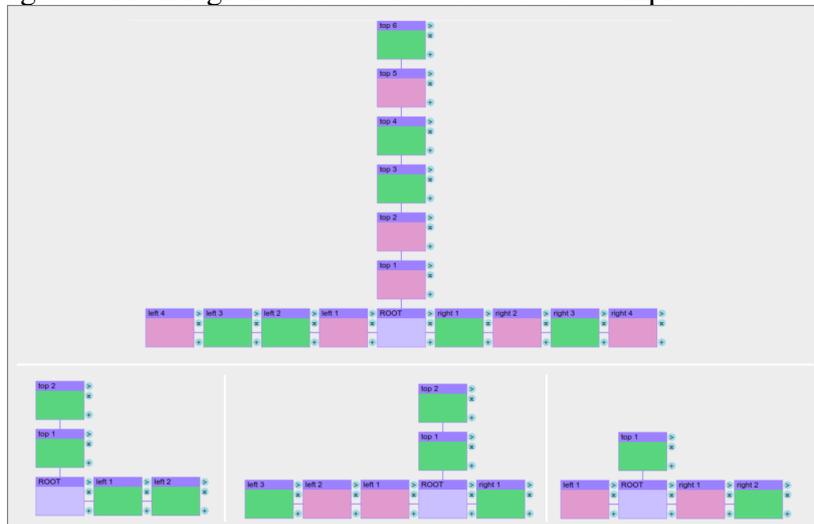
Tabela 3.2: Questões submetidas aos participantes.

| Questão               | Categoria                     | Respostas possíveis         |
|-----------------------|-------------------------------|-----------------------------|
| Quem realizou a ação? | Awareness de “quem/usuário”   | Jane/Piano ou Peter/Violão  |
| O que foi realizado?  | Awareness de “o quê/ação”     | Inclusão ou Exclusão        |
| Em que direção?       | Awareness de “onde/direção”   | Esquerda, Direta, ou Frente |
| A que distância?      | Awareness de “onde/distância” | Número entre 1 e 5 links    |

Fonte: autor.

Os artefatos resultantes de uma sessão de teste podem ser vistos na Figura 3.11. Os posicionamentos, rótulos e opções de colorização foram usados pelo autor/moderador, enquanto os entrevistados perceberam todas as ações apenas por meio da interface auditiva.

Figura 3.11: Diagramas resultantes de sessões completas de teste.



Fonte: autor

#### D) Coleta de Dados

Os dados das sessões de teste foram tabulados segundo o  $\log^{10}$  da ferramenta e comparados às respostas dos participantes. Um excerto das respostas pode ser observado na Tabela 3.3.

Tabela 3.3: Últimos quatro casos de teste do 2º participante.

| Usuário        |                | Ação           |                | Direção        |                | Distância      |                |             |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------|
| <i>Efetivo</i> | <i>Perceb.</i> | <i>Efetivo</i> | <i>Perceb.</i> | <i>Efetivo</i> | <i>Perceb.</i> | <i>Efetivo</i> | <i>Perceb.</i> | <i>Erro</i> |
| Jane           | Jane           | Inc.           | Inc.           | Esq.           | Esq.           | 3              | 3              | 0           |
| Jane           | Jane           | Inc.           | Inc.           | Frente         | Frente         | 3              | 2              | 1           |
| Peter          | Peter          | Exc.           | Exc.           | Esq.           | Esq.           | 4              | 2              | 2           |
| Peter          | Peter          | Inc.           | Inc.           | Esq.           | Esq.           | 3              | 2              | 1           |

Fonte: autor.

Os resultados são provenientes da tabulação de cinco sessões de teste com 20 casos de teste, no total de 100 atividades. Contudo, há um caso excepcional: em alguns testes

<sup>10</sup>Registro persistente dos eventos da ferramenta.

ocorreu de um ou mais artefatos serem incluídos ou excluídos a cinco ou mais *hops* de distância, o que não produziu sinal auditivo — situação resolvida na segunda iteração. Essas iterações silenciosas foram ignoradas; assim, algumas sessões acabaram com 19 casos de teste efetivos. No entanto, as ações ignoradas não influenciaram as médias.

### 3.5 Segunda Iteração: Atualização da Interface Auditiva e Opções Baseadas em Fala

Esta seção aborda as atualizações da interface auditiva e protótipo após primeira análise. Foram projetadas duas *sonificações* para comunicar a distância das ações, entre outras mudanças na sinalização dos elementos de *workspace awareness*. Também foram projetadas e desenvolvidas duas opções verbais, baseadas em fala sintetizada e natural.

#### 3.5.1 Atualização da Interface Auditiva

Com base nos resultados da primeira avaliação (Seção 4.1), entendeu-se a necessidade de reavaliar a *sonificação* de “*onde*” uma ação está ocorrendo, especificamente quanto a distância entre o artefato selecionado pelo usuário cego e o artefato que recebeu a ação dos outros usuários presentes no *workspace*. Adicionalmente, introduziu-se a ação de alteração, pois apenas a inclusão e exclusão estavam disponíveis na primeira etapa.

A *sonificação* da primeira interface comunicava a distância por meio da variação da intensidade dos sons, reduzindo em 20% o volume para cada *link* adicional de distância entre o objeto focado e a interação (Fig. 3.8). Os resultados da avaliação (Fig. 4.1) demonstraram que as cinco variações de volume (100, 80, 60, 40 e 20%) não eram percebidas corretamente, embora houvesse uma noção aproximada (Fig. 4.3). Portanto, foram projetadas duas atualizações para *sonificar* a distância: (A) alterações no desvanecimento da intensidade e (B) inclusão de uma nova técnica, com o alongamento das notas musicais (no formato de *beeps*).

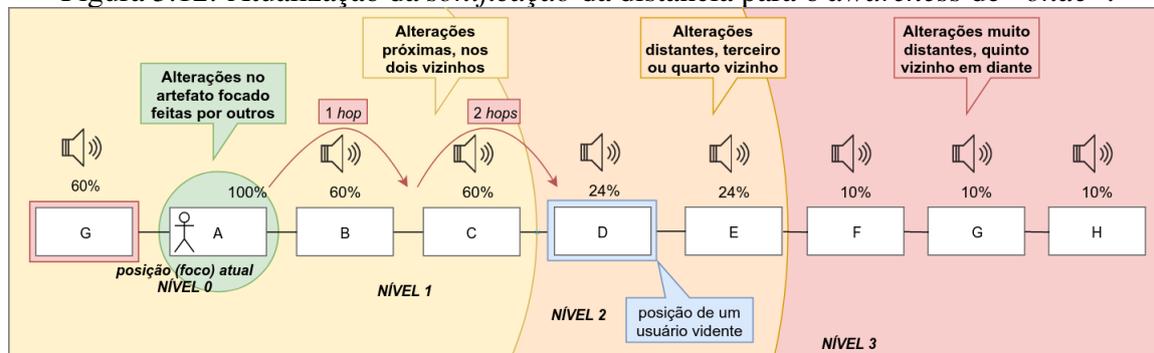
##### A) Atualização do Desvanecimento para Sinalizar Distância

O efeito de desvanecimento da intensidade do som foi atualizado para três níveis (eram cinco no projeto inicial – Figura 3.8). As diferenças também foram mais espaçadas, pois os saltos de 20% projetados na primeira iteração não introduziam variação suficiente para distinção. Por exemplo, uma ação realizada a 3 *links* de distância, e reproduzida com

60% de intensidade, não era corretamente distinguida da anterior, com 80% de intensidade a 2 *links* (25% de decréscimo de volume). No entanto, ações à 4 *links* de distância e reproduzidas com 40% de intensidade eram melhor diferenciadas das de 2 *links* com 80% de intensidade (50% de decréscimo de volume). Deste modo, foram definidos 3 níveis de intensidade para *sonificar* a distância, mais um nível de 100% se a ação for realizada no artefato atual (focado), que neste segundo protótipo pode sofrer alterações (Figura 3.12). O efeito de desvanecimento da intensidade foi projetado para reduzir o volume em 40% em relação ao nível anterior. Os três níveis podem ser vistos na lista a seguir:

- *100% de volume*: para ações que afetam o artefato focado — distância 0;
- *60% de volume*: para ações imediatas que acontecem no artefato imediato ou até dois *links* de distância;
- *24% de volume*: para ações próximas que acontecem no artefato a três ou quatro *links* de distância;
- *10% de volume*: para ações distantes que acontecem nos artefatos a mais de quatro *links* de distância;

Figura 3.12: Atualização da *sonificação* da distância para o *awareness* de “onde”.



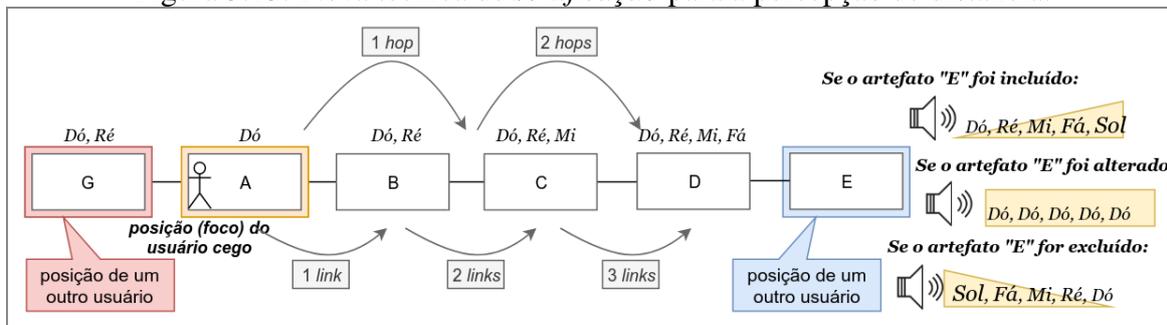
Fonte: autor.

### B) Reprodução de Notas Musicais Proporcionais à Distância

Uma nova *sonificação* foi projetada para sinalizar o *awareness* da distância por meio da variação da quantidade de notas musicais reproduzidas. Foram projetados seis níveis, reproduzidos com as notas *Dó-Ré-Mi-Fá-Sol-Lá-Si* em altura ascendente nas inclusões, descendente (*Si-Lá-Sol ...*) nas exclusões, e constante ( $n$  links + 1 vezes a nota *Dó*) nas alterações. Ainda, nesta estratégia, a intensidade acompanha a altura das notas em partes iguais, que produz os efeitos de *fade-in* e *fade-out* nas inclusões e exclusões, respectivamente. Isto é, as notas mais altas são reproduzidas com mais intensidade e as mais baixas com menor, à medida que a sequência é reproduzida. Por exemplo, conforme a Figura 3.13, se um usuário inclui um artefato *E* à direita e à 4 *links* de distância,

o usuário local posicionado no artefato A ouvirá a sequência de notas *Dó, Ré, Mi, Fá, Sol* reproduzida com o efeito de *fade-in*. A mesma sequência é reproduzida inversamente e com efeito de *fade-out* nas exclusões. No caso das alterações, a nota *Dó* é reproduzida *distância+1* vezes com volume constante (Figura 3.13).

Figura 3.13: Nova técnica de *sonificação* para a percepção de distância.



Fonte: autor.

### 3.5.2 Projeto de Interfaces Baseadas em Fala

Nesta segunda iteração também foram projetadas duas interfaces verbais, isto é, baseadas em fala, que adicionaram uma nova dimensão e serviram de referência para comparação com as interfaces auditivas não-verbais. A primeira foi projetada sobre a bem conhecida (A) fala sintetizada, isto é, com recursos de TTS e de modo que todos os elementos (*quem fez o quê e onde*) fossem narrados para o usuário. A segunda foi projetada para reproduzir uma (B) fala natural gravada de maneira que os elementos fossem narrados com a voz dos coparticipantes da sessão de modelagem, para reduzir o tempo de enunciação.

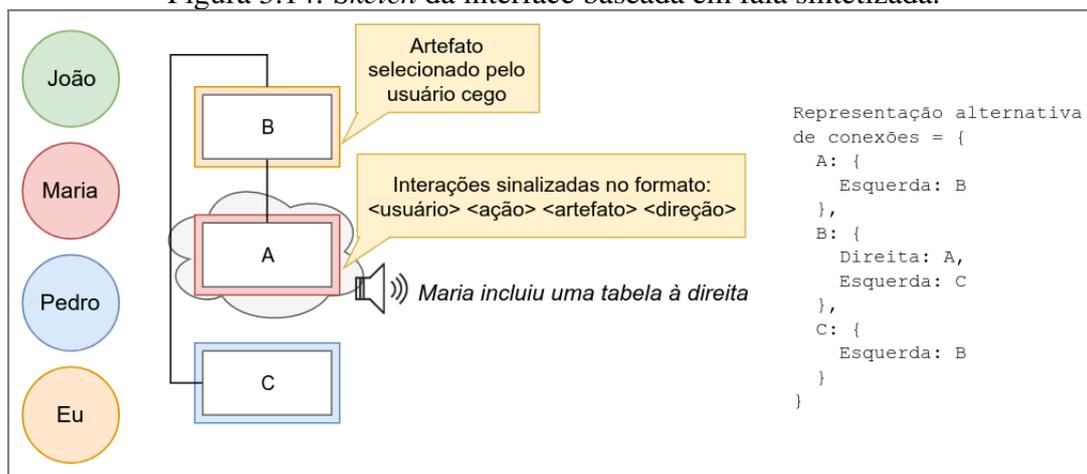
#### A) Interface com Fala Sintetizada

Enquanto os *earcons* são sons de curta duração, a representação por fala exige mais tempo para reprodução. Existem estudos (DINGLER; LINDSAY; WALKER, 2008; BEATTIE; BAILLIE; HALVEY, 2015) que comparam as situações em que cada técnica é mais adequada e discutem vantagens e desvantagens de cada uma. Por exemplo, acerca das soluções baseadas em fala, para Brewster (2003) a sua natureza serial é o motivo pela baixa velocidade ao apresentar informação, de modo que um usuário teria que ouvir a mensagem do início ao fim e que muitas palavras teriam que ser assimiladas até a mensagem ser entendida, no entanto, embora os *earcons* sejam curtos e ouvidos mais

rapidamente, ainda assim o usuário precisa aprender o significado das amostras, já que o significado não está inerentemente contido na mensagem.

A sinalização das categorias de *awareness* com fala sintetizada foi projetada com a seguinte sintaxe: <usuário> {incluiu | excluiu | alterou} {um(a) | esse(a)} <artefato> à {esquerda | frente | direita}. Por exemplo, em um editor de diagramas de entidade-relacionamento poderia ser anunciado: “*Maria incluiu uma tabela à direita*” (Fig. 3.14).

Figura 3.14: *Sketch* da interface baseada em fala sintetizada.



Fonte: autor.

