

Proposta e Avaliação de uma Ferramenta Web para Vocalização de Gráficos de Barras

Wallace Pacheco Pereira
Universidade Federal do Pará
Belém, Brasil
wallacecovic@gmail.com

Cynthya Letícia Teles de
Oliveira
Universidade de São Paulo
São Carlos, Brasil
cynthya@usp.br

Marcelle Pereira Mota
Universidade Federal do Pará
Belém, Brasil
mpmota@ufpa.br

ABSTRACT

Charts are visual representations of numerical data to transmit information efficiently, facilitating understanding. However, the charts present on the internet often do not have textual descriptions that screen reader programs can process. Alternatively, it is incomplete for understanding when this information exists and, consequently, excludes people, such as the visually impaired. This article presents a web tool that vocalizes the textual description of simple and grouped bar charts, using validated description models, one for each chart type. The evaluation verified the proposal's feasibility to help people understand the information in the charts through speech synthesis. Tests were applied with participants without visual impairment and with visual impairment. The analysis of the collected data shows a positive evaluation of the users about using the tool.

KEYWORDS

Gráfico de Barras, Descrição Textual de Gráficos, Síntese de Voz, Acessibilidade, Avaliação com Usuários.

1 INTRODUÇÃO

Os gráficos têm a capacidade de apresentar sinteticamente um enorme volume de dados. Deste modo, a justificativa para seu uso deve-se à eficiência para transmitir visualmente informações de forma resumida. Assim, os gráficos são muito utilizados em diferentes áreas e contextos no cotidiano, como artigos científicos, revistas, jornais e sites. Com a expansão da internet, associações de diferentes tipos de segmentos da sociedade começaram a divulgar vários tipos de informações na rede mundial de computadores. Entretanto, recursos visuais, como figuras e gráficos, são apresentados na internet, muitas das vezes, sem recurso de acessibilidade, excluindo pessoas com diferentes problemas visuais [1].

No Brasil, de acordo com o último censo realizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE), realizado em 2010, 6,6 milhões de pessoas declararam que apresentam algum tipo de deficiência visual severa, ou grande dificuldade de enxergar, ou que não conseguem de modo algum. Dentre o total dessa população, 506,3 mil declararam serem cegos [2]. Neste contexto, torna-se primordial a utilização de Tecnologia Assistiva (TA) que, conforme o Comitê de Ajudas Técnicas (CAT) em conjunto com a Secretaria Especial dos Direitos Humanos da Presidência da República (SEDR/PR), estabeleceu o conceito brasileiro de Tecnologia Assistiva como “uma área do conhecimento, de característica interdisciplinar, que engloba produtos, recursos, metodologias, estratégias, práticas e serviços que objetivam promover a funcionalidade, relacionada à atividade

e participação, de pessoas com deficiência, incapacidades ou mobilidade reduzida, visando sua autonomia, independência, qualidade de vida e inclusão social.” [3].

Conforme Kleina [4], Tecnologia Assistiva pode sugerir uma variedade de tecnologia para ajudar os deficientes visuais a ter maior possibilidade de qualidade de vida e inclusão social, por meio da expansão de sua comunicação, mobilidade, suas habilidades de aprendizagem e relações com outros indivíduos. Um exemplo de Tecnologia Assistiva são softwares leitores de telas [5–8], que são essenciais para pessoas com deficiência visual, analfabetas ou com dificuldades de aprendizagem, e também auxiliam no manuseio de dispositivos tecnológicos fornecendo informações por meio de síntese de voz. Porém, muitas vezes os recursos visuais não possuem descrições textuais que possam ser processadas pelos leitores de tela ou, quando possuem, as informações são incompletas para a compreensão. Assim, não basta adicionar somente textos alternativos às figuras, esses textos devem transmitir uma comunicação compreensiva para seus usuários [9]. Quando não existe descrição textual de imagens de gráficos, utilizam-se outros recursos para extração de dados de forma automática, como técnicas de visão computacional e de reconhecimento óptico de caracteres, que permitem armazenar os dados em um tipo de formato de arquivo [10]. A partir dos dados extraídos do gráfico, sugere-se a criação de um modelo de descrição textual com essas informações para serem repassadas ao usuário, por meio de síntese de voz, para a compreensão do recurso visual.

Este trabalho apresenta um módulo de uma ferramenta web que vocaliza uma descrição textual para gráfico de barras simples e agrupadas. Esses gráficos foram escolhidos por serem os instrumentos estatísticos mais utilizados em documentos, principalmente no meio acadêmico [11] e na internet [12]. Assim, o objetivo deste trabalho foi propor e avaliar uma ferramenta para auxiliar deficientes visuais a terem um bom entendimento das informações de gráficos de barras, por meio da vocalização de uma descrição textual. A implementação do módulo web adotou dois modelos de descrições textuais para vocalização, e as linguagens de programação utilizadas foram Python, HTML, CSS e JavaScript. Após a implementação, foram aplicados testes utilizando o protótipo da ferramenta com participantes com e sem deficiência visual. Por fim, foi realizada análise quantitativa e qualitativa dos dados, os resultados mostraram uma avaliação da positiva sobre a ferramenta.

