

# Audio description on Instagram: evaluating and comparing two ways of describing images for visually impaired

João Marcelo dos Santos Marques<sup>1</sup>, Luiz Fernando Gopi Valente<sup>1</sup>, Simone Bacellar Leal Ferreira<sup>1</sup>, Claudia Cappelli<sup>1</sup> and Luciana Salgado<sup>2</sup>

<sup>1</sup>*Department of Applied Informatics, Federal University of the State of Rio de Janeiro, Av. Pasteur 458 - Urca, Rio de Janeiro, Brazil*

<sup>2</sup>*Institute of Computing, Federal Fluminense University, Av. Gal. Milton Tavares de Souza, s/n - São Domingos, Niterói, Brazil*

{[joao.marques](mailto:joao.marques), [luiz.valente](mailto:luiz.valente), [simone](mailto:simone), [claudia.cappelli](mailto:claudia.cappelli)}@uniriotec.br, [luciana@ic.uf.br](mailto:luciana@ic.uf.br)

**Keywords:** Accessibility, Audio Description, Instagram.

**Abstract:** The social network Instagram encourages interactions among users around audio-visual content (pictures and short duration videos). However, this type of content still presents itself as a barrier for the visually impaired. To mitigate this problem, screen readers can be used, but those only work for images which have texts in the form of subtitles. Audio description, on the other hand, is a technique that describes visual images into words, allowing the comprehension of these elements. This technique has been used in many fields, fostering a scenario of inclusion and opportunities for this public. The objective of this paper is to evaluate and compare these two forms of describing images published on Instagram: one utilizing the descriptive text read by the screen reader and another utilizing audio description recorded by the image's own author. Through an empirical study, we have identified the form of image description preferred by the visually impaired participants and if the use of audio description on Instagram would encourage its use by this public.

## 1. 1 INTRODUCTION

With the evolution of information and communication technology, new opportunities and possibilities are open for people who have some type of disability. Nowadays it is common to find people with visual impairment using computers to surf the Internet with the help of screen reading software. This public is not small. According to the data from the 2010 IBGE (Brazilian Institute of Geography and Statistics) census, concerning those with severe visual disabilities (those with great difficulty in seeing or who cannot see at all), more than 6.6 million people have claimed to have this type of disability. Of these 6.6 million, 506.3 thousand have claimed to be blind. (IBGE, 2010).

Not all interactive systems, however, are designed to meet the needs of this portion of the population. Among the types of systems that present accessibility problems are social networks. Accessibility is the term used to indicate the possibility of anyone enjoying the benefits of life in society, and among

them, is the use of the internet (NBR 9050, 1994; Nicholl, 2001). Despite the advances, studies show (Piovesan et al., 2013) there is still much to be done to meet all of the accessibility criteria.

One of the prominent features of virtual social networks is the frequent use of user published images and videos, a behavior which is becoming more popular in the last few years. In order for information to be accessible to all, there are accessibility guidelines and recommendations, some of which are specific for images and videos. Guidelines that focus on audio-visual resources include those which determine that all non-textual content must be displayed in an alternative format, that is, images must come with an alternative text describing it. (W3C, 2008). Images accompanied by text can be understood by the visually impaired, as the screen reader reads the text.

In social networks which are completely based in pictures and videos, such as Instagram, accessibility issues it's a barrier, which can prevent people with seeing disabilities from using them. For this specific public, there are two fundamental issues: Instagram

must be completely accessible, according to accessibility standards which have already been published, and the user published images should be accompanied by a descriptive text which can be read by a screen reader.

This paper has as an objective to evaluate and compare two forms of describing images on Instagram, one through the reading of an image's descriptive text by the screen reader, and another through audio description recorded by the image's own author which could be heard through the execution of an audio file, permitting us to identify if the use of audio description would encourage this public to have a greater participation in online image and video based social networks.

This paper is divided into following fashion: section 2 presents the theoretical framework and describes the main concepts involved in this research; section 3 describes how this research was planned and executed; the results are analyzed in section 4 and section 5 discusses this research's conclusions.