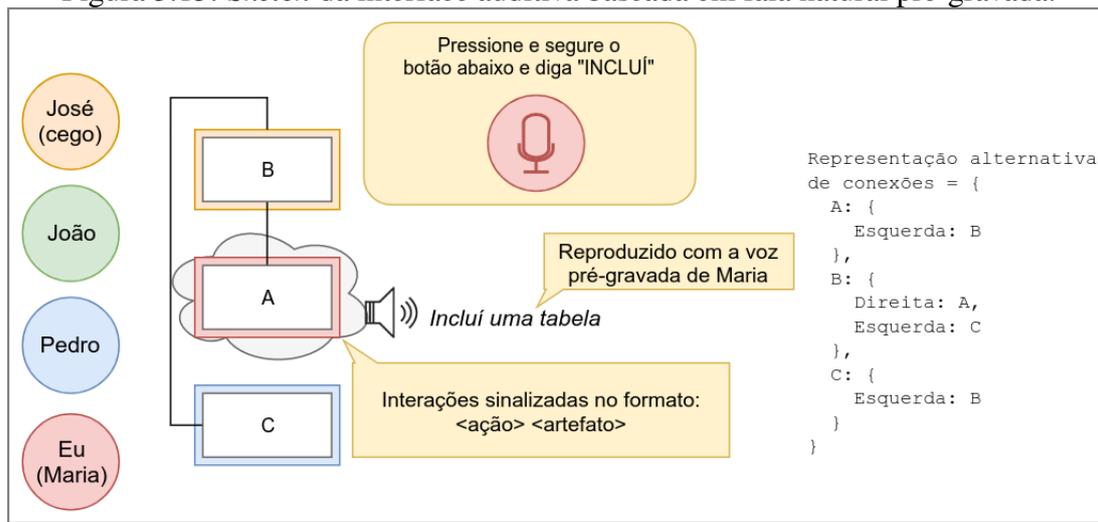
A sinalização da distância não foi codificada no enunciado, pois o tornaria mais extenso e exigiria mais tempo para reprodução. Por exemplo, o enunciado “*Maria incluiu uma tabela seis "links" à direita*” tem a duração de 4 s na velocidade padrão (1.0), enquanto sem enunciar a distância tem a duração de 3.2 s — 20% menor. Assim, manteve-se o efeito de desvanecimento da intensidade (Fig. 3.12) para comunicar a distância aproximada (imediate, perto, longe), que foi comparada às soluções *sonificadas* projetadas anteriormente. Por fim, a comunicação da direção foi codificada no enunciado e não por meio do efeito estereofônico como nas interfaces auditivas.

### B) Interface com Fala Natural

Foi projetada uma segunda interface baseada em fala, alternativa à fala sintetizada, por meio da gravação dos enunciados com as vozes naturais dos coparticipantes. É uma solução comparável aos *earcons* em tempo de reprodução, com enunciados de curta duração, aproximadamente 2 s para reprodução na velocidade padrão (30% menor que na fala sintetizada) e entre 1 s e 1,5 s se acelerada. O tempo é reduzido por omitir a comunica-

ção do usuário e direção no enunciado, que são codificados, em vez, no timbre da voz do colega, para comunicar a *awareness* de “quem”, e no efeito estereofônico para sinalizar a *awareness* de direção. Não obstante, mantém a vantagem das interfaces *faladas*, de codificar o significado na mensagem.

Figura 3.15: *Sketch* da interface auditiva baseada em fala natural pré-gravada.



Fonte: autor.

Os enunciados foram projetados com a sintaxe <ação> <artefato> e são reproduzidos com as falas pré-gravadas dos coparticipantes, conforme Figura 3.15. Nesta interface, manteve-se o efeito estereofônico para comunicar a direção e o efeito de desvanecimento para comunicar distância (Fig. 3.12), logo, parte da informação está codificada no enunciado (ação, artefato) enquanto outra parte usa recursos de *sonificação* (usuário e local).

### 3.5.3 Considerações para Avaliação

Junto com a interface auditiva projetada durante a primeira iteração deste estudo, agora são três não-verbais *sonificadas* com *earcons*. Nas seções de avaliação e resultados elas são referenciadas como *SONIV1*, *SONIV2* e *SONIV3*. As duas interfaces baseadas em fala foram nomeadas, para referência, como *FALASIN* e *FALANAT*. As cinco opções de interface e suas diferentes decisões de projeto podem ser conferidas na Tabela 3.4.

Tabela 3.4: Comparação das decisões de projeto para cada interface.

| <b>Interface</b> | <b>Usuário</b>                         | <b>Ação</b>  | <b>Direção</b>               | <b>Distância</b>  |
|------------------|--|--|------------------------------|---|
| <b>SONiv1</b>    | Timbre do instrumento                  | Reprodução de três notas de altura ascendente nas inclusões e descendente para exclusões   | Efeito estereofônico         | Efeito de desvanecimento com 5 níveis de intensidade: 100, 80, 60, 40 e 20% de volume       |
| <b>SONiv2</b>    | Timbre do instrumento                  | Reprodução de uma nota para alterações e duas notas de altura ascendente nas inclusões e descendente nas exclusões                                 | Efeito estereofônico         | Efeito de desvanecimento com 3 níveis de intensidade: 60, 24 e 10% (artefato focado 100%)   |
| <b>SONiv3</b>    | Timbre do instrumento                  | Sequências constantes da nota <i>Dó</i> nas alterações, sequências de <i>Dó</i> a <i>Si</i> nas inclusões e de <i>Si</i> a <i>Dó</i> nas exclusões | Efeito estereofônico         | Reprodução das notas proporcional à distância em <i>links</i> com efeito de desvanecimento  |
| <b>FALASIN</b>   | Narrado por fala sintetizada           | Narrada por fala sintetizada   | Narrada por fala sintetizada | Efeito de desvanecimento com 3 níveis de intensidade: 60, 24 e 10% (artefato focado a 100%) |
| <b>FALANAT</b>   | Timbre da voz narrado por fala natural | Narrada por fala natural   | Efeito estereofônico         | Efeito de desvanecimento com 3 níveis de intensidade: 60, 24 e 10% (artefato focado a 100%) |

Fonte: autor.

### 3.5.4 Atualizações do Protótipo

Foram desenvolvidas funcionalidades de interação do usuário com o diagrama que não estavam disponíveis na primeira iteração. Esses recursos enriqueceram a experiência da ferramenta e permitiram que os usuários avaliassem melhor a carga de trabalho<sup>11</sup>, ao contrário se apenas recebessem os sinais passivamente, como foi na primeira etapa. A lista com os novos recursos e suas finalidades está a seguir:

<sup>11</sup> Eles também foram implementados com a meta de viabilizar o protótipo como produto após essa pesquisa.

- *Usuários Presentes*: a tecla “U” narra os usuário presentes com os instrumentos que os representam no caso das interfaces auditivas com *earcons*. Este recurso permite que o usuário local saiba quem está presente no *workspace*;
- *Descreve Artefato*: a tecla “Barra de Espaços” narra o nome do artefato focado, por exemplo, "Tabela alunos", e permite que o usuário conheça os detalhes de cada artefato/objeto/*node*;
- *Vizinhos*: a tecla “V” narra os vizinhos do objeto focado, por exemplo, "Tabela turmas à direita, tabela matrículas à esquerda". Esse recurso permite que o usuário saiba quais são os artefatos/objetos/*nodes* no entorno e tenha uma noção de contexto;
- *Conexões*: a tecla “C” narra as conexões (direções) disponíveis, por exemplo, "Não há conexões", "Conexão à esquerda". Permite que os usuários saibam para quais direções podem navegar, isto é, mover o foco;
- *Movimento*: as teclas “A, W e D” movem o foco para o artefato à esquerda, frente e direita, respectivamente. Foi incluído, também, um ícone auditivo para comunicar colisão, reproduzido se não for possível mover-se em alguma direção;
- *Aumentar/Reduzir Velocidade da Fala*: com as teclas “+” e “-” o usuário pode alterar a velocidade da fala, sintetizada e natural, nos seguintes níveis: 0.5, 0.7, 1.0, 1.3, 1.6, 1.9. O volume padrão utilizado foi 1.3, isto é, 130% de velocidade — ou 30% mais rápido que o normal;
- *Replay/Repetir Última Sinalização*: a tecla “S”, localizada entre as teclas de movimento, permite que o usuário ouça novamente o último sinal, isto é, o *earcon* ou enunciado disparado pela ação dos outros usuários. Uma característica relevante deste recurso, que foi aplicada durante o teste de usabilidade, é que se o usuário se mover no diagrama e pressionar *replay*, a sinalização é realizada novamente, porém relativa ao novo artefato focado. Logo, esse recurso permite que o usuário persiga o local exato da ação por meio da aproximação constante.

### 3.5.5 Avaliação

Como na primeira avaliação (Seção 3.4.2), a ferramenta inicia a simulação com dois usuários previamente registrados e *logados* (neste caso, *Márcio* e *Tiago*), além do usuário local. O protótipo da primeira iteração dispunha de um artefato fixo, enquanto

nesta versão o usuário pode navegar pelos artefatos e seguir os eventos, isto é, mover-se na direção do local aonde as ações estão ocorrendo. Por fim, um diferencial importante deste protótipo foi a possibilidade de repetir o último sinal auditivo.

O modelo genérico é baseado em *node/links*, com suporte à muitos tipos específicos de diagrama, enquanto nesta versão do protótipo os *nodes* são descritos concretamente como “*Tabelas*” e os *links* como relacionamentos.

Foram conduzidas quatro sessões de teste, uma para cada interface projetada, rotuladas como *SONIV2*, *SONIV3*, *FALASIN* e *FALANAT*, sobre as quais os participantes ouviram as sinalizações *sonificadas* e *faladas* que, então, eram verificadas se foram percebidas corretamente e distinguíveis por meio do método de taxa de sucesso por tarefa (ver Seção 3.2.2) e, ao fim da sessão, segundo as respostas às questões subjetivas do formulário NASA-TLX (ver Seção 3.2.3). Os detalhes acerca da seleção dos (A) participantes, (B) preparação do ambiente e (C) procedimentos são descritos nas seções a seguir.

#### A) *Participantes*

Nesta segunda iteração foram recrutados três usuários videntes, nas mesmas condições da primeira avaliação (Seção 3.4.2), e três participantes cegos, alunos de cursos superiores relacionados à área da computação, um com cegueira congênita e dois com adquirida. Todos os participantes possuíam conhecimento sobre informações diagramáticas e alguns com experiência em diagramas específicos, por exemplo, grafos, entidade-relacionamento, UML. Um participante cego, em particular, possuía experiência em diagramas elétricos, também conhecido como esquema elétrico, de um curso de eletrotécnica anterior não concluído. A experiência com informações conectadas era necessária para a realização dos testes e para que o treinamento fosse dedicado apenas aos quesitos de sistema e não à capacitação do que seriam informações diagramáticas.

#### B) *Preparação do Ambiente*

Foram usados os mesmos equipamentos da primeira avaliação (Seção 3.4.2) com a adição de um teclado *wireless* QWERTY padrão ABNT, que foi usado pelos participantes para interagir com diagrama.

Como na primeira avaliação, o *laptop* foi posicionado de modo que os participantes videntes não tivessem acesso à interface gráfica e obtivessem o retorno apenas por áudio, no entanto, diferente da primeira iteração, podiam usar o teclado como recurso de

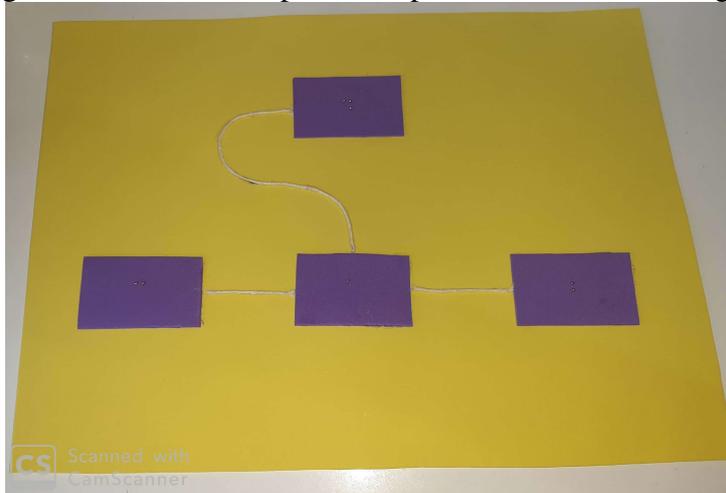
entrada.

O ambiente de teste foi instalado nas dependências do Centro de Ciências Computacionais para o teste com videntes, nas instalações do IFSul Pelotas para atender um estudante cego<sup>12</sup> e na Associação Escola Loius Braille em Pelotas<sup>13</sup> para atender os outros dois estudantes cegos, evitando assim que eles tivessem de se deslocar até Rio Grande.

### C) Procedimento

As sessões de teste foram precedidas de um tempo para treinamento para os participantes terem contato com as quatro interfaces, além de conhecer e experimentar as teclas de atalho fundamentais. O modelo direcional-relativo foi apresentado visualmente aos videntes, no entanto, aos usuários cegos foi necessário apresentar por meio de um modelo físico tátil, confeccionado com EVA, corda para os *links* e alfinetes para rótulos em *Braille* (Figura 3.16), de modo que compreendessem o motivo de poder se movimentar apenas para esquerda, direita e frente, diferente de um espaço bidimensional que permitiria a movimentação vertical cima/baixo.

Figura 3.16: Modelo háptico tátil para treinamento dos cegos.



Fonte: autor.

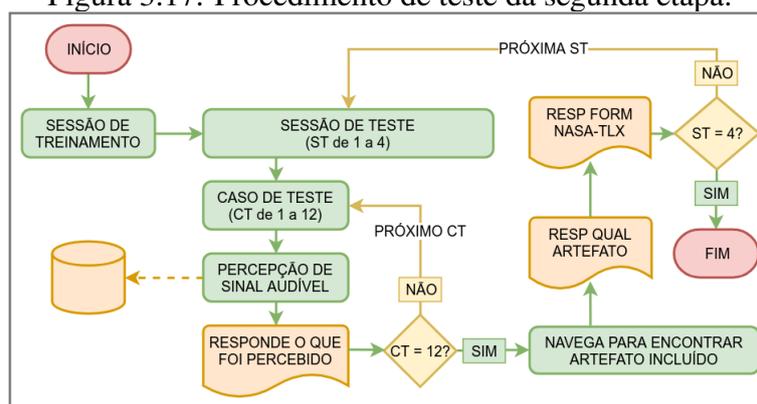
Em cada uma das quatro Sessões de Teste (ST), os participantes foram expostos à 12 tarefas, aqui referenciadas como Casos de Teste (CT). Em cada CT, o moderador, no papel de um usuário simulado, executa uma ação que dispara um saída audível, conforme interface sob teste, que comunica as quatro informações de *awareness*: “*quem/usuário*”, “*o quê/ação*”, “*onde/direção*” e “*onde/distância*”. Após cada CT, o participante, então,

<sup>12</sup>Em espaço cedido gentilmente pela coordenação do curso de Sistemas para Internet.

<sup>13</sup>Em espaço cedido gentilmente pela direção da associação.

responde o que percebeu. Ao fim da ST é adicionado um último artefato e dada ao participante a tarefa de encontrá-lo com o auxílio das funcionalidades de movimento, *replay* e descrição. Essa tarefa busca ampliar a experiência do participante, exigir um esforço físico e contribuir para o preenchimento das dimensões exigidas segundo as escalas da ferramenta NASA-TLX (ver Seção 3.2.3). O procedimento completo pode ser visto na Figura 3.17.