O trabalho está organizado em oito seções: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 apresenta os modelos de descrição textual; a Seção 4 apresenta uma ferramenta web de tecnologia assistiva; a Seção 5 apresenta os testes realizados; a Seção 6 apresenta as descobertas da pesquisa; e a Seção 7 apresenta as considerações finais e trabalhos futuros.

2 TRABALHOS RELACIONADOS

Na literatura foram encontrados trabalhos que transformam dados visuais em recursos acessíveis para pessoas sem ou com algum tipo de deficiência. Como exemplo, o trabalho de [13] que desenvolveu uma aplicação que permite a interação com os usuários, com e sem deficiência visual, por meio do teclado, e o *feedback* da aplicação utiliza voz sintética. Essa ferramenta é composta por módulos de extração, geração de informação e interface para facilitar a interação com o usuário. Aplicou-se testes com gráficos de linha tirados de um jornal, e a ferramenta descrevia informações desses gráficos para os participantes que eram leitores desse noticiário. Também cita-se o trabalho de [14], com a proposta de desenvolver um sistema para pessoas com deficiência visual que melhorava o acesso sobre documentos multimodais contendo gráficos de linha. O sistema tinha a finalidade de ajudar o leitor de tela a reconhecer e transmitir, de forma resumida, as informações de gráficos presentes em documentos.

Observa-se que existem diferentes metodologias capazes de extrair e melhorar a qualidade de difundir dados informativos, de uma forma mais compreensível. Nesse contexto, o trabalho de [15] propôs a utilização de uma linguagem controlada em sua aplicação, para verificar a adequação de textos alternativos para imagens gráficas de dados numéricos, levando em consideração um conjunto de regras orientadas à acessibilidade ajustado ao idioma francês. Já o aplicativo proposto por [16] converte automaticamente imagens no formato vetorial SVG (*Scalable Vector Graphics*) em uma descrição textual para usuários com deficiência visual, utilizando síntese de voz, empregando do lado do cliente um TTS (*Text-to-Speech*), para que a descrição da imagem pudesse ser lida pelo aplicativo e para realizar testes com usuários, comparando descrições produzidas automaticamente. Há também métodos não automatizados que possibilitam a interatividade entre usuários, [17] listou uma série de aplicativos para cegos e destacou o "Be My Eyes", um aplicativo que permite pessoas sem deficiência visual a ajudarem pessoas cegas que precisam da descrição de imagens, gráficos, tabelas, dentre outras possibilidades, por meio de chamada de vídeo.

Verifica-se na literatura consultada diversas propostas de trabalhos direcionadas para pessoas com deficiência visual, bem como o desenvolvimento de aplicações voltadas em tornar dados visuais mais acessíveis para o público em geral, utilizando métodos de descrição textual e síntese de voz, podendo ser de forma automática ou manual, essas metodologias são base para o progresso desta pesquisa. O presente artigo apresenta a implementação de um módulo da ferramenta web, modelos validados aplicados de descrição textual por voz sintética no idioma português brasileiro, para gráficos de barras simples e agrupadas, procedimentos de testes com o público-alvo, assim como, avaliação da experiência dos usuários com o intuito de melhorar a ferramenta e padronizar modelos textuais, adotando uma abordagem diferente dos trabalhos citados anteriormente.

3 MODELOS DE DESCRIÇÃO TEXTUAL PARA VOCALIZAÇÃO DE GRÁFICOS DE BARRAS

O modelo de descrição textual dos gráficos utilizado nesse trabalho, foi proposto e validado por [18], que em sua pesquisa realizaram

diversos testes com usuários. A seguir são apresentados os modelos para gráficos de barras simples e agrupadas, respectivamente:

- Este é um gráfico de barras verticais. Seu título é {título}. A legenda do eixo y é {rótulo do eixo y}. A legenda do eixo x é {rótulo do eixo x}. Os valores das barras variam entre {valor de mínimo} e {valor de máximo}, e tem cor {cor}. A primeira barra é {nome da primeira barra} e possui valor {valor da primeira barra}. (...) A n-ésima barra é {nome da n-ésima barra} e possui valor {valor da n-ésima barra}. A média dos valores é {média}. O desvio padrão é {desvio padrão}.
- Este é um gráfico de barras verticais agrupadas. Seu título é {título}. A legenda do eixo y {rótulo do eixo y}. A legenda do eixo x {rótulo do eixo x}. O gráfico é composto pelos grupos de barras {nomes dos grupos}. Cada grupo contém {quantidade de barras do grupo} barras: {nomes das barras}, que serão apresentadas nas cores {cor da primeira série}, (...) e {cor da n-ésima série}, nessa ordem. Os valores das barras variam entre {valor de mínimo} e {valor de máximo}. O primeiro conjunto de barras é {nome do primeiro grupo} e possui valores {valor da primeira barra}, (...), {valor da n-ésima barra}. (...) O enésimo conjunto de barras é {nome do n-ésimo grupo} e possui valores {valor da primeira barra}, (...), {valor da n-ésima barra}. A média dos valores é {média}. O desvio padrão é {desvio padrão}.

A ferramenta proporciona vocalizar todo o texto de forma contínua em único áudio, ou cada elemento separado. O usuário pode interagir e escolher o dado de acordo com as funções dos comandos por meio do teclado.

4 FERRAMENTA WEB

O presente trabalho utiliza arquitetura de descrição automática de gráficos de dados mencionada por [18], que propôs criar e avaliar modelos de descrição textual de gráficos de barras simples e agrupadas utilizando síntese de voz. Se espera que tenha ocorrido uma fase prévia de leitura do gráfico e geração de um documento em formato JSON que contenha as informações do gráfico separadas, elemento por elemento, isso deve ter sido feito de modo automatizado com ferramentas de processamento de imagens baseadas em visão computacional.

Assim, com os elementos do gráfico organizados no arquivo JSON, foi criado um script em Python para criar arquivos MP3 com a voz sintetizada reproduzindo a descrição textual dos gráficos, baseadas nos modelos de barras simples e agrupadas, de acordo com o tipo de gráfico. Em outras palavras, o script Python, tem a tarefa de ler o arquivo com a informação do modelo de descrição textual do gráfico de barras, enviado do módulo de extração de dados em formato JSON. Como os dados desse arquivo são textos agrupados em uma grande cadeia de caracteres em uma única linha, o script organiza essa cadeia, dividindo o texto todas as vezes que encontra um ponto no final de frase passada como argumento, assim, criando uma outra cadeia de caracteres em uma nova linha. Logo após esse processo, inicia a vocalização por síntese de voz utilizando GTTS (*Google Text-to-Speech*), que é uma biblioteca Python que permite gravar dados de arquivos em áudio (formato MP3). Inicialmente, é gravado o áudio de todo o texto do arquivo e, depois, por meio

de uma estrutura de repetição, é gravado o áudio de cada linha do arquivo equivalente a cada um dos elementos do gráfico.

Os áudios gravados são salvos para serem referenciados na página web aberta no navegador, esta página é gerada automaticamente para cada gráfico de entrada e utiliza HTML, CSS e JavaScript. Como a ferramenta vocaliza a descrição textual de gráficos de barras simples e agrupadas, temos dois modelos de páginas web que podem ser gerados e comandos específicos para cada tipo de gráfico. O código da página utiliza JavaScript para responder a ação do usuário por meio do teclado usando o conceito de eventos, onde cada evento espera que uma tecla seja pressionada para tocar um arquivo de áudio. Optou-se por um design mais simples e minimalista, e todos os comandos nas páginas são reconhecidos por software leitores de tela.

5 TESTES

Foram realizados testes com usuários seguindo orientações de [19] para preparação, escolha de participantes, definição de tarefas a serem realizadas, coleta de dados, interpretação e relato dos resultados. Além disso, foi utilizada metodologia similar de testes realizados por [18] que avaliou modelos de descrição textuais de gráficos de barras.

5.1 Participantes

Participaram dos testes 20 discentes da Universidade Federal do Pará, sem deficiência, sendo 10 discentes da área de ciências exatas e 10 discentes da área da saúde. Além destes, também participaram 3 deficientes visuais integrantes da Associação de Discentes com Deficiência da Universidade Federal do Pará, sendo 2 discentes da área de exatas com cegueira, e 1 discente da área de ciências humanas com baixa visão. Todos os 23 participantes afirmaram ter conhecimento sobre gráficos de barras.

5.2 Procedimento

Foram escolhidos para a realização dos testes os modelos de descrição textual completos, com informações de gráficos de barras simples e agrupados para serem vocalizados, assim como, comandos utilizados na interação por meio do teclado. Posteriormente, foram definidas tarefas para serem realizadas pelos participantes, como ouvir os áudios das descrições textuais por comandos do teclado. Em seguida, o participante deveria desenhar o gráfico com base nessas descrições e explicar o gráfico desenhado. Vale ressaltar que os voluntários com deficiência visual não executavam a tarefa de desenhar os gráficos, a tarefa definida para esse grupo era explicar o gráfico com base nos áudios com as descrições textuais, e depois era apresentado o gráfico em Braille, permitindo ao usuário comparar e emitir sua opinião sobre a descrição do gráfico. Por fim, todos os participantes eram entrevistados para opinar sobre a ferramenta, as entrevistas foram registradas com um gravador de áudio. Para coletar de dados, além dos desenhos feitos pelos voluntários, foram feitas anotações observando o usuário interagindo com a ferramenta para identificar características de interação. Os gráficos utilizados nos testes podem ser visualizados nas Figuras 1 e 2, e os comandos que podiam ser utilizados pelos usuários podem ser verificados nas Tabelas 1 e 2.