Figura 3.17: Procedimento de teste da segunda etapa.



Fonte: autor.

#### D) Coleta de Dados

Os dados das STs foram armazenados em uma base de dados em tempo-real orientada a documentos conhecida como Google Firestore. As informações foram organizadas pelas sessões, nas quais são armazenadas os eventos, isto é, cada atividade e sinalização reproduzida ao usuário, incluindo as repetições (*replays*) se acionadas. Como pode ser visto na Figura 3.18, o registro de cada sinal contém a sequência (*sequence*), o *id* do usuário (*user*), a ação (*action*), direção (*pan*) e distância (*distance*).

Figura 3.18: Excerto da base de dados (*log*) das sessões de treinamento.

```

{
  "sessions": {
    "40ozNmhEU0NIctEbAbtu": {
      "name": "sujeito-1-teste-fala-sintetizada",
      "timestamp": {
        "_seconds": 1576613520,
        "_nanoseconds": 308000000
      }
    },
    "subCollection": {
      "sessions/40ozNmhEU0NIctEbAbtu/interactions": {
        "7vjmmZ2L035dwp3804IX": {
          "timestamp": {
            "_seconds": 1576613591,
            "_nanoseconds": 341000000
          },
          "pan": 0,
          "sequence": 8,
          "distance": 1,
          "action": "addition",
          "user": 1
        }
      }
    }
  }
}
  
```

Fonte: autor.

Os dados foram tabulados por participante e tipo de interface, como pode ser visto na Figura 3.19. As colunas representam o número da sequência, usuário, ação, direção (*pan*) e distância, que foram recuperados dos registros na base de dados. O nível de distância depende do projeto de cada interface, especificamente. Na interface auditiva *SONiv2*, a distância é comunicada pela intensidade do áudio reproduzido de acordo com os níveis 1:60%, 2:24% ou 3:10%. Os *replays* foram agrupados a partir das sinalizações repetidas em sequência encontradas nos registros e consideradas como indicadores na etapa de análise. Cada sessão obteve uma taxa de sucesso por categoria de *awareness* (*quem, o quê, onde* — direção e distância) e geral.

Figura 3.19: Tabulação dos registros de sinais executados e percebidos.

| Participante #1            |         |      |         |           |       |        |                         |     |      |     |         |     |           |     |  |
|----------------------------|---------|------|---------|-----------|-------|--------|-------------------------|-----|------|-----|---------|-----|-----------|-----|--|
| INTERFACE AUDITIVA SONiv2  |         |      |         |           |       |        |                         |     |      |     |         |     |           |     |  |
| SINALIZAÇÕES REALIZADAS    |         |      |         |           |       |        | SINALIZAÇÕES PERCEBIDAS |     |      |     |         |     |           |     |  |
| Seq                        | Usuário | Ação | Direção | Distância | Nível | Replay | Usuário                 |     | Ação |     | Direção |     | Distância |     |  |
|                            |         |      |         |           |       |        | Vlr                     | Pts | Vlr  | Pts | Vlr     | Pts | Vlr       | Pts |  |
| 2                          | 1       | I    | 1       | 1         | 1     | 0      | 1                       | 1   | I    | 1   | 1       | 1   | 1         | 1   |  |
| 3                          | 1       | I    | 1       | 2         | 1     | 0      | 1                       | 1   | I    | 1   | 1       | 1   | 1         | 1   |  |
| 4                          | 2       | I    | 1       | 3         | 2     | 0      | 2                       | 1   | I    | 1   | 1       | 1   | 1         | 0   |  |
| 5                          | 2       | I    | 1       | 4         | 2     | 0      | 2                       | 1   | I    | 1   | 1       | 1   | 1         | 0   |  |
| 6                          | 1       | I    | -1      | 1         | 1     | 2      | 1                       | 1   | I    | 1   | -1      | 1   | 1         | 1   |  |
| 9                          | 1       | I    | -1      | 2         | 1     | 0      | 1                       | 1   | I    | 1   | -1      | 1   | 1         | 1   |  |
| 10                         | 2       | A    | -1      | 2         | 1     | 10     | 2                       | 1   | E    | 0   | -1      | 1   | 1         | 1   |  |
| 21                         | 2       | A    | 1       | 4         | 2     | 0      | 2                       | 1   | E    | 0   | 1       | 1   | 2         | 1   |  |
| 22                         | 1       | I    | 1       | 5         | 3     | 5      | 1                       | 1   | I    | 1   | 1       | 1   | 3         | 1   |  |
| 28                         | 1       | I    | 0       | 1         | 1     | 2      | 1                       | 1   | I    | 1   | 0       | 1   | 1         | 1   |  |
| 31                         | 1       | I    | 0       | 2         | 1     | 0      | 1                       | 1   | I    | 1   | 0       | 1   | 1         | 1   |  |
| 32                         | 2       | E    | -1      | 2         | 1     | 0      | 2                       | 1   | E    | 1   | -1      | 1   | 2         | 0   |  |
| Casos de Teste: 12         |         |      |         |           |       |        | Pontuação:              |     | 12   | 10  | 12      | 9   |           |     |  |
| Replays: 19                |         |      |         |           |       |        | Taxa de Sucesso:        |     | 100% | 83% | 100%    | 75% |           |     |  |
| Total de interações: 31    |         |      |         |           |       |        |                         |     |      |     |         |     |           |     |  |
| Taxa de sucesso geral: 90% |         |      |         |           |       |        |                         |     |      |     |         |     |           |     |  |

Fonte: autor.

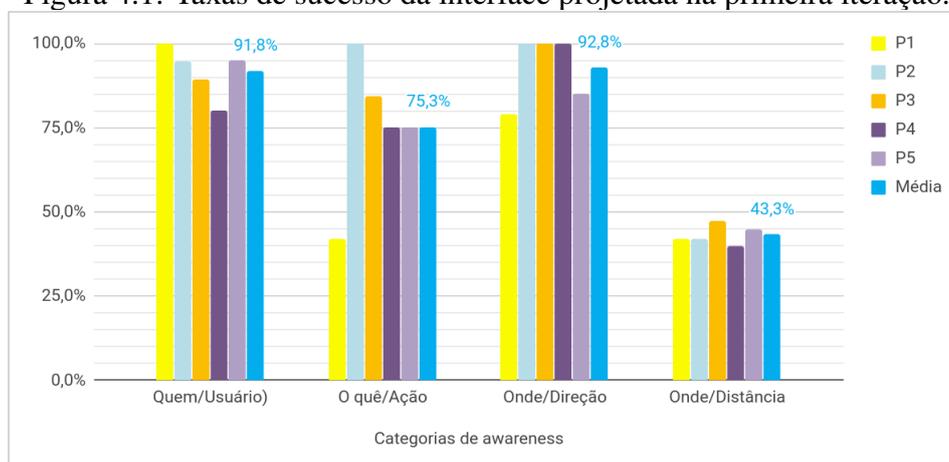
Os indicadores foram sintetizados e submetidos à avaliação estatística de modo a verificar a dispersão dos dados e reportar médias, valores mínimos, máximos e desvio padrão. Foi conduzido um Teste *t* pareado bicaudal para assegurar a significância na comparação das interfaces projetadas nessa iteração, e um Teste *t* independente para comparações entre a interface auditiva projetada na iteração anterior (*SONiv1*) e as projetadas nessa iteração. Os resultados, gerais e específicos, são apresentados no capítulo a seguir.

## 4 RESULTADOS

### 4.1 Primeira Iteração

O gráfico comparativo pode ser visto na Figura 4.1, que demonstra a taxa de sucesso individual dos participantes (P1,...,P5) e destaca a média da taxa entre eles por categoria de *awareness*. O teste demonstrou uma taxa de sucesso de pelo menos 75% para as primeiras três categorias (*quem? o quê? onde (direção)?*). A exceção foi a percepção de distância, pertencente ao *awareness* de “*onde?*”, que obteve apenas 43% de respostas exatas em média. Todas as taxas de sucesso podem ser vistas na Tabela 4.1.

Figura 4.1: Taxas de sucesso da interface projetada na primeira iteração.



Fonte: autor.

Tabela 4.1: Taxas de sucesso por participante e categoria de *awareness*

| Participante        | Quem       | Ação       | Direção    | Dist. Exata | Dist Aprox. ( $\pm 1$ ) |
|---------------------|------------|------------|------------|-------------|-------------------------|
| P1                  | 100%       | 42%        | 79%        | 42%         | 58%                     |
| P2                  | 95%        | 100%       | 100%       | 42%         | 61%                     |
| P3                  | 89%        | 84%        | 100%       | 47%         | 68%                     |
| P4                  | 80%        | 75%        | 100%       | 40%         | 58%                     |
| P5                  | 95%        | 75%        | 85%        | 45%         | 63%                     |
| <b>Média</b>        | <b>92%</b> | <b>75%</b> | <b>93%</b> | <b>43%</b>  | <b>61%</b>              |
| <b>Desv. Padrão</b> | <b>8%</b>  | <b>21%</b> | <b>10%</b> | <b>3%</b>   | <b>4%</b>               |

Fonte: autor.

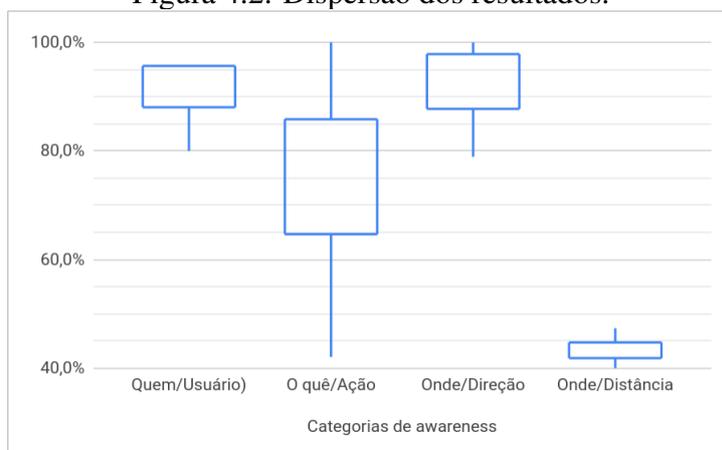
Quanto ao *awareness* de “*quem*”, houve um mínimo de 80% de percepções corretas, máximo de 100%, e baixa dispersão (8% de desvio padrão). É importante mencionar que os instrumentos usados, piano e violão, possuem timbres semelhantes; logo, esses resultados foram obtidos próximo ao pior cenário, logo, a taxa de sucesso média de 92% é bastante significativa.

Os sujeitos precisaram perceber se os acordes foram reproduzidos nas direções ascendente ou descendente para saber *o quê* aconteceu, isto é, qual foi a ação realizada

(inclusão ou exclusão). O pior resultado apresentou uma taxa de sucesso de 42%, enquanto o melhor foi de 100%, com desvio padrão de 21%. Apesar dos resultados não serem insatisfatórios, essa categoria precisa de atenção na segunda iteração da pesquisa.

A categoria “*onde*”, *awareness* da direção em que a ação aconteceu, obteve o melhor resultado entre todas, com 93% de taxa de sucesso média. Além disso, três dos cinco sujeitos atingiram a pontuação máxima. As percepções do usuário e direção foram as menos dispersas, enquanto a categoria “*o quê/ação*” obteve resultados com maior dispersão (Figura 4.2).

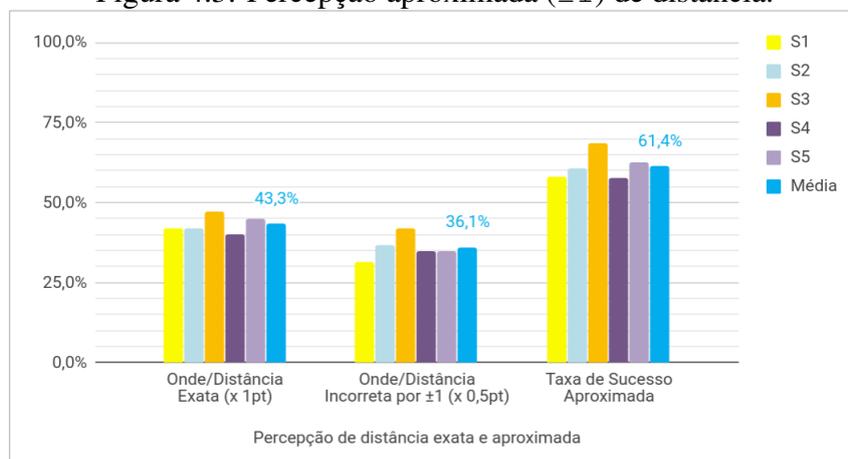
Figura 4.2: Dispersão dos resultados.



Fonte: autor.

A percepção da distância obteve resultados insatisfatórios, com apenas 43% de percepções exatas. As respostas incorretas foram analisadas e notou-se que 36% desviaram por um *link*. Isto é, embora as respostas não tenham sido exatas, foram razoavelmente precisas. Ao atribuir 0,5 ponto às percepções na margem de erro de  $\pm 1$  obteve-se uma taxa de sucesso média de 61% (Fig. 4.3). Contudo, esses resultados obtidos com a *sonificação* da distância demonstraram que essa categoria precisava de melhorias na segunda iteração.

Figura 4.3: Percepção aproximada ( $\pm 1$ ) de distância.

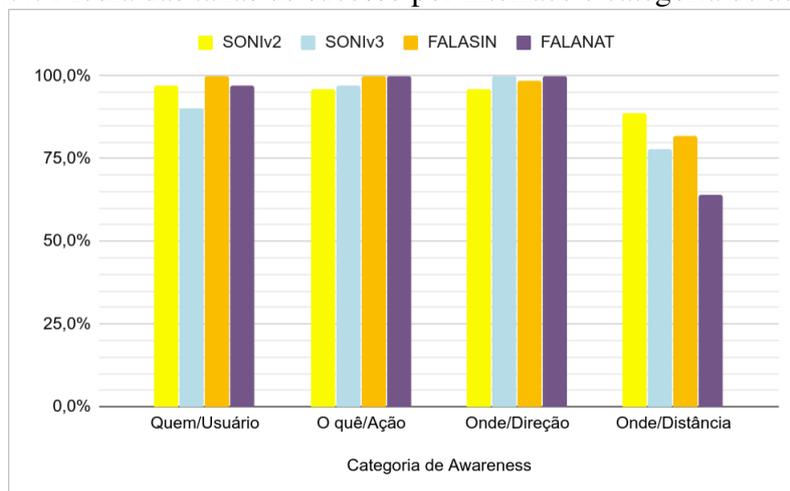


Fonte: autor

## 4.2 Segunda Iteração

Na segunda iteração as quatro interfaces projetadas foram testadas por meio de interações as quais os participantes precisavam identificar. A quantidade de percepções corretas resultou na taxa de sucesso. O quadro geral desses resultados pode ser visto na Figura 4.4. As quatro interfaces produziram resultados satisfatórios para as três categorias de *awareness* “quem”, “o quê/qual” e “onde/direção” e resultados razoáveis na avaliação de “onde/distância”. Exames mais detalhados são apresentados no fim dessa Seção.