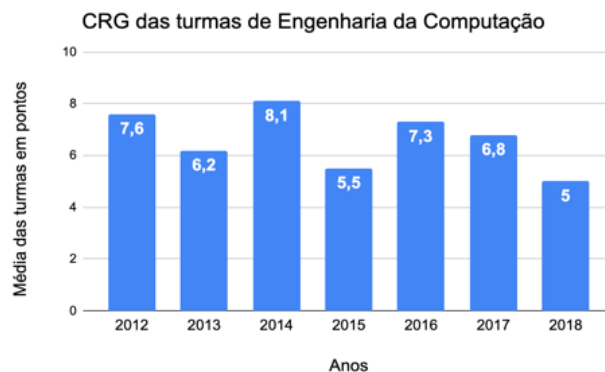


Figura 1: Gráfico A utilizado nos testes

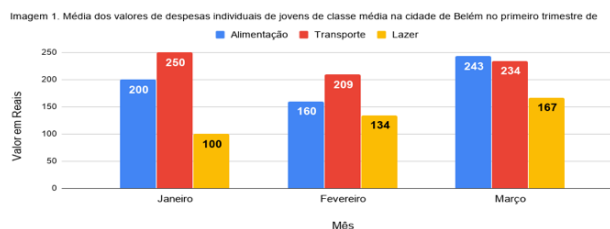


Figura 2: Gráfico B utilizado nos testes

Tabela 1: Comandos utilizados no gráfico A

Comando	Tecla
Ler descrição completa	D
Ler categoria	C
Ler título	T
Ler legenda do eixo y	Y
Ler legenda do eixo x	X
Ler valor mínimo e máximo	M
Ler dados da primeira barra	1
Ler dados da segunda barra	2
Ler dados da terceira barra	3
Ler dados da quarta barra	4
Ler dados da quinta barra	5
Ler dados da sexta barra	6
Ler dados da sétima barra	7
Ler média e desvio padrão	E

Para aplicação do teste, primeiramente, foi explicado quais tarefas o voluntário deveria realizar, como ouvir um conjunto de áudios de descrições textuais utilizando os comandos expostos na tela do computador com o leitor de tela NVDA (*NonVisual Desktop Access*), executado em ambientes *Windows*, utilizado no teste para ler os comandos da ferramenta na página web no navegador *Google Chrome* e interagindo por meio do teclado, executando comandos no máximo duas vezes para ouvir cada um dos áudios. O objetivo era que o participante desenhasse o gráfico, se baseando na descrição ouvida nos áudios, em uma folha de papel entregue para ser preenchido com as informações. Ressaltando que a tarefa de desenhar foi para pessoas sem deficiência. Após a explicação das tarefas, foi ligado o gravador e iniciada a entrevista. Quando o voluntário terminou de

Tabela 2: Comandos utilizados no gráfico B

Comando	Tecla
Ler descrição completa	D
Ler categoria	C
Ler título	T
Ler legenda do eixo y	Y
Ler legenda do eixo x	X
Ler grupos e séries	G
Ler valor mínimo e máximo	M
Ler dados do primeiro grupo	1
Ler dados do segundo grupo	2
Ler dados do terceiro grupo	3
Ler média e desvio padrão	E

desenhar o primeiro gráfico de barras simples, foi pedido para ele explicar o seu desenho com o máximo de detalhes e foi perguntado se ele teria alguma informação que ele considerava essencial para o entendimento do gráfico e, se não estivesse disponível, que ele apontasse qual informação seria. Esse mesmo procedimento foi feito para o gráfico de barras agrupadas.

6 RESULTADOS

A partir dos dados coletados, foi feita uma análise individual de cada participante destacando erros e observações, logo após, foi comparada cada informação destacada com o objetivo de buscar uma recorrência para ter uma análise geral de todo o conteúdo verificado. Também, foram adotadas siglas, como PE (participante da área de exatas), PS (participante de área da saúde), PD (participante com deficiência) e um número para diferenciar cada indivíduo na pesquisa.

6.1 Análise quantitativa

O grupo de participantes com deficiência visual terminou o teste com o tempo médio aproximado de 14,3 minutos, e o grupo dos discentes sem deficiência visual terminou com o tempo médio de aproximadamente 16,7 minutos. O limite inferior de tempo de realização foi de 11 minutos, realizado por um aluno de exatas, e o limite superior foi de 23 minutos, realizado por um discente da saúde. Observou-se que os voluntários que finalizaram em menor tempo apresentaram pequenos erros, logo, tiveram uma explicação mais curta do seu desenho. Ao contrário, os participantes que finalizaram com maior tempo de duração, desenharam os gráficos com todos os elementos e explicaram com mais detalhes, porém nenhum participante deixou de falar dados dos gráficos.

Notou-se que cerca de 80% dos participantes seguiram a ordem dos comandos de cima para baixo, começando com a descrição completa e terminando no comando de desvio padrão, principalmente, o grupo com deficiência visual onde todos seguiram essa ordem por conta do leitor de tela ler dessa forma. Para o gráfico A, aproximadamente 58% dos participantes não utilizaram a possibilidade de reproduzir todos os áudios duas vezes. Já com o gráfico B, a maioria, 85%, utilizaram duas vezes os comandos para ouvir as descrições textuais. O grupo de discentes da área de saúde cometeu um pouco mais de erros no desenho dos gráficos, porém todos desenharam e explicaram os gráficos de acordo com sua percepção (ver Fig. 3).

Em relação ao desenho dos gráficos, cerca de 95% dos participantes desenharam o gráfico de barras simples com todos os elementos

esperados, baseados na descrição textual informada. Apenas uma pessoa desenhou sem os nomes dos eixos X e Y, e sem os valores das barras. Porém, a ausência dos valores foi em decorrência das marcações do eixo Y definida no papel de desenho (ver Fig. 4). É possível observar que o participante PS10 cometeu o mesmo erro do gráfico de barra simples e não conseguiu anotar a legenda completa. Outros participantes confundiram os valores de máximo e mínimo, pois anotaram somente um valor para grupo, e essa informação não estava na descrição textual, logo, tinha que ser percebida verificando os dados. Em outro caso, PS9 não escreveu corretamente o título do gráfico no desenho do gráfico B (ver Fig. 5)

Porém, nos desenhos do gráfico de barras agrupadas, por conterem mais informações, aconteceram mais erros, a acurácia foi de 75%. O erro mais ocorrido foi não anotar toda legenda do gráfico, PE8 justificou esse erro declarando: “*eu achei rápido, deveria ser mais pausado, porque não dá tempo de escutar e escrever tudo (...)*”. O segundo equívoco mais recorrente foi confundir valores de mínimo e máximo. Os erros e as quantidades estão na Fig. 6.

Ao final da entrevista, após todo o processo do teste, os participantes atribuíram uma nota de 1 a 10 indicando a sua avaliação e o grau de satisfação em relação a ferramenta utilizada. A Fig. 7 ilustra o bom resultado alcançado pela ferramenta.

6.2 Análise qualitativa

Nessa análise, foram destacados comentários dos participantes a respeito da interpretação dos gráficos e avaliação dos usuários sobre a experiência utilizando a ferramenta, principalmente do grupo de deficientes visuais, para pensar em possíveis melhorias para desenvolver um melhor design da interação e novas funcionalidades para esse público.

6.2.1 Interpretações do Gráfico A. Sobre o entendimento do gráfico A, voluntários que finalizaram mais rápido as tarefas deram uma explicação mais breve sobre o gráfico, como PS10 que falou do gráfico A: “*Esse aqui é um gráfico de barras simples que mostra o CRG das turmas de engenharia da computação de cada ano a partir de 2012 até 2018. Mostra o mínimo e o máximo, média e desvio padrão*”, essa foi a fala mais breve. Também teve uma explicação muito detalhada sobre esse gráfico como a do participante PE2 que disse: “*Bom, aqui analisado é o CRGs das turmas de engenharia da computação. Tem como interesse a variável Y, média das turmas, e ele tem a média das distribuições por anos. E a gente vê que, a partir de 2012, foi de 7,6, em 2013 teve uma queda, teve seu máximo em 2014. Enfim, depois desceu em 2015 e foi indo, em 2018 teve seu menor CRG e olhando mais algum detalhe aqui, nós vimos que os valores que foram acima da média dessas notas ‘foi’ os anos de 2021, 2014, 2016 e 2017, e os demais anos foram abaixo da média (...)*”. Assim como o grupo de pessoas cegas que também teve uma excelente explicação e análise do conteúdo apresentado nesse primeiro gráfico, o participante PD1 comentou: “*Entendi que no eixo de X eu tenho os anos, e de Y a variação dos CRGs medidos em média para cada ano e cada barra marca o ano e a média dos CRGs dos alunos naquele ano. Eu consigo entender ‘cê’ tá medindo a média dos CRGs dos alunos e consigo até te dizer que no ano tal teve uma queda acentuada para cinco ponto alguma coisa, sendo que no ano anterior era de 8 (...)*”. O participante PD3 também falou de uma forma bem detalhada, ele disse: “*É um gráfico de coluna né, apresenta o CRGs das turmas de engenharia da*



Figura 3: Tempo de teste em minutos e erros por participante

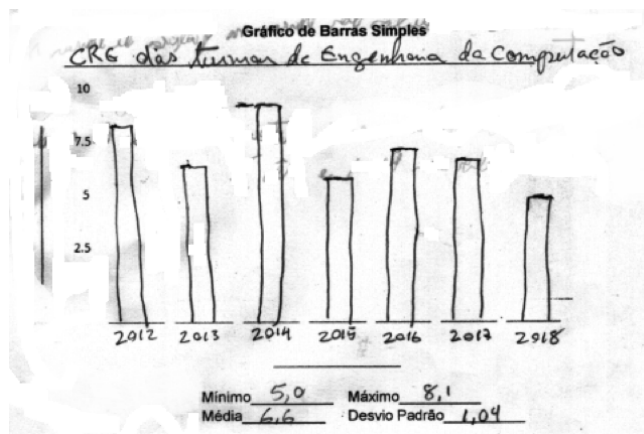


Figura 4: Gráfico de barras simples do participante PS10

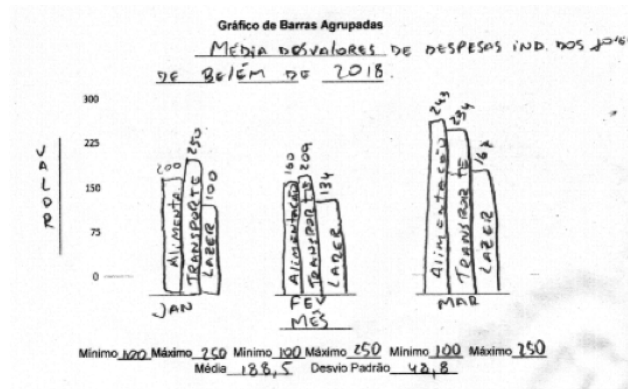


Figura 5: Gráfico de barras agrupadas do participante PS9

computação, e o maior valor foi em 2014 das médias da turma 8,1. O eixo vertical apresenta as médias das turmas e o eixo horizontal os anos, e ele segue uma ordem de anos que vai de 2012 até 2018, o valor mínimo é 5 e o máximo é 8,1, a média é 6,6 e o desvio padrão é 1,04.”