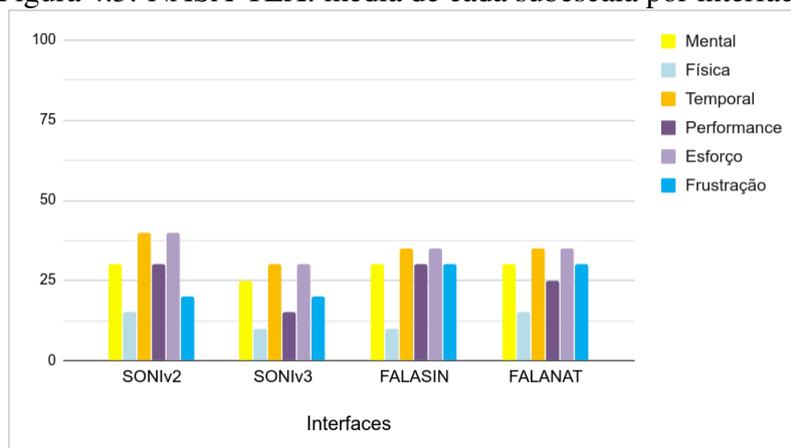
Figura 4.4: Média das taxas de sucesso por interface e categoria de *awareness*.



Fonte: autor.

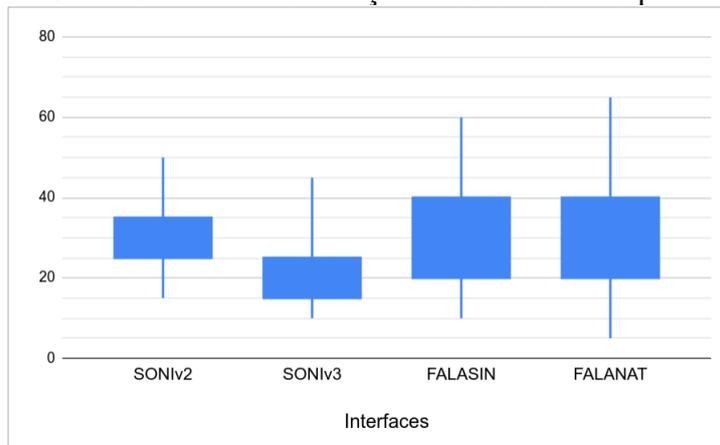
Como segundo instrumento de avaliação foi medida a carga de trabalho percebida, com respostas às seis subescalas, segundo o padrão NASA Task Load Index (Seção 3.2) sem pesos. O quadro geral das respostas pode ser visto na Figura 4.5. Os participantes informaram sua percepção de esforço para obter a performance objetiva descrita pela taxa de sucesso, logo, quanto menor o valor, menor a carga e melhor o resultado. Desta forma, uma interface com melhores resultados objetivos, como a fala sintetizada (*FALASIN*), não necessariamente será reportada com um carga de trabalho menor. Este é o caso, por exemplo, da interface *SONiv3* que, embora com taxa regular de sucesso, apresentou valores baixos na maioria das subescalas de avaliação subjetiva. Inclusive, ambas interfaces baseadas em fala obtiveram resultados próximos, mesmo que os resultados objetivos da taxa de performance (Fig. 4.4) tenham sido diferentes.

Figura 4.5: NASA-TLX: média de cada subescala por interface.



Fonte: autor.

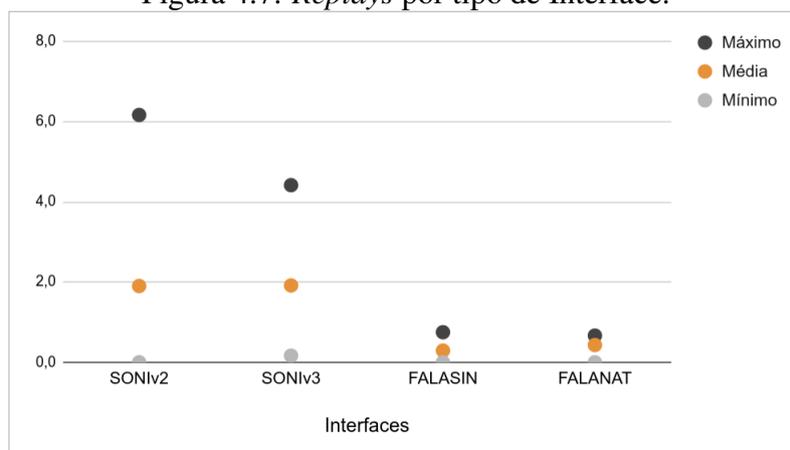
O *score* final do NASA-TLX também foi favorável à terceira interface *sonificada*, como pode ser visto na Figura 4.6. Ambas interfaces auditivas produziram resultados mais coesos, isto é, com baixa dispersão. As interfaces *faladas*, especialmente a fala natural, obtiveram resultados mais dispersos.

Figura 4.6: NASA-TLX: distribuição dos *scores* finais por interface.

Fonte: autor.

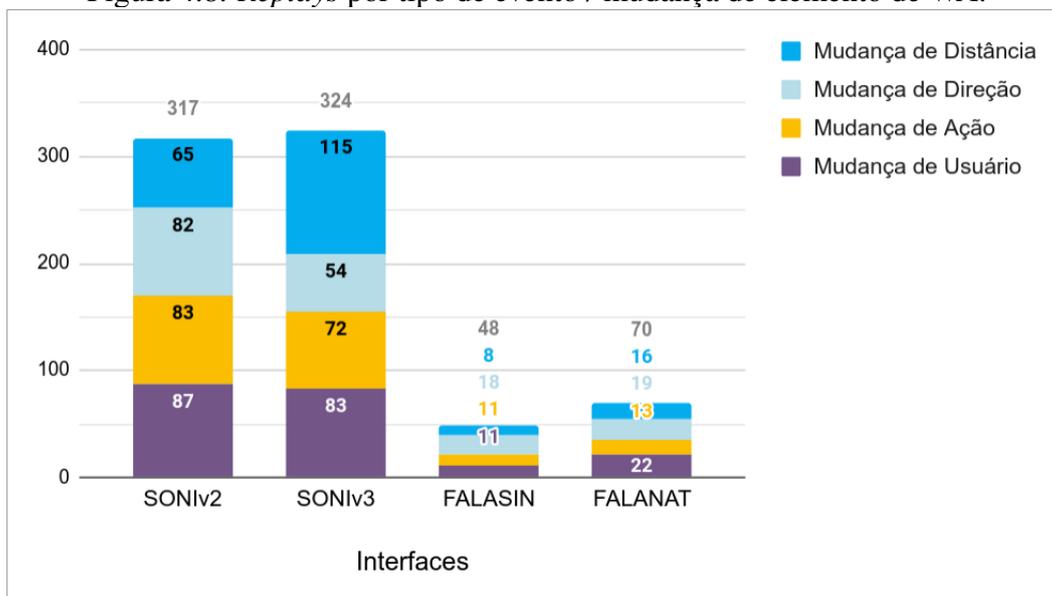
Durante a sessão de teste de usabilidade, os participantes podiam repetir o sinal recebido, seja por não compreensão ou confirmação. A quantidade de *replays* foi salva no *log*<sup>1</sup> de ações. Considerou-se *replay* todos os sinais reproduzidos em sequência em que não houveram mudanças no estado da *workspace*. Houve um equilíbrio intragrupos, isto é, as interfaces auditivas produziram resultados similares entre si, enquanto as interfaces *faladas* também produziram resultados semelhantes em seu grupo, como pode ser visto na Figura 4.7. Comparando entre os grupos, foram executadas duas repetições em média nas interfaces auditivas, enquanto nas interfaces baseadas em fala as repetições foram raras.

<sup>1</sup>Registrado no banco de dados.

Figura 4.7: *Replays* por tipo de Interface.

Fonte: autor.

As mudanças de estado no *workspace* também foram salvas no *log*, de modo que foi possível correlacionar com a taxa de *replays*. Buscou-se saber quais mudanças de estado, elemento de *workspace awareness* que sofreu mudança, que estimulam o participante a solicitar a repetição da sinalização. O gráfico apresentado na Figura 4.8 discrimina os *replays* por tipo de evento. Em outras palavras, qual sinalização foi diferente no momento do *replay*. Na interface auditiva *SONiv2*, os *replays* foram frequentes após mudanças de usuário e ação, enquanto na *SONiv3* as repetições foram frequentes após mudanças de distância. Outro ponto observado na *SONiv2* foi a baixa taxa de repetição após mudanças de direção. As interfaces baseadas em fala, no entanto, produziram taxas de *replay* equilibradas.

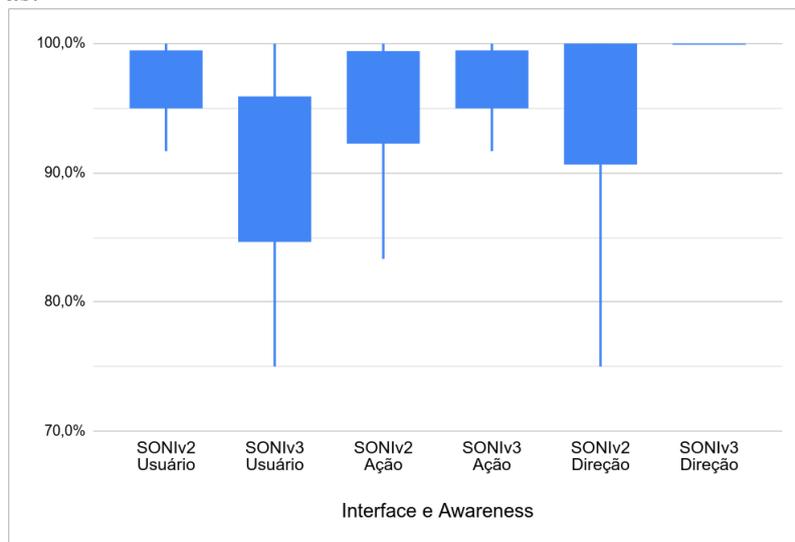
Figura 4.8: *Replays* por tipo de evento / mudança de elemento de WA.

Fonte: autor.

### 4.2.1 Interfaces Auditivas

As interfaces auditivas *SONiv2* e *SONiv3*, projetadas na segunda iteração, também foram *sonificadas* usando a técnica de *earcons*, como na primeira versão (*SONiv1*). Ambas apresentaram taxa de sucesso partir de 90%, como pode ser visto na Figura 4.9. A interface *SONiv2* ( $M=0,9722$ ;  $DP=0,0430$ ) apresentou resultados melhores que *SONiv3* ( $M=0,9028$ ;  $DP=0,1108$ ) na percepção do usuário, porém não significante ( $t(5)=0,19$ ;  $p > 0,05$ ). A interface *SONiv3* obteve resultados ligeiramente melhores ( $t(5)=0,70$ ;  $p > 0,05$ ) que *SONiv2* na identificação da ação. Quanto a noção de direção, os participantes alcançaram a taxa máxima de sucesso na interface *SONiv3* ( $M=1,00$ ;  $DP=0$ ), resultados que embora melhores que *SONiv2* ( $M=0,9583$ ;  $DP=0,1021$ ) não foram considerados significativos ( $t(5)=0,70$ ;  $p > 0,05$ ). A distância, que obteve resultados dispersos, é relatada separadamente.

Figura 4.9: Média das taxas de sucesso das interfaces auditivas projetadas na primeira e segunda etapas.

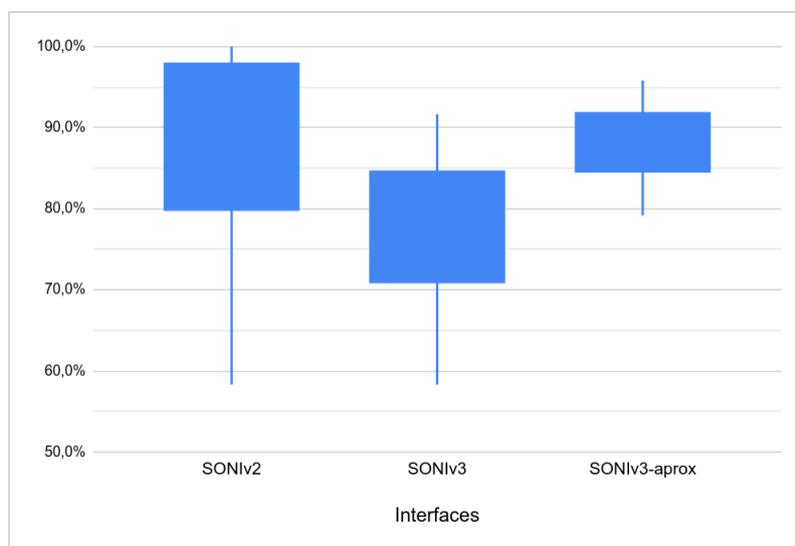


Fonte: autor.

A percepção da distância foi um dos pontos fracos na primeira interface auditiva projetada (*SONiv1*). Por isso, ela foi reprojetaada nessa segunda etapa para ter menos e mais espaçadas variações de intensidade sonora na *SONiv2*, e mais níveis e sinalizados por mais notas em *SONiv3*. Neste sentido, a interface auditiva *SONiv3* equipara-se à *SONiv1* na profundidade, isto é, um usuário pode perceber uma ação como à cinco *links* de distância em ambas. Ainda, a interface *SONiv3* foi submetida à duas métricas: (1) eram pontuadas apenas as percepções exatas, isto é, não foi atribuído um valor para aproximações, como identificar quatro níveis de distância quando foram cinco de fato;

(2) chamada de *SONiv3-aprox*, atribuiu-se 0,5 ponto nos erros por um, como identificar cinco níveis quando foram seis de fato. Como pode ser visto na Figura 4.10, os resultados da interface *SONiv3* foram inferiores aos da *SONiv2* nas percepções exatas, enquanto na percepção aproximada os resultados foram melhores e menos dispersos. A diferença não é significativa de *SONiv2* ( $M=0,8889$ ;  $DP=0,18$ ) para *SONiv3* ( $M=0,7778$ ;  $DP=0,1361$ ), nem de *SONiv2* para *SONiv3-aprox* ( $M=0,8819$ ;  $DP=0,0718$ ), ambos com  $p > 0,05$ . No entanto, *SONiv3-aprox* foi significativamente melhor que *SONiv3* ( $t(5)=0,01$ ;  $p < 0,05$ ).

Figura 4.10: Taxas de sucesso na percepção de distância em *SONiv2*, *SONiv3* e *SONiv3-aprox*.

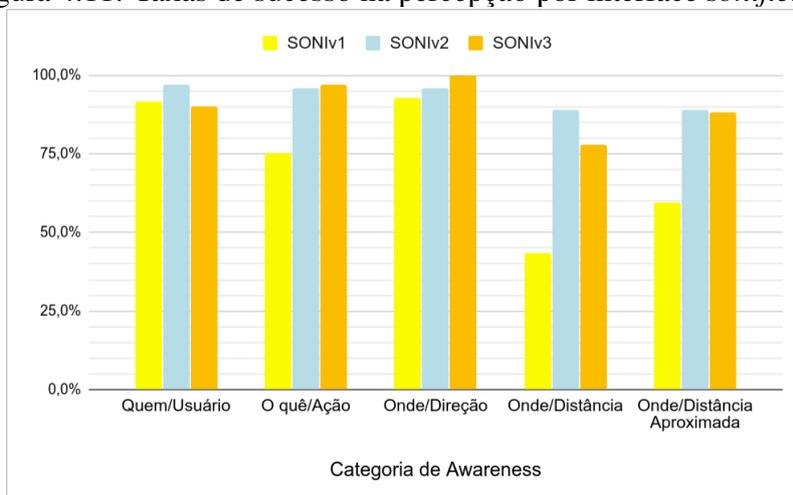


Fonte: autor.

Neste ponto, pretende-se resgatar a primeira interface auditiva, referenciada como *SONiv1*, e compará-la aos resultados das interfaces auditivas projetadas nessa etapa, *SONiv2* e *SONiv3*. Na identificação de usuário, ação e direção, as novas interfaces *sonificadas* obtiveram melhores resultados que a *SONiv1* (Figura 4.11). Examinando os resultados quanto a distância, cabe lembrar que a interface *SONiv1* utilizou, inicialmente, cinco níveis de intensidade sonora para representá-la. No primeiro teste de usabilidade ela foi avaliada segundo a percepção exata dos participantes, que recebia um ponto, e aproximada, que recebia 0,5 ponto se as percepções estivessem no intervalo aproximado de  $\pm 1$ , neste caso referenciada como *SONiv1-aprox*. A interface *SONiv2* representa a distância com três níveis de intensidade. A interface *SONiv3* reproduz várias notas, até sete, em sequência para sinalizar distância. Assim, são duas interfaces com vários níveis, *SONiv1* e *SONiv3*, com as técnicas de desvanecimento e variação de quantidade de notas reproduzidas para sinalizar a distância, respectivamente. As outras três, *SONiv1-aprox*, *SONiv2*

e *SONiv3-aprox*, são simplificadas, isto é, ou com menos variações de intensidade, ou avaliadas aproximadamente, e foram tratadas como se tivessem menos níveis de distância sinalizados ou percebidos.

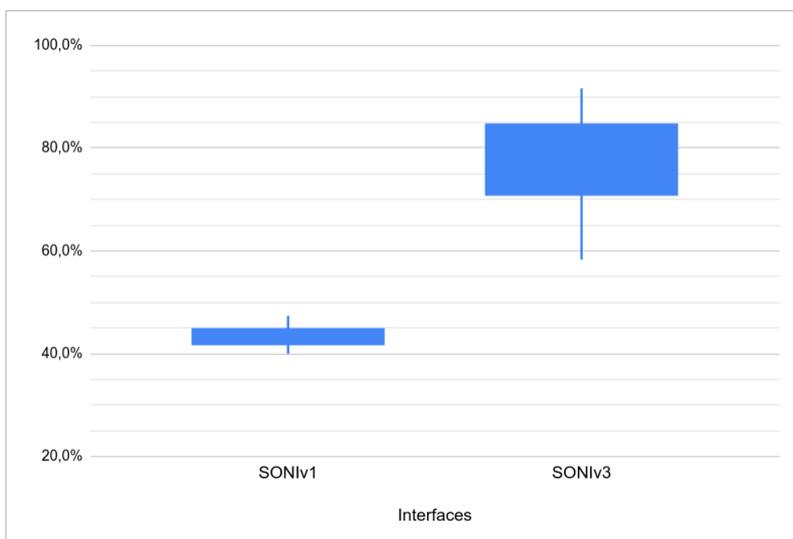
Figura 4.11: Taxas de sucesso na percepção por interface *sonificada*.



Fonte: autor.

A seguir, na Figura 4.12, são apresentados os resultados das interfaces com mais gradações na sinalização da distância, pela variação da intensidade (v1) e do número e altura das notas reproduzidas (v3). Ambas foram avaliadas segundo a percepção exata, isto é, foi atribuído um ponto às percepções corretas e zero às incorretas. Os resultados da interface *SONiv3* ( $M=0,4332$ ;  $DP=0,0288$ ) foram significativamente melhores que da *SONiv1* ( $M=0,7778$ ;  $DP=0,1361$ ) para  $t(4, 5)=0,0012$  e  $p < 0,01$ , nível de confiança de 99%.

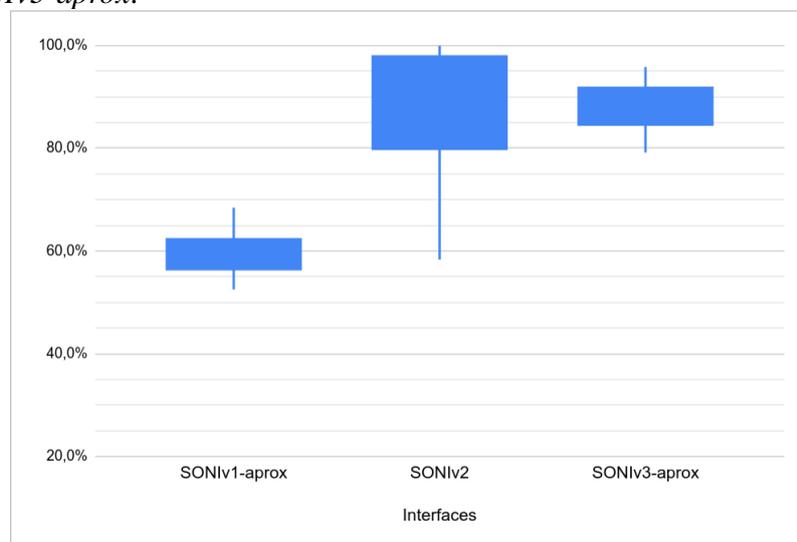
Figura 4.12: Taxas de sucesso na percepção de distância exata nas interfaces *SONiv1* e *SONiv3*.



Fonte: autor.

Considerando menos graduações de distância (Fig. 4.13), ambas interfaces *SONiv2* ( $M=0,8889$ ;  $DP=0,18$ ) e *SONiv3-aprox* ( $M=0,8819$ ;  $DP=0,0718$ ) obtiveram resultados significativamente ( $M=0,8889$ ;  $DP=0,18$ ) melhores que *SONiv1-aprox* ( $M=0,5937$ ;  $DP=0,0583$ ), se considerado  $t(5)=0,0085$ ;  $p < 0,01$  e  $t(5)=0,00004$ ;  $p < 0,0001$  respectivamente — a primeira com 99% e a segunda com 99,99% de confiança.

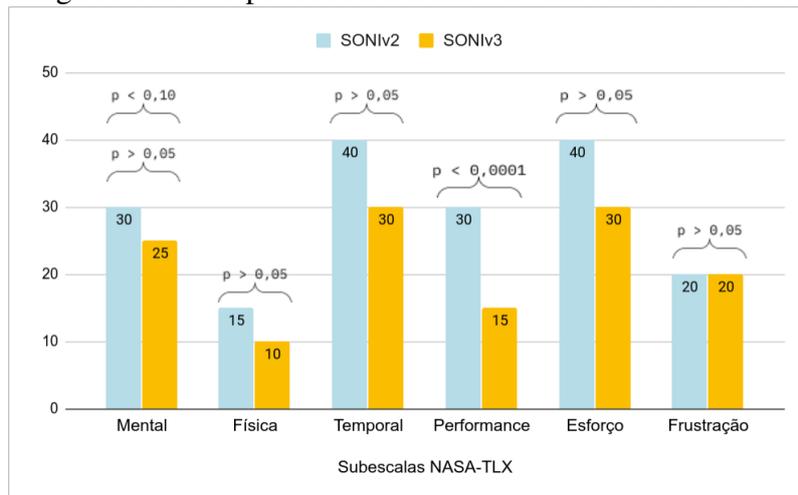
Figura 4.13: Taxas de sucesso na percepção de distância nas interfaces *SONiv2*, *SONiv1-aprox* e *SONiv3-aprox*.



Fonte: autor.

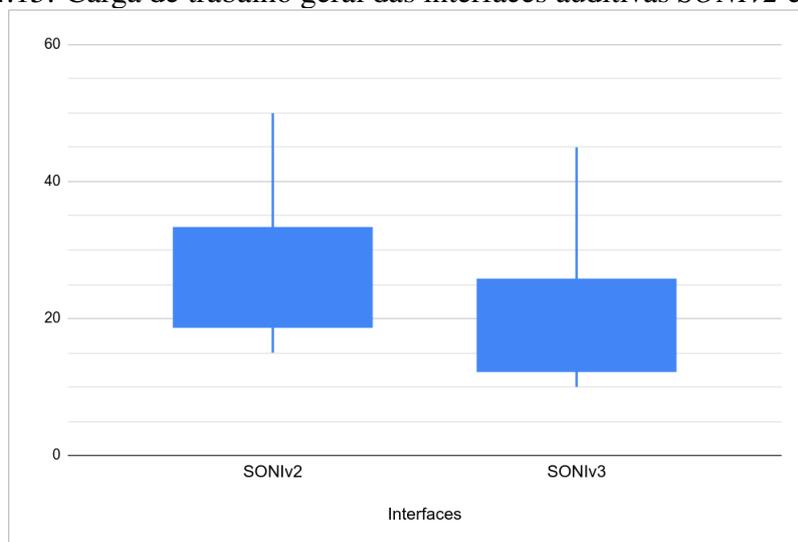
A carga de trabalho das interfaces *SONiv2* e *SONiv3* foi avaliada segundo as subescaladas da ferramenta NASA-TLX conforme Figura 4.14. A interface auditiva *SONiv3* obteve um resultado melhor do que a *SONiv2*, em todas subescalas exceto *frustração*. O destaque é a subescala de *performance*, na qual os participantes relatam quão bem-sucedidos sentiram-se e sua satisfação com os resultados alcançados. Nesta métrica, a interface *SONiv3* ( $M=29,17$ ;  $DP=13,93$ ) foi significativamente melhor avaliada que *SONiv2* ( $M=15,83$ ;  $DP=14,29$ ), com 99,99% de confiança ( $t(5)=0,00005$ ;  $p < 0,0001$ ). Com 90% de confiança ( $t(5)=0,08$ ;  $p < 0,1$ ), a carga mental percebida foi menor na *SONiv3*,  $M=22,5$ ;  $DP=11,29$  contra  $M=30,83$ ;  $DP=12,01$  da *SONiv2*. A carga de trabalho geral foi melhor na *SONiv3* (Fig. 4.15), mas com baixa significância ( $t(5)=0,3$ ).

Figura 4.14: Carga de trabalho por subescala das interfaces auditivas *SONiv2* e *SONiv3*.



Fonte: autor.

Figura 4.15: Carga de trabalho geral das interfaces auditivas *SONiv2* e *SONiv3*.

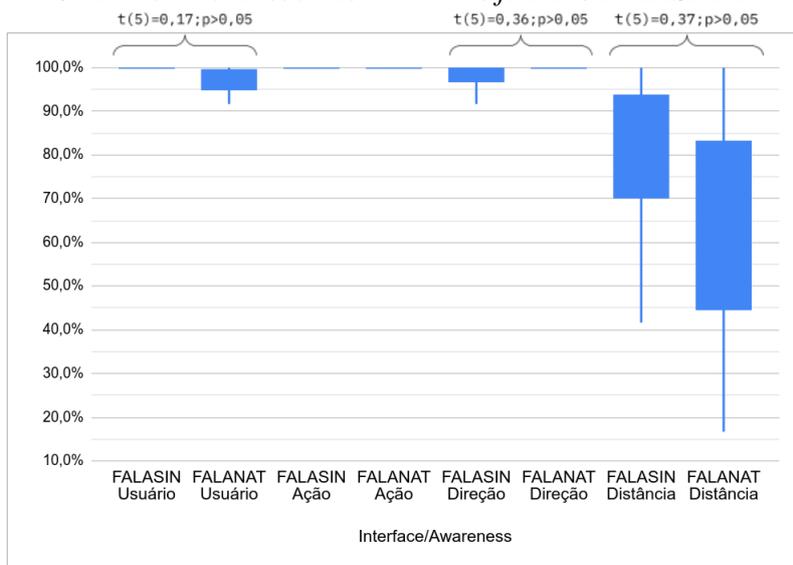


Fonte: autor.

#### 4.2.2 Interfaces Baseadas em Fala

As interfaces *faladas* produziram taxas de sucesso equivalentes entre si (Fig. 4.16), exceto na categoria “*onde/distância*”, na qual a taxa de sucesso foi de 82% da *FALASIN*, mais expressiva que a taxa de 64% da *FALANAT*. No entanto, essa diferença não foi significativamente melhor.

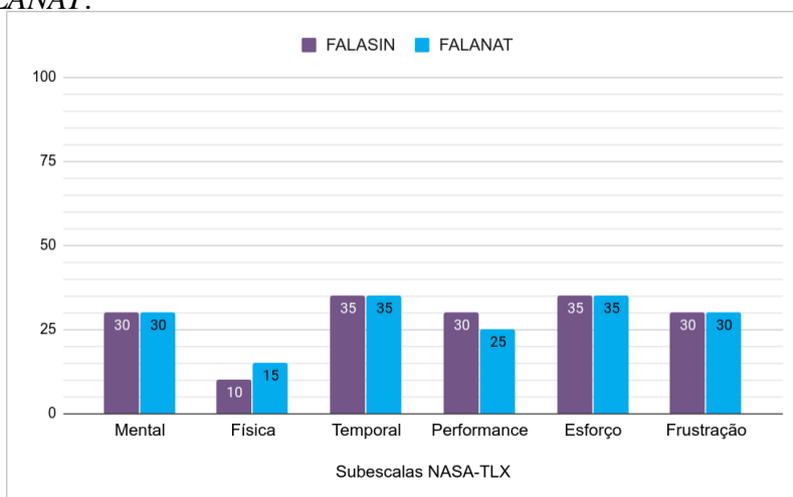
Figura 4.16: Taxas de sucesso nas interfaces *faladas* FALASIN e FALANAT.



Fonte: autor.

As subescalas da carga de trabalho subjetiva também foram equilibradas (Fig. 4.17). As diferenças obtidas na subescala de demanda física e performance não foram significativas, com  $t(5)=0,47$  e  $t(5)=0,33$  respectivamente, ambos  $p > 0,05$ .

Figura 4.17: Carga de trabalho por subescala da NASA-TLX das interfaces *faladas* FALASIN e FALANAT.