Essas falas demonstram que todos os grupos tiveram um bom entendimento do gráfico A. Quando perguntados se faltava alguma informação, todos responderam que não, e aprovaram a descrição

textual e a forma que ela é passada pelo áudio. PD2 falou: “Não, acho que para esse tipo de gráfico, não. A gente consegue ter leituras críticas do gráfico, principalmente, nesse recurso de clicar 1 e ler a barra 1, clicar 2 ler a barra 2, e assim vai... entendeu?”.

6.2.2 *Interpretações do gráfico B.* Apesar de o gráfico B ser mais difícil de interpretar do que o gráfico A, todos tiveram uma boa compreensão. Porém, fizeram algumas observações sobre a descrição. O PE4 fez a seguinte explicação: “Aqui, é um gráfico de barras agrupadas, fala sobre os valores das despesas dos jovens durante os meses de janeiro, fevereiro e março, com alimentação, transporte e lazer. Em janeiro, eles gastaram mais com transporte e menos com lazer, em fevereiro gastaram mais com transporte e menos com lazer também, já em março gastaram mais com alimentação e menos com lazer. A média foi de 188,5 e o desvio padrão de 48,8.”. Mesmo desenhando o gráfico com todos os elementos, PE4 expôs que “ele fala que isso aqui é legenda, eu botei aqui. No caso, eu botaria isso como título (...). Mas deu para entender. E ela falou muito rápido a legenda porque o texto é maior. Achei que ela falou muito rápido.”. Outros participantes também expressaram a mesma opinião, como PD1, que disse: “a função legenda deveria ser outra informação como a cor, tipo barras vermelha significa isso... barra azul significa outra coisa e assim vai. Mas é uma boa descrição, só que o nome que tá errado como legenda, poderia ser o título”.

Todos do grupo com deficiência conseguiram entender o gráfico B, mas como há muitas informações, eles não falaram exatamente todos os dados. PD2 explicou sobre seu entendimento: “O gráfico mede... hum... o gasto, o quanto que esses jovens estão gastando por ano em 3 categorias – alimentação, transporte e lazer. Custo por ano, no eixo X tem Meses e no eixo Y tem a média dos valores em reais, e aí no primeiro que eu lembro eles gastaram muito mais com o transporte e muito menos com lazer, no transporte foi 250 reais e no lazer foi só 100, no intermediário disso foi alimentação (...).” O único participante a falar sobre a informação que falta no gráfico foi PD3 que sugeriu colocar a informação da fonte em algum comando, pois acredita que esse dado é muito importante para o entendimento do gráfico. Sobre faltar informações para entender esse tipo de gráfico, todos responderam que não, como PD3 que afirmou: “Não faltou nenhuma informação, tem todas as informações necessárias, inclusive, as cores das barras.”

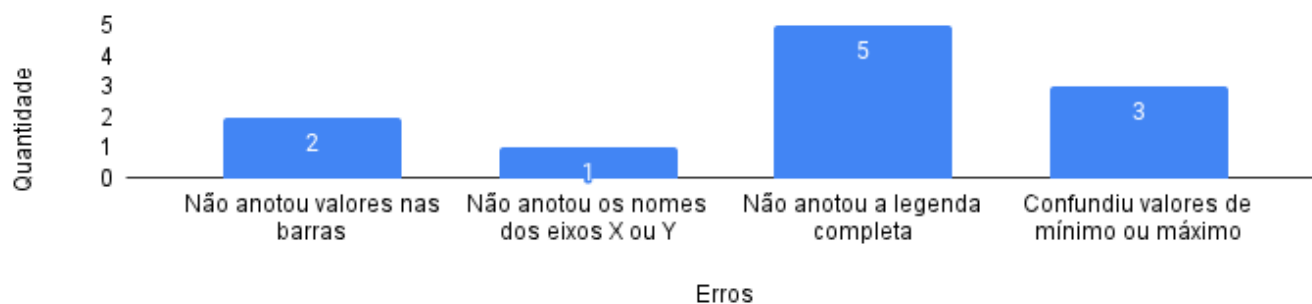


Figura 6: Erros observados nos desenhos do gráfico de barras agrupadas

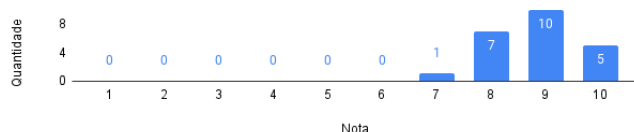


Figura 7: Grau de satisfação dos participantes

6.2.3 *Entrevistas.* Por fim, foram realizadas entrevistas e o participante era livre para opinar sobre a ferramenta. Alguns participantes sem deficiência fizeram críticas relacionadas à velocidade do áudio reproduzido, reportando essa reclamação ao ouvirem a descrição do gráfico de barras agrupadas. Já as pessoas com deficiência visual acharam a velocidade um pouco lenta, conforme PD1: “a velocidade da descrição ‘tá’ ok, mas eu usaria mais rápido porque já estou acostumada com o meu leitor de tela.” O participante PD2 também relatou: “se pudesse ajustar a velocidade dos áudios, eu botava em uma velocidade mais rápida.” Lembrando que as descrições dos áudios foram vocalizadas com a velocidade padrão da biblioteca GTTS. Uma crítica recebida foi relacionada à falta de comandos para pausar e voltar, PD1 mencionou: “Poderia ter um pause e voltar a informação aonde eu queria, para ter essa informação, e o comando pause poderia ser a tecla espaço porque é mais convencional e fácil de achar.” Outros problemas citados pelo PD3 foram o tamanho da tabela para o leitor de tela encontrar mais rápido quando o usuário passar o mouse por cima, e também permitir que o usuário mude as cores das páginas web para uma pessoa com baixa visão poder personalizar de acordo com sua necessidade.

Apesar das críticas mencionadas, a ferramenta obteve elogios de todos os participantes, em vários aspectos como na fala de PS7, que disse: “Achei que estão bem compreensíveis, nem muito rápido e nem muito lento.” Da mesma forma, PD1 afirmou que: “a descrição geral é boa, mas eu gostei mais com os comandos particulares (separados). Como por exemplo, eu tenho muita dificuldade de memorizar os eixos. Como ‘cê’ tem tudo junto é legal porque tem um entendimento amplo. Mas eu gostei muito em saber o valor das barras individualmente, quando eu teço 1 tem o valor do x e y na barra 1, e a leitura dos eixos quando eu aperto X e Y e eu gostei disso.”. Isso indica uma boa interação entre a ferramenta e o usuário, o relato de PD3 reforça a ideia, pois ele comenta que: “conforme eu aperto os comandos e

escuto as descrições, vou montando o gráfico mentalmente com os dados (...)”.

A pesquisa revelou que a vocalização do modelo de descrição aplicado com a ferramenta auxiliou os usuários com e sem deficiência visual a terem uma boa compreensão das informações dos gráficos A e B. Porém, é notado que ainda há dificuldade de interpretação dos gráficos quando os voluntários não conseguem acertar os valores de máximo e mínimo que são solicitados. Uma possibilidade para solucionar esse problema seria inserir explicitamente esses valores na vocalização com os termos ‘máximo’ e ‘mínimo’. Para os participantes com problemas nos títulos e legendas, pode-se verificar a possibilidade de produzir textos menores para esses valores nos gráficos, como um recurso de acessibilidade para todos.

Com o aprendizado adquirido durante a pesquisa, é possível fazer algumas considerações sobre as adversidades encontradas e melhorias na ferramenta. São elas: a) controle de velocidade dos áudios – é viável a implementação de ajuste da velocidade de acordo com a necessidade do usuário; b) ausência de comandos para pausar, parar, voltar e prosseguir – é possível inserir novas funções de comando em uma espécie de tocador mp3; c) tabela de comandos com o tamanho pequeno – deve-se permitir o ajuste da tabela de comandos para um tamanho maior, fazendo com que o leitor de tela encontre mais rápido as informações dos comandos e as pessoas com baixa visão possam aumentar a fonte; d) falta de ajuste de cores – deve-se implementar uma função para que o usuário possa aplicar alto contraste nas páginas, promovendo acessibilidade às pessoas com baixa visão.

As análises dos testes indicaram que a ferramenta demonstra um grande potencial como tecnologia assistiva para tornar as informações de gráficos de barras simples e agrupadas mais acessíveis aos usuários, principalmente para deficientes visuais, auxiliando a terem um bom entendimento desses recursos visuais. Além disso, pode-se sugerir a utilização da ferramenta também no processo de ensino-aprendizagem e letramento de usuários sobre gráficos de barras.

7 CONSIDERAÇÕES FINAIS

Este trabalho relatou o módulo de vocalização de descrição textual para gráficos de barras simples e agrupadas, com objetivo principal de propor e analisar a ferramenta para auxiliar deficientes visuais a terem um bom entendimento por meio da vocalização, já que as vezes esses dados são inacessíveis para esses usuários. Os resultados

obtidos com os testes indicaram alguns problemas de interação que são limitações deste trabalho, como ausência de alguns comandos, controle de velocidade dos áudios e falta de ajuste de cores. Questões técnicas que podem ser facilmente consideradas em versões futuras da ferramenta. Também foi possível perceber aspectos comuns aos voluntários, como a maior parte interage seguindo a ordem dos comandos de cima para baixo. Em relação ao entendimento dos gráficos de barra simples, a maior parte não teve dificuldade nenhuma de entender e explicar o conteúdo dos dados. Porém, o gráfico de barras agrupadas, por conter mais informações, teve pequenos erros de interpretação.