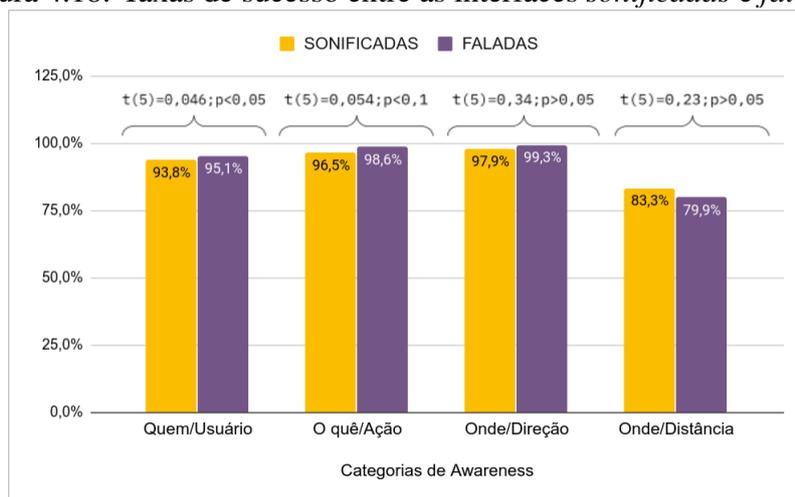
Fonte: autor.

### 4.2.3 Interfaces Sonificadas e Faladas

Os resultados das interfaces *sonificadas* e *faladas* foram agrupados para comparação, isto é, considerou-se os casos de teste para uma dimensão de interfaces SONIFICADAS, enquanto os resultados de FALASIN e FALANAT foram concatenados e rotulados

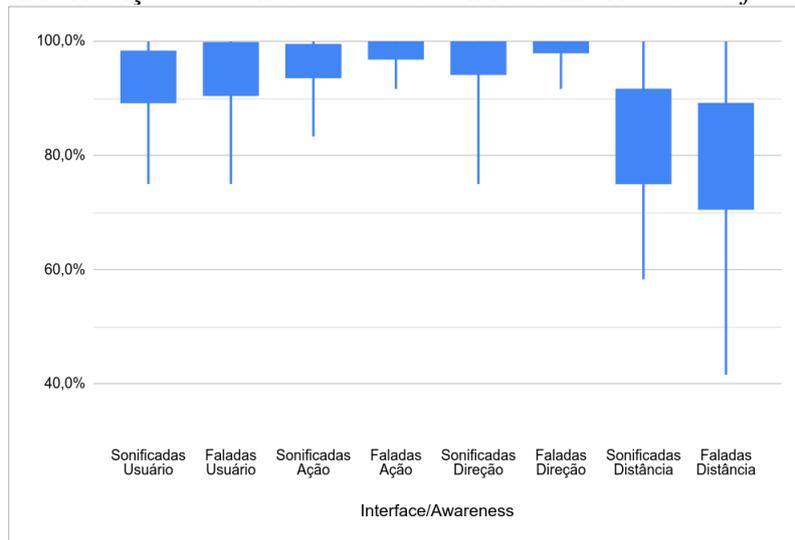
como *FALADAS*. Conforme as Figuras 4.18 e 4.19, os grupos obtiveram taxas de sucesso próximas, no entanto, significativamente melhores para as interfaces *faladas* na percepção das categorias “*quem/usuário*” e “*o quê/ação*”, com 95% e 90% de confiança respectivamente.

Figura 4.18: Taxas de sucesso entre as interfaces *sonificadas* e *faladas*.



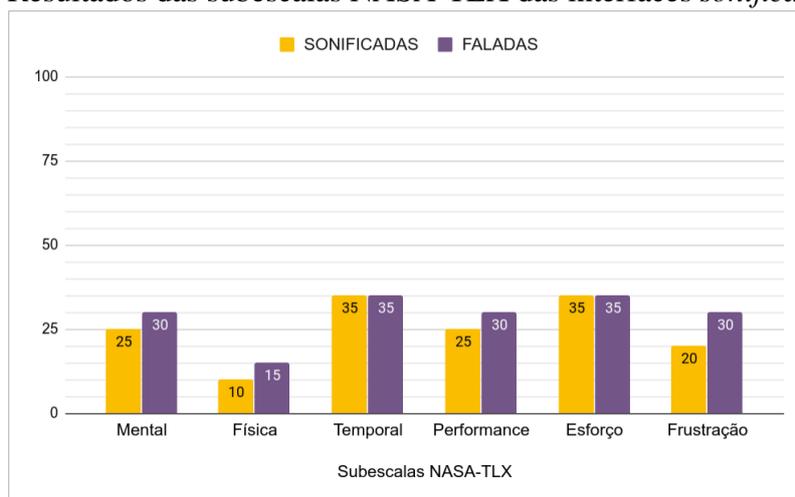
Fonte: autor.

Figura 4.19: Distribuição das taxas de sucesso entre as interfaces *sonificadas* e *faladas*.



Fonte: autor.

As interfaces *sonificadas* tiveram uma carga de trabalho inferior a das interfaces *faladas* na maioria das subescalas (Fig. 4.20), no entanto, os resultados foram dispersos, conforme Tabela 4.2, e não pode-se assumir um nível adequado de confiança.

Figura 4.20: Resultados das subescalas NASA-TLX das interfaces *sonificadas* e *faladas*.

Fonte: autor.

Tabela 4.2: Significância dos resultados do NASA-TLX: interfaces *sonificadas* e *faladas*.

| Subescala          | Sonificada |      | Falada |      | Teste <i>t</i>          |
|--------------------|------------|------|--------|------|-------------------------|
|                    | Média      | DP   | Média  | DP   |                         |
| <i>Mental</i>      | 26,7       | 11,9 | 29,6   | 18,0 | T (11) = 0,67; p > 0,05 |
| <i>Física</i>      | 11,3       | 14,9 | 12,9   | 16,8 | T (11) = 0,76; p > 0,05 |
| <i>Temporal</i>    | 32,5       | 21,1 | 34,6   | 22,1 | T (11) = 0,81; p > 0,05 |
| <i>Performance</i> | 22,5       | 15,2 | 27,5   | 23,8 | T (11) = 0,31; p > 0,05 |
| <i>Esforço</i>     | 33,3       | 22,4 | 35,8   | 21,7 | T (11) = 0,79; p > 0,05 |
| <i>Frustração</i>  | 19,6       | 20,1 | 30,4   | 26,5 | T (11) = 0,23; p > 0,05 |

Fonte: autor.

### 4.3 Discussão dos Resultados

Os resultados foram satisfatórios em ambas interfaces, auditivas e baseadas na fala. A taxa de sucesso média da percepção foi maior que 90%, exceto para a percepção da categoria “*onde/distância*” com 80% (Fig. 4.21). Examinando a carga total de trabalho, segundo a métrica NASA-TLX, nenhuma foi maior que 30 pontos em média (Fig. 4.6), em uma escala até 100. O maior desafio, segundo os resultados das duas fases da pesquisa, foi de comunicar a distância e representar onde uma ação está ocorrendo (Figuras 4.4 e 4.10), que obteve resultados regulares inclusive nas interfaces *faladas*.

Considerando apenas as interfaces auditivas, as mudanças na estratégia para a comunicação de distância, projetadas na segunda etapa, interferiu na performance das outras categorias de *awareness*. Por exemplo, embora a *sonificação* projetada para a percepção do agente, *awareness* de “*quem*”, tenha sido a mesma nas interfaces *SONiv1*, *SONiv2* e *SONiv3*, os resultados foram mais expressivos para a interface *SONiv2*, que reproduz

duas notas na sinalização, menos do que *SONIv3*, que reproduz até sete, e *SONIv1* que reproduz três (Fig. 4.11). São necessários mais testes para compreender esse fenômeno, mas sabe-se de antemão que reproduções mais longas podem reduzir a performance.

Contudo, o *awareness* de “o quê/ação” foi melhor percebido na *SONIv3*. A técnica de *sonificação* para esta categoria foi a mesma nas três interfaces auditivas, a saber, a reprodução constante, crescente e decrescente das notas, indicando alterações, inclusões e exclusões respectivamente. Pode ser que a reprodução de mais notas e alongadas tenham facilitado a identificação da ação, isto é, permitiram a percepção de que as mesmas estão em tom constante, ascendente ou descendente. Os participantes também perceberam com 100% de eficiência a direção na *SONIv3*, e aventa-se que também foi influenciado pelo tempo de reprodução.

Examinando a percepção de distância, a quantidade de níveis perceptíveis eram diferentes nas interfaces *SONIv2* e *SONIv3*. A percepção exata foi maior na *SONIv2*, que reproduzia apenas três níveis com diferentes intensidades sonoras. Pensa-se que o mapeamento mental, a associação do volume do áudio à distância, foi mais simples do que a contagem dos *beeps* exigida na interface *SONIv3*, com até sete níveis e exigia a contagem mental. No entanto, a média de *replays* foi equilibrada (Fig. 4.7), e faz parecer que a noção aproximada era suficiente para os participantes, mesmo na *SONIv3*. Isto é, perceber uma ação como *longe* parece ter sido suficiente em vez de conhecer a quantidade de *links* de distância.

Analisando as respostas do NASA-TLX, os participantes perceberam um melhor rendimento, isto é, uma melhor performance utilizando a interface *SONIv3* (Fig. 4.14). O esforço cognitivo também foi considerado menor na *SONIv3*, mesmo com a contagem de *beeps*. Portanto, pode ser que alongar a reprodução das notas proporcione uma experiência melhor que apenas variar a intensidade sonora.

Comparando as interfaces *sonificadas* desenvolvidas na segunda etapa com a primeira versão, os resultados na percepção do *awareness* de “onde/distância” foram aproximadamente duas vezes melhores (Fig. 4.11). Pode ser que essa diferença tenha sido causada por dois fatores: (1) a mudança de estratégia de *sonificação*, na *SONIv2* para menos níveis de distância e volumes mais espaçados para representá-los, e (2) a possibilidade de repetição da última sinalização.

Considerando as interfaces *faladas*, esperava-se a vantagem para a fala sintetizada, como os usuários estão familiarizados com esse tipo de interface. Todas as informações foram codificadas (exceto a distância) no discurso, e por esse motivo obteve 100% nos

resultados de percepção de usuário, ação e direção. Na opção com voz natural, apenas o *awareness* de “o quê/ação” é comunicado no discurso, que obteve 100% nos resultados, no entanto, mesmo sem serem comunicados no discurso, os participantes puderam identificar o *awareness* de “quem” pelo timbre da voz, e o *awareness* de “onde/direção” pelo efeito estereofônico, ambos resultados comparáveis aos obtidos com a fala sintetizada (Fig. 4.16).

Quanto a avaliação da percepção de distância nas interfaces baseadas em fala, o efeito de desvanecimento da intensidade sonora foi melhor reconhecido na fala sintetizada do que na natural (Fig. 4.16). Esperava-se esse resultado, considerando que na interface projetada com a fala natural (*FALANAT*) a qualidade da gravação das vozes dos coparticipantes depende do *hardware* e condições ambientais, como o tipo de microfone e isolamento acústico, além da própria intensidade aplicada pelo coparticipante no momento da captura. Portanto, isso pode dificultar a percepção de distância por meio do efeito de desvanecimento, já que as vozes foram armazenadas com diferentes intensidades e qualidade geral. Para solucionar esse problema as gravações poderiam ser submetidas a um processo de normalização de áudio, até que a amplitude média das amostras alcançasse um valor normal predefinido, antes de serem armazenadas no banco de enunciados. Estima-se que, entre as opções baseadas em fala, a voz natural seja uma boa opção, por terem enunciados curtos, com percepções comparáveis à fala sintetizada e, especialmente, se considerar que as vozes de conhecidos (colega, professor) do usuário cego possam ser gravadas, pode proporcionar uma melhor experiência geral aos usuários cegos.

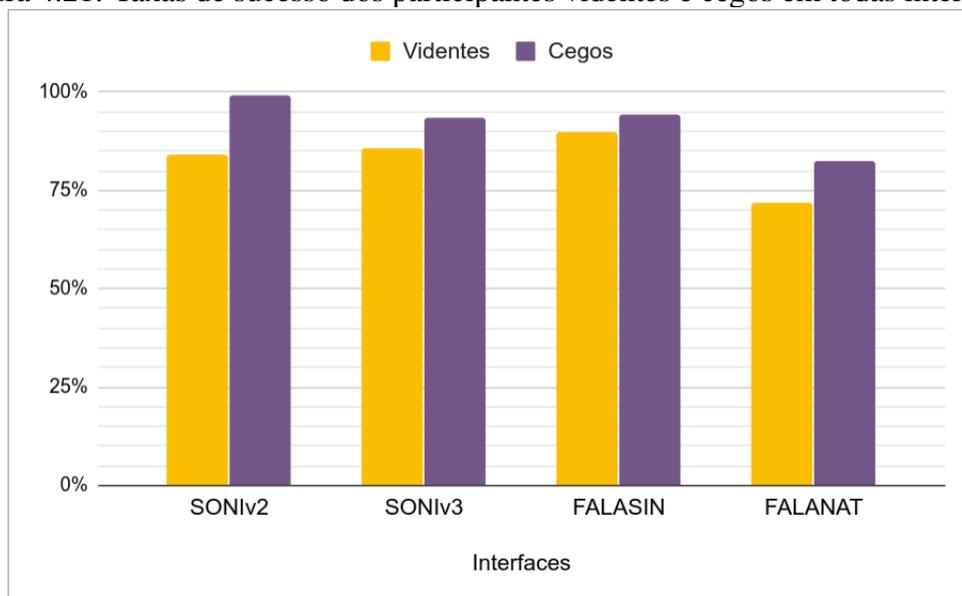
Ao contrapor as interfaces *sonificadas* com as *faladas*, é necessário observar que nas *sonificadas* o usuário precisa lembrar do significado associado ao sinal, enquanto nas *faladas* o significado está no sinal, isto é, o usuário precisa apenas esperar o fim do enunciado para interpretar. Esse efeito é observável, por exemplo, na taxa de repetição, de maneira que nas *faladas* a funcionalidade de *replay* é menos usada (Fig. 4.7). No entanto, embora a interface com fala sintetizada tenha obtido uma maior taxa de sucesso, as interfaces *sonificadas* foram consideradas mais suaves, isto é, os participantes reportaram menor carga de trabalho ao usá-las (Fig. 4.17).

Como opção, pode-se combinar os recursos de *sonificação* e fala para, por exemplo, comunicar informações sobre o *workspace* por meio da reprodução de enunciados e *earcons* conforme os resultados positivos e negativos de cada.

### 4.3.1 Limitações

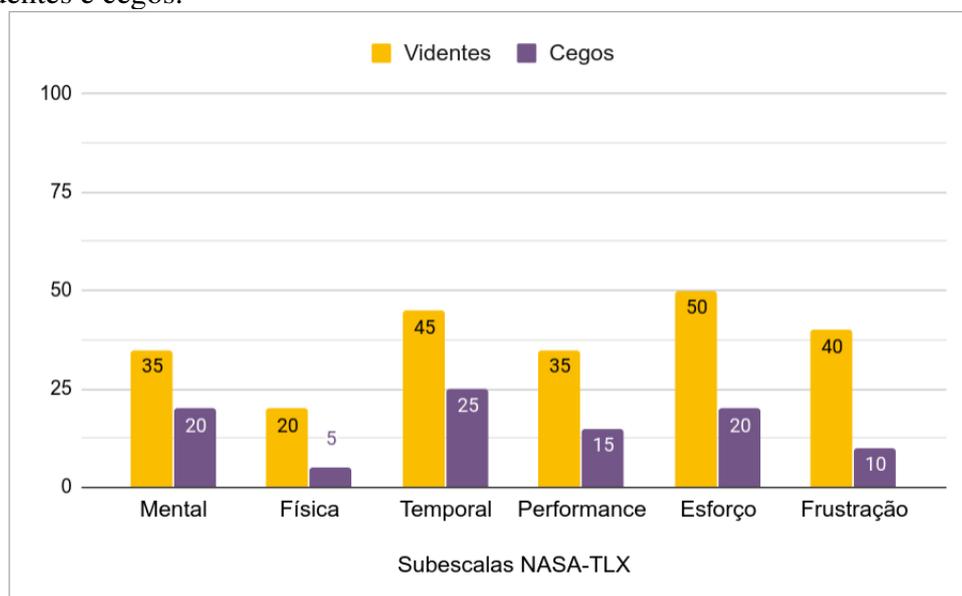
A taxa de sucesso dos participantes cegos nos experimentos foi melhor do que a dos videntes em todas interfaces (Fig. 4.21). Ademais, os participantes cegos reportaram menor carga de trabalho (Fig. 4.22), isto é, a percepção de menor esforço e de que as tarefas foram menos exigentes, comparados aos resultados dos videntes. Essa discrepância pode ter influenciado a dispersão.

Figura 4.21: Taxas de sucesso dos participantes videntes e cegos em todas interfaces.



Fonte: autor.

Figura 4.22: Carga de trabalho por subescala da NASA-TLX comparativo dos participantes videntes e cegos.



Fonte: autor.

O recurso de *replay* estava disponível em todas as interfaces, *sonificadas* e *faladas*, desenvolvidas nessa segunda etapa. Entretanto, o recurso não estava disponível na interface *SONiv1*, projetada e desenvolvida na primeira iteração dessa pesquisa. Assim, os resultados obtidos na segunda etapa podem ter sido influenciados pela disponibilidade desse recurso, entendendo-se como uma vantagem que não estava disponível na primeira, e que pode ter influenciado as taxas de sucesso e carga de trabalho percebida.

Finalmente, deve-se dizer que alguns testes foram elencados, mas não conduzidos nessa pesquisa. Por exemplo, não foram conduzidos casos de teste especiais para a situações em que:

- as reproduções se sobrepõem;
- há mais de três usuários presentes;
- os participantes cometem as ações;
- os coparticipantes estão *on-line*;
- a sessão de modelagem acontece em sala de aula.

## 5 CONCLUSÃO

Este trabalho discutiu como a falta de funcionalidades colaborativas pode restringir o acesso dos estudantes cegos aos recursos pedagógicos relevantes, como a aprendizagem colaborativa, e mitigar as interações sociais com colegas mais experientes e desperdiçar oportunidades de desenvolvimento em potencial. Logo, como pesquisa se propôs projetar uma interface auditiva que tornasse os recursos colaborativos acessíveis aos cegos nas ferramentas de modelagem colaborativa de diagramas.

Concluiu-se que a *sonificação* dos elementos de *workspace awareness* permite implementar uma interface auditiva acessível aos usuários cegos, para que percebam as atividades dos coparticipantes em sessões de modelagem colaborativa e sigam as atualizações das informações diagramáticas, conscientes de quem está agindo, o que está fazendo e onde, num espaço de trabalho compartilhado. O protótipo desenvolvido nesta pesquisa também demonstrou a possibilidade de implementar saídas audíveis acessíveis por meio de linguagens e bibliotecas disponíveis e distribuídas abertamente na plataforma *web*.

Da experiência dos autores nesse projeto e no trabalho com os cegos, percebeu-se que há muita demanda de material adaptado e acessível, sejam estes para uso cotidiano ou didático. Logo, espera-se que os resultados dessa pesquisa auxiliem novos desenvolvimentos na direção de recursos colaborativos acessíveis. Os recursos de *sonificação* projetados e avaliados podem ser usados em vários tipos de diagramas, assim como podem ser adaptados a outros contextos, como ambientes virtuais de aprendizagem, salas e fóruns *on-line*, por exemplo. Toda a implementação foi disponibilizada publicamente como código aberto sob a Licença Pública Geral<sup>1</sup> GNU e está disponível sob a organização do grupo de Informática na Educação na plataforma GitHub neste endereço <<https://github.com/FURG-Infoeduc/audicolabr>>.

### 5.1 Trabalhos Futuros

As interfaces projetadas precisam de mais testes, por exemplo, em situações complexas e de uso intensivo, como sessões de modelagem com vários usuários simultâneos e com mais ações. Também poderiam ser avaliados os efeitos no aprendizado, a experiência em sala de aula, e os efeitos da inclusão. O *workspace awareness* acessível poderia ser estudo do ponto de vista dos professores ou tutores. Ainda, é possível trabalhar em

---

<sup>1</sup>GNU General Public License v3

pesquisas derivadas, tais como:

- outros tipos de *awareness*, como *social*, *group*, *task and informal awareness*;
- a acessibilidade do *awareness* em outros espaços colaborativos, como Ambientes Virtuais de Aprendizagem (AVAs);
- outras técnicas de *sonificação*, como ícones auditivos e *spearcons*;
- interfaces híbridas, que combinam recursos de *sonificação* e fala, ou multimodais mistas com recursos hápticos.

Embora não se tenha concluído um produto final, o protótipo desenvolvido mostrou-se promissor, independente do tipo de recurso audível utilizado. Não foi usado um método formal para a avaliação da satisfação dos usuários, no entanto os participantes cegos demonstraram bastante interesse. Ao fim das sessões foram coletados comentários e sugestões dos participantes, que narraram a sua experiência com a ferramenta. A principal observação foi quanto à surpresa dos participantes ao poder navegar nas conexões do diagrama e perceber o conteúdo. Isto é, os usuário cegos pareciam contentes em perceber e interagir com a informação diagramática sem ser por modo textual, que todos já conheciam por terem estudado dessa forma nas disciplinas em que tiveram contato com diagramas. As sugestões propostas pelos participantes são listadas a seguir:

- *Reproduzir um sinal com 100% de intensidade no início da sessão de modelagem*: um dos participantes sugeriu reproduzir sinalizações de exemplo com 100% do volume logo no início da sessão e antes de qualquer ação, para referência na percepção dos efeitos de desvanecimento posteriores;
- *Manter ambas interfaces disponíveis, sonificada e falada*: todos sugeriram manter a possibilidade de alternar entre as interfaces *sonificadas* e *faladas* para aproveitar as vantagens de cada uma;
- *Maior intervalo entre as notas na interface SONIv3*: um participante sugeriu deixar as notas mais espaçadas na indicação da distância, isto é, alongar o tempo de reprodução e facilitar para o usuário realizar a contagem;
- *Usar o efeito estereofônico mesmo na fala sintetizada*: foi sugerido que, mesmo se direção for narrada (ex.: “tabela à direita”), o sinal fosse reproduzido no canal estéreo correspondente e oferecesse, assim, duas pistas audíveis ao usuário.

## REFERÊNCIAS

- BALAN, O.; MOLDOVEANU, A.; MOLDOVEANU, F. Navigational audio games: an effective approach toward improving spatial contextual learning for blind people. **International Journal on Disability and Human Development**, De Gruyter, v. 14, n. 2, p. 109–118, 2015.
- BALIK, S. M. **Combinatorial Graph Creation and Navigation for Blind People**. Thesis (PhD) — North Carolina State University, 2014.
- BEATTIE, D.; BAILLIE, L.; HALVEY, M. A comparison of artificial driving sounds for automated vehicles. In: **Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing**. [S.l.: s.n.], 2015. p. 451–462.
- BENNETT, D. J. **Presenting diagrams in sound for blind people**. Thesis (PhD) — University of York, 1999.
- BENNETT, D. J. Effects of navigation and position on task when presenting diagrams to blind people using sound. In: SPRINGER. **International Conference on Theory and Application of Diagrams**. [S.l.], 2002. p. 161–175.
- BLATTNER, M. M.; SUMIKAWA, D. A.; GREENBERG, R. M. Earcons and icons: Their structure and common design principles. **Human-Computer Interaction**, Taylor & Francis, v. 4, n. 1, p. 11–44, 1989.
- BORGES, M.; PINO, J. Awareness mechanisms for coordination in asynchronous cscw. In: **Procs. of the 9th Workshop on Information Technologies and Systems**. [S.l.: s.n.], 1999. p. 69–74.
- BORNSCHEIN, J.; PRESCHER, D. Visual to non-visual collaboration on a dynamic tactile graphics display. In: **Collaboration Meets Interactive Spaces**. [S.l.]: Springer, 2016. p. 443–461.
- BRASIL. **Lei 9394 de 20 de dezembro de 1996: estabelece as diretrizes e bases da educação nacional**. 1996. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/Leis/L9394.htm](http://www.planalto.gov.br/ccivil_03/Leis/L9394.htm)>. Acesso em: 22 nov. 2018.
- BRASIL. **Lei 10.098 de 19 de dezembro de 2000: Estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida**. 2000. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/110098.htm](http://www.planalto.gov.br/ccivil_03/leis/110098.htm)>. Acesso em: 22 nov. 2018.
- BRASIL. **Lei 12.711 de 29 de agosto de 2012: dispõe sobre o ingresso nas universidades federais e nas instituições federais de ensino técnico de nível médio**. 2012. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2012/lei/112711.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/112711.htm)>. Acesso em: 22 nov. 2018.
- BRASIL. **Lei 12.409 de 28 de dezembro de 2016: altera a Lei nº 12.711, de 29 de agosto de 2012, para dispor sobre a reserva de vagas para pessoas com deficiência nos cursos técnico de nível médio e superior das instituições federais de ensino**. 2016. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2016/Lei/L13409.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/Lei/L13409.htm)>. Acesso em: 31 jan. 2020.

BREWSTER, S. Nonspeech auditory output. **The human computer interaction handbook: Fundamentals, evolving technologies and emerging applications**, Mahwah, New Jersey, USA: Lawrence Erlbaum Associates, Inc, p. 220–239, 2003.

BRIMHALL, O. D. Exoskeleton with kinesthetic feedback and robotic control. **Acoustical Society of America Journal**, v. 100, p. 2901, 1996.

BROOKSHIRE, R. G. Teaching uml database modeling to visually impaired students. **Issues in Information Systems**, v. 7, n. 1, p. 98–101, 2006.

BROWN, A.; PETTIFER, S.; STEVENS, R. Evaluation of a non-visual molecule browser. In: **Proceedings of the ACM SIGACCESS conference on Computers and accessibility - ASSETS '04**. [S.l.]: ACM Press, 2004.

BROWN, A.; STEVENS, R.; PETTIFER, S. Issues in the non-visual presentation of graph based diagrams. In: IEEE. **null**. [S.l.], 2004. p. 671–676.

BRUFFEE, K. A. Sharing Our Toys: *Cooperative Learning Versus Collaborative Learning*. **Change: The Magazine of Higher Learning**, v. 27, n. 1, p. 12–18, feb. 1995. ISSN 0009-1383, 1939-9146.

BUZZI, M. C. et al. Collaborative Editing: Collaboration, Awareness and Accessibility Issues for the Blind. In: MEERSMAN, R. et al. (Ed.). **On the Move to Meaningful Internet Systems: OTM 2014 Workshops**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. v. 8842, p. 567–573. ISBN 978-3-662-45549-4 978-3-662-45550-0.

CLAUDIA, M.; BUZZI, M.; MORI, B. L. G. Designing E-Learning Collaborative Tools for Blind People. In: PONTES, E. (Ed.). **E-Learning - Long-Distance and Lifelong Perspectives**. [S.l.]: InTech, 2012. ISBN 978-953-51-0250-2.

COHEN, R. F.; MEACHAM, A.; SKAFF, J. Teaching graphs to visually impaired students using an active auditory interface. In: ACM. **ACM SIGCSE Bulletin**. [S.l.], 2006. v. 38, n. 1, p. 279–282.

COLLAZOS, C. et al. Introducing shared-knowledge awareness. In: **IASTED International Conference: Information and Knowledge Sharing**. [S.l.: s.n.], 2002. p. 13–18.

CRYER, H. et al. Teaching stem subjects to blind and partially sighted students: Literature review and resources. **RNIM Centre for Accessible Information**, 2013.

CSAPÓ, Á.; WERSÉNYI, G. Overview of auditory representations in human-machine interfaces. **ACM Computing Surveys (CSUR)**, ACM, v. 46, n. 2, p. 19, 2013.

DAS, M. Understanding Collaborative Writing Practices of People with Visual Impairments. In: **Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers - UbiComp '18**. Singapore, Singapore: ACM Press, 2018. p. 1744–1749.

DAVIDSON, N.; MAJOR, C. H. Boundary crossings: Cooperative learning, collaborative learning, and problem-based learning. **Journal on excellence in college teaching**, v. 25, 2014.

DILLENBOURG, P. **What do you mean by collaborative learning?** [S.l.]: Elsevier, 1999.

DINGLER, T.; LINDSAY, J.; WALKER, B. N. Learnability of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech. In: **INTERNATIONAL COMMUNITY FOR AUDITORY DISPLAY**. [S.l.], 2008.

DOOLITTLE, P. E. Understanding cooperative learning through vygotsky's zone of proximal development. ERIC, 1995.

DOURISH, P.; BELLOTTI, V. Awareness and coordination in shared workspaces. In: **Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW '92**. Toronto, Ontario, Canada: ACM Press, 1992. p. 107–114. ISBN 978-0-89791-542-7.

ELLIS, C. A.; GIBBS, S. J.; REIN, G. Groupware: some issues and experiences. **Communications of the ACM**, ACM, v. 34, n. 1, p. 39–58, 1991.

FREIRE, A. P. et al. Revealing the whiteboard to blind students: An inclusive approach to provide mediation in synchronous e-learning activities. **Computers & Education**, Elsevier, v. 54, n. 4, p. 866–876, 2010.

GAVER, W. Auditory Icons: Using Sound in Computer Interfaces. **Human-Computer Interaction**, v. 2, n. 2, p. 167–177, jun. 1986. ISSN 0737-0024.

GODFREY, A. J. R.; MURRELL, P.; SORGE, V. An accessible interaction model for data visualisation in statistics. In: **Lecture Notes in Computer Science**. [S.l.]: Springer International Publishing, 2018. p. 590–597.

GOKHALE, A. A. Collaborative Learning Enhances Critical Thinking. **Journal of Technology Education**, v. 7, n. 1, sep. 1995. ISSN 1045-1064.

GRILLO, F. D. N. **Uma ferramenta acessível de apoio à modelagem de software na Web**. Thesis (PhD) — Universidade de São Paulo, 2014.

GRUDIN, J. Computer-supported cooperative work: history and focus. **Computer**, v. 27, n. 5, p. 19–26, may 1994. ISSN 0018-9162.

GUTWIN, C.; GREENBERG, S. Workspace awareness for groupware. In: ACM. **Conference Companion on Human Factors in Computing Systems**. [S.l.], 1996. p. 208–209.