Os participantes avaliaram a ferramenta e afirmaram ter um alto grau de satisfação, isto indica o seu grande potencial em ajudar pessoas com deficiência visual a ter acesso a informação desse tipo, auxiliando em suas interpretações de dados nos gráficos. Por fim, conclui-se que logo após as correções dos problemas expostos e a disponibilidade dos códigos dessa aplicação web em um repositório público para outros desenvolvedores utilizarem, essa ferramenta servirá como base e incentivo para outras pesquisas sobre tecnologias assistivas, assim, permitindo a inclusão de mais pessoas.

Como trabalhos futuros pretende-se implementar todas as alterações sugeridas e realizar novos testes com mais usuários com deficiência visual, para fazer uma validação da ferramenta com um número mais expressivo de voluntários, descobrindo novos requisitos para mais melhorias. Outra função para ser adicionada em uma nova pesquisa é a interação por voz, ao invés de pressionar teclas do teclado. Pretende-se também adaptar diferentes descrições textuais para vários tipos de gráficos, como pizza, barras horizontais e dispersão, propondo uma nova lista de comandos para esses gráficos aumentando a flexibilidade da ferramenta.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Agradecemos também a todos os voluntários que participaram da pesquisa.

REFERÊNCIAS

- [1] João Marcelo Santos Marques, Simone Bacellar Leal Ferreira, and Claudia Cappelli. Identificando as principais dificuldades na compreensão de gráficos pelos cidadãos cegos. *Brazilian Journal of Development*, 6(11):88683–88704, 2020.
- [2] Censo Demográfico de 2010, 2010. URL <https://www.ibge.gov.br/estatisticas/sociais/populacao/9662-censo-demografico-2010.html?=&t=destaques>. Acessado em 12 de novembro de 2019.
- [3] Secretaria Especial dos Direitos Humanos da Presidência da República. Ata vii – comitê de ajudas técnicas – cat. 2007.
- [4] Cláudio Kleina. Tecnologia assistiva em educação especial e educação inclusiva. *Curitiba Inter saberes*, 2012.
- [5] Jaws, 2019. URL <http://www.freedomscientific.com/Products/Blindness/JAWS>.
- [6] Projeto dosvox, 2019. URL <http://intervox.nce.ufrj.br/dosvox/>.
- [7] Nv access, 2019. URL <https://www.nvaccess.org/>.
- [8] Virtual vision, 2019. URL <http://www.virtualvision.com.br/>.
- [9] Rafael José Geraldo and Renata Pontin Fortes. Dificuldades de usuários cegos na interação com a web: uma análise sobre as pesquisas. *Revista de Sistemas e Computação-RSC*, 3(2), 2013.
- [10] Daekyoung Jung, Wonjae Kim, Hyunjoo Song, Jeong-in Hwang, Bongshin Lee, Bohyoung Kim, and Jinwook Seo. Chartsense: Interactive data extraction from chart images. In *Proceedings of the 2017 chi conference on human factors in computing systems*, pages 6706–6717, 2017.
- [11] Huahai Yang, Yunyao Li, and Michelle X Zhou. Understand users' comprehension and preferences for composing information visualizations. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(1):1–30, 2014.
- [12] Leilani Battle, Peitong Duan, Zachery Miranda, Dana Mukusheva, Remco Chang, and Michael Stonebraker. Beagle: Automated extraction and interpretation of visualizations from the web. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2018.
- [13] Leo Ferres, Gitte Lindgaard, Livia Sumegi, and Bruce Tsuji. Evaluating a tool for improving accessibility to charts and graphs. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(5):1–32, 2013.
- [14] Charles F Greenbacker, Peng Wu, Sandra Carberry, Kathleen F McCoy, Stephanie Elzer, David D McDonald, Daniel Chester, and Seniz Demir. Improving the accessibility of line graphs in multimodal documents. In *Proceedings of the Second Workshop on Speech and Language Processing for Assistive Technologies*, pages 52–62, 2011.
- [15] Sílvia Rodríguez Vázquez and Sabine Lehmann. Acrolinx: a controlled-language checker turned into an accessibility evaluation tool for image text alternatives. In *Proceedings of the 12th Web for All Conference*, pages 1–2, 2015.
- [16] Vítor Carvalho and Diamantino Freitas. Automatic description of svg images for the visually impaired: a gestaltic approach. *Procedia Computer Science*, 67:2–11, 2015.
- [17] Aline Aparecida Nascimento Frazão, Livia da Conceição Costa Zaquero, Ísis de Paula Santos Mendonça, Thays Nayara Frazão Silva, and Francisca Morais da Silveira. Tecnologia assistiva: Aplicativos inovadores para estudantes com deficiência visual. *Brazilian Journal of Development*, 6(11):85076–85089, 2020.
- [18] Cynthia Leticia Teles De Oliveira, Alan Trindade De Almeida Silva, Erick Modesto Campos, Tiago Davi Oliveira Araújo, Marcelle Pereira Mota, Bianchi Serique Meiguins, and Jefferson Magalhães De Morais. Proposal and evaluation of textual description templates for bar charts vocalization. In *2019 23rd International Conference Information Visualisation (IV)*, pages 163–169. IEEE, 2019.
- [19] Simone Barbosa and Bruno Silva. *Interação humano-computador*. Elsevier Brasil, 2010.