GUTWIN, C.; GREENBERG, S. Design for individuals, design for groups: tradeoffs between power and workspace awareness. In: ACM. **Proceedings of the 1998 ACM conference on Computer supported cooperative work**. [S.l.], 1998. p. 207–216.

GUTWIN, C.; GREENBERG, S. A descriptive framework of workspace awareness for real-time groupware. **Computer Supported Cooperative Work (CSCW)**, Springer, v. 11, n. 3-4, p. 411–446, 2002.

GUTWIN, C.; PENNER, R.; SCHNEIDER, K. Group awareness in distributed software development. In: ACM. **Proceedings of the 2004 ACM conference on Computer supported cooperative work**. [S.l.], 2004. p. 72–81.

HART, S. G. Nasa task load index (tlx). volume 1.0; paper and pencil package. 1986.

HERMANN, T. Taxonomy and definitions for sonification and auditory display. In: INTERNATIONAL COMMUNITY FOR AUDITORY DISPLAY. [S.l.], 2008.

HERMANN, T.; HUNT, A.; NEUHOFF, J. G. **The sonification handbook**. [S.l.]: Logos Verlag Berlin, 2011.

HEVNER, A.; CHATTERJEE, S. Design science research in information systems. In: **Design research in information systems**. [S.l.]: Springer, 2010. p. 9–22.

IBGE. **XII Censo Demográfico**. 2010. Disponível em: <<http://www.censo2010.ibge.gov.br/>>. Acesso em: 22 nov. 2018.

INEP. **Resumo Técnico: Censo da Educação Superior de 2010**. Ministério da Educação, 2018. Disponível em: <[http://download.inep.gov.br/download/superior/centso/2010/resumo\\_tecnico\\_censo\\_educacao\\_superior\\_2010.pdf](http://download.inep.gov.br/download/superior/centso/2010/resumo_tecnico_censo_educacao_superior_2010.pdf)>. Acesso em: 28 jan. 2020.

INEP. **Resumo Técnico: Censo da Educação Superior de 2016**. Ministério da Educação, 2018. Disponível em: <[http://download.inep.gov.br/educacao\\_superior/centso\\_superior/resumo\\_tecnico/resumo\\_tecnico\\_censo\\_da\\_educacao\\_superior\\_2016.pdf](http://download.inep.gov.br/educacao_superior/centso_superior/resumo_tecnico/resumo_tecnico_censo_da_educacao_superior_2016.pdf)>. Acesso em: 28 jan. 2020.

JOHANSEN, R. **Groupware: Computer support for business teams**. [S.l.]: The Free Press, 1988.

JOHNSON, D. W.; JOHNSON, R. T.; SMITH, K. A. Cooperative Learning Returns To College What Evidence Is There That It Works? **Change: The Magazine of Higher Learning**, v. 30, n. 4, p. 26–35, jul. 1998. ISSN 0009-1383, 1939-9146.

KENNEL, A. R. Audiograf: A diagram-reader for the blind. In: ACM. **Proceedings of the second annual ACM conference on Assistive technologies**. [S.l.], 1996. p. 51–56.

KING, A. R. **Re-presenting visual content for blind people**. Manchester: University of Manchester, 2007.

KITCHENHAM, B. Procedures for performing systematic reviews. **Keele, UK, Keele University**, v. 33, n. 2004, p. 1–26, 2004.

KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing Systematic Literature Reviews in Software Engineering**. [S.l.], 2007.

KOHLMANN, W. Identifying Barriers to Collaborative Learning for the Blind. In: HUTCHISON, D. et al. (Ed.). **Computers Helping People with Special Needs**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. v. 7382, p. 84–91. ISBN 978-3-642-31521-3 978-3-642-31522-0.

KYSH, L. Difference between a systematic review and a literature review. In: **Medical Library Group of Southern California & Arizona (MLGSCA) and the Northern California and Nevada Medical Library Group (NCNMLG) Joint Meeting**. [S.l.: s.n.], 2013.

LAAL, M.; GHODSI, S. M. Benefits of collaborative learning. **Procedia - Social and Behavioral Sciences**, v. 31, p. 486–490, 2012. ISSN 18770428.

- LAAL, M.; LAAL, M. Collaborative learning: what is it? **Procedia - Social and Behavioral Sciences**, v. 31, p. 491–495, 2012. ISSN 18770428.
- LIPPONEN, L. Exploring foundations for computer-supported collaborative learning. In: **CSCL**. [S.l.: s.n.], 2002. v. 2, p. 72–81.
- LOITSCH, C. et al. UML4all Syntax – A Textual Notation for UML Diagrams. In: MIESENBERGER, K.; KOUROUPETROGLOU, G. (Ed.). **Computers Helping People with Special Needs**. Cham: Springer International Publishing, 2018. v. 10896, p. 598–605. ISBN 978-3-319-94276-6 978-3-319-94277-3.
- LOOMIS, J. M.; LEDERMAN, S. J. Tactual perception. **Handbook of perception and human performances**, Wiley, v. 2, p. 2, 1986.
- LUQUE, L.; BRANDÃO, L.; BRANDÃO, A. Model2gether: uma ferramenta de apoio ao ensino e a aprendizagem de modelos por cegos. In: **Anais dos Workshops do V Congresso Brasileiro de Informática na Educação (CBIE 2016)**. [S.l.]: Sociedade Brasileira de Computação - SBC, 2016.
- LUQUE, L. et al. Can we work together? on the inclusion of blind people in uml model-based tasks. In: **Inclusive Designing**. [S.l.]: Springer, 2014. p. 223–233.
- MACHADO, R. P.; CONFORTO, D.; SANTAROSA, L. Sound chat: Implementation of sound awareness elements for visually impaired users in web-based cooperative systems. In: IEEE. **Computers in Education (SIIE), 2017 International Symposium on**. [S.l.], 2017. p. 1–6.
- MASSIRONI, M. et al. **The psychology of graphic images: Seeing, drawing, communicating**. [S.l.]: Psychology Press, 2001.
- MCINNERNEY, J. M.; ROBERTS, T. S. Collaborative or cooperative learning? **Online Collaborative Learning: Theory and Practice**, p. 203, 2004.
- METATLA, O. Workspace Awareness in Collaborative Audio-Only Interaction with Diagrams. In: **Proceedings of the First African Conference on Human Computer Interaction - AfriCHI'16**. Nairobi, Kenya: ACM Press, 2016. p. 165–169. ISBN 978-1-4503-4830-0.
- METATLA, O. et al. Cross-modal collaborative interaction between visually-impaired and sighted users in the workplace. In: GEORGIA INSTITUTE OF TECHNOLOGY. [S.l.], 2012.
- MILLER, D. **Can We Work Together?** Thesis (PhD) — University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2009. AAI3352649.
- MILLS, K. L. Computer-supported cooperative work. In: CITSEER. **ENCYCLOPEDIA OF LIBRARY AND INFORMATION SCIENCES (2ND EDITION)**. [S.l.], 2003.
- MIURA, T. et al. Diagram presentation using loudspeaker matrix for visually impaired people: Sound characteristics for their pattern recognition. In: **Proceedings of the 7th Augmented Human International Conference 2016**. [S.l.: s.n.], 2016. p. 1–2.

- MOLL, J.; HUANG, Y.; SALLNÄS, E.-L. Audio makes a difference in haptic collaborative virtual environments. **Interacting with Computers**, v. 22, n. 6, p. 544–555, nov. 2010. ISSN 09535438.
- MÜLLER, K. How to make unified modeling language diagrams accessible for blind students. In: SPRINGER. **International Conference on Computers for Handicapped Persons**. [S.l.], 2012. p. 186–190.
- NEES, M. A.; WALKER, B. N. **Auditory Interfaces and Sonification**. 2009.
- NIELSEN, J. Guerrilla hec: Using discount usability engineering to penetrate the intimidation barrier. **Cost-justifying usability**, Academic Press New York, p. 245–272, 1994.
- NIELSEN, J. Success rate: the simplest usability metric. **Jakob Nielsen's Alertbox**, v. 18, p. 3–5, 2001.
- NIELSEN, J. Why you only need to test with 5 users, 2000. **Jakob Nielsen's Alertbox**. Available on line from [www.useit.com/alertbox/20000319.html](http://www.useit.com/alertbox/20000319.html), 2012.
- NUERNBERG, A. H. Contribuições de vigotski para a educação de pessoas com deficiência visual. **Psicologia em estudo**, SciELO Brasil, v. 13, n. 2, p. 307–316, 2008.
- ODOM, S. L. et al. Research in special education: Scientific methods and evidence-based practices. **Exceptional children**, SAGE Publications Sage CA: Los Angeles, CA, v. 71, n. 2, p. 137–148, 2005.
- OKOLO, C. M.; BOUCK, E. C. Research about assistive technology: 2000–2006. what have we learned? **Journal of Special Education Technology**, SAGE Publications Sage CA: Los Angeles, CA, v. 22, n. 3, p. 19–33, 2007.
- OLOFSSON, S. **Designing interfaces for the visually impaired: Contextual information and analysis of user needs**. 2018.
- OMS. **Make listening safe brochure**. 2015. Disponível em: <[https://apps.who.int/iris/bitstream/handle/10665/177884/WHO\\_NMH\\_NVI\\_15.2\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/177884/WHO_NMH_NVI_15.2_eng.pdf)>. Acesso em: 28 jan. 2020.
- ORTIZ, L.; ARISTIZABAL, A.; CARABALLO, F. Uml diagramming software: Definition of accessibility criteria required for the construction of diagrams by visually impaired users. **Actas de Ingeniería**, v. 2, p. 344–351, 2016.
- OWEN, C. B.; COBURN, S.; CASTOR, J. Teaching modern object-oriented programming to the blind: an instructor and student experience. In: **ASEE Annual Conference**. [S.l.: s.n.], 2014.
- OXFORD, R. L. Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. **The Modern Language Journal**, Wiley Online Library, v. 81, n. 4, p. 443–456, 1997.
- OXFORD, R. L. Cooperative Learning, Collaborative Learning, and Interaction: Three Communicative Strands in the Language Classroom. **The Modern Language Journal**, v. 81, n. 4, p. 443–456, dec. 1997. ISSN 00267902.

PANITZ, T. **Collaborative versus Cooperative Learning: A Comparison of the Two Concepts Which Will Help Us Understand the Underlying Nature of Interactive Learning**. [S.l.]: ERIC, 1999.

PANSANATO, L. T. E. et al. Projeto d4all: acesso e manipulação de diagramas por pessoas com deficiência visual. In: BRAZILIAN COMPUTER SOCIETY. **Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems**. [S.l.], 2012. p. 33–36.

PASCOLINI, D.; MARIOTTI, S. P. Global estimates of visual impairment: 2010. **British Journal of Ophthalmology**, BMJ Publishing Group Ltd, v. 96, n. 5, p. 614–618, 2012.

PEFFERS, K. et al. A design science research methodology for information systems research. **Journal of management information systems**, Taylor & Francis, v. 24, n. 3, p. 45–77, 2007.

PREECE, J. et al. **Human-computer interaction**. [S.l.]: Addison-Wesley Longman Ltd., 1994.

RAJAN, R. et al. “roger that!”—the value of adding social feedback in audio-mediated communications. In: SPRINGER. **IFIP Conference on Human-Computer Interaction**. [S.l.], 2013. p. 471–488.

RAMA, J.; BISHOP, J. A survey and comparison of cscw groupware applications. In: SOUTH AFRICAN INSTITUTE FOR COMPUTER SCIENTISTS AND INFORMATION TECHNOLOGISTS. **Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries**. [S.l.], 2006. p. 198–205.

SÁNCHEZ, J.; BALOIAN, N. Issues in implementing awareness in collaborative software for blind people. In: SPRINGER. **International Conference on Computers for Handicapped Persons**. [S.l.], 2006. p. 1318–1325.

SANTOS, L. G. dos et al. Recursos de acessibilidade para auxiliar a navegação de estudantes cegos em um editor de diagramas. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. [S.l.: s.n.], 2012. v. 23, n. 1.

SCHMIDT, K.; BANNON, L. Taking CSCW seriously: Supporting articulation work. **Computer Supported Cooperative Work (CSCW)**, v. 1, n. 1-2, p. 7–40, mar. 1992. ISSN 0925-9724, 1573-7551.

SCHOEBERLEIN, J. G.; WANG, Y. Accessible collaborative writing for persons who are blind: a usability study. In: **Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '12**. Boulder, Colorado, USA: ACM Press, 2012. p. 267. ISBN 978-1-4503-1321-6.

SILVA, C. E.; PANSANATO, L.; FABRI, J. A. Ensinando diagramas uml para estudantes cegos. **XVIII CIESC–XXXVI CLEI**, 2010.

SMITH, B. L.; MACGREGOR, J. T. **What is collaborative learning**. 1992.

SILVA, C. E.; PANSANATO, L.; FABRI, J. A. Ensinando diagramas uml para estudantes cegos. **XVIII CIESC–XXXVI CLEI**, 2010.

SMITH, B. L.; MACGREGOR, J. T. **What is collaborative learning**. 1992.

SONZA, A. P. et al. Acessibilidade e tecnologia assistiva: pensando a inclusão sociodigital de pessoas com necessidades especiais. **Bento Gonçalves: BBB**, 2013.

SPRINGER, L.; STANNE, M. E.; DONOVAN, S. S. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. **Review of educational research**, Sage Publications, v. 69, n. 1, p. 21–51, 1999.

STAHL, G.; KOSCHMANN, T.; SUTHERS, D. D. Computer-supported collaborative learning: An historical perspective. In: . [S.l.: s.n.], 2006.

STOLL, E. L. An introduction to collaborative learning: From theory to application. **Communication Education**, Taylor & Francis, v. 45, n. 3, p. 260–263, 1996.

SYAL, P.; CHATTERJI, S.; SARDANA, H. K. DiGVis: A system for comprehension and creation of directed graphs for the visually challenged. **Universal Access in the Information Society**, v. 15, n. 2, p. 199–217, jun. 2016. ISSN 1615-5289, 1615-5297.

TAYLOR, P. **Text-to-speech synthesis**. [S.l.]: Cambridge university press, 2009.

THAPA, R. B.; FERATI, M.; GIANNOUMIS, G. A. Using non-speech sounds to increase web image accessibility for screen-reader users. In: **ACM. Proceedings of the 35th ACM International Conference on the Design of Communication**. [S.l.], 2017. p. 19.

TURNER, C. W.; LEWIS, J. R.; NIELSEN, J. Determining usability test sample size. **International encyclopedia of ergonomics and human factors**, Citeseer, v. 3, n. 2, p. 3084–3088, 2006.

UN, U. N. **Convention on the Rights of Persons with Disabilities Article 24**. 2007. Disponível em: <<https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-24-education.html>>.

VYGOTSKI, L. **Obras Escogidas de Vygotski - V: Fundamentos de defectología**. [S.l.]: Antonio Machado Libros, 2015. (Machado Nuevo Aprendizaje). ISBN 9788491140894.

VYGOTSKY, L. S. **Mind in society: The development of higher psychological processes**. [S.l.]: Harvard university press, 1980.