## **2.2 THEORETICAL FRAMEWORK**

### **2.1 Virtual Social Networks**

Broadly speaking, social networks are any type of relationship among people, mediated or not by computerized systems. Such relationships involve interactions which aim to change people's lives, for the collective or organizations, since such interactions can occur for private interests, in defense of others or in the name of organizations (Tavares and de Paula, 2015). Among the types and formats of social networks, those which are established in cyberspace, denominated "virtual social networks", represent a new and complex universe of communicative, social and discursive phenomena (Recuero, 2014). For Lévy (Lévy, 1999), a virtual social network "is built on the affinity of interests, knowledge, of mutual projects, in a process of cooperation or exchange, all of which are independent of the geographical proximity and institutional affiliation.

According to a report produced by the Brazilian Media Research, in 2015 (SECOM, 2015), almost half of Brazilians use the Internet regularly. The use of social networks grows each year: in 2014, the most accessed social networks were Facebook (mentioned by 83% of people), Whatsapp (58%), Youtube (17%), Instagram (12%) and Google+ (8%).

One of the fastest growing social networks in the world and in Brazil, Instagram, focuses exclusively on the publishing of images and short duration videos (Quadros, 2015). Its main objective is the sharing of this content among its participants. Users can explore published images and choose the people or organizations they wish to follow. In this network, the act of following a person or organization establishes a bond which represents, at least, and interest in keeping up with their publications. Interaction between users can also happen when one user likes or comments on a picture or video published by another user.

Instagram was created in 2010 and reached the milestone of a million users in the first two months in which it was available on the Apple platform (Paschoal, 2015). This success is, in great part, due to the ability of applying filters to pictures before publishing them, which allows users to simulate a variety of effects on their images. In 2012 Instagram was also made available for the Android platform, where it was installed by 1 million users in just 24 hours. Brazil is the second country with the most active users in this network, following the United States (Ribeiro, 2015).

### **2.2 Instagram Web: promoting Accessibility**

As a vehicle for communication with the internet, through which a variety of information is transmitted to people spread in many regions of the world (Agha, 2008), the interface of systems and applications such as Instagram must enable the access by any person, regardless of their physical-motor and perceptive, culture and social abilities. That is, they must be designed in conformity with accessibility guidelines.

However, obtaining interfaces that meet the needs of many users is not a trivial task, since there are a variety of people with distinct limitations. In order to orientate developers in the elaboration of accessible systems, there are recommendations and guidelines, such as the "Web Content Accessibility Guidelines" proposed by the W3C international committee, which regulate issues related to the internet. These guidelines address issues which hinder the access to websites by users with access characteristics or limitations (W3C, 2008). These efforts enabled the Internet to play a key role in the daily lives of people with disabilities, allowing them to create new forms of relationships, find job opportunities and leisure options (Queiroz, 2012).

In early 2013, Instagram announced the launch of a Web version so that its users could access the social network through the computer and not only through the mobile application. In this Web version, the user can visualize, like and comment on pictures and videos. However, in order to maintain its initial strategy, the publication of new images on Instagram would still be an exclusivity of the mobile application. This decision was made to preserve the application's basic characteristics. According to one of its executives, "Instagram is about taking pictures on the spot, in the real world, in real time" (Stivanin, 2015).

The creation of a Web version for Instagram opened up new possibilities of use, mainly for those who have seeing disabilities, as they can use their computer and screen reading software for access while not having to depend on a smartphone. Nevertheless, as sight is the main form of interaction in this type of network, users with accentuated or total seeing disabilities require an assistive technology capable of capturing the interfaces and making them accessible. Therefore, regardless of how well designed an interface is, if it's not accessible, it will be a barrier to the social inclusion of the visually impaired. Furthermore, these users' access also depends on the characteristics of these assistive technologies (Ferreira and Nunes, 2008).

Assistive technology is the term used to identify any tool or resource (like a cane) which provides or expands the functional abilities of people with impairments and thus promotes greater autonomy (Ferreira and Nunes, 2008). In the case of a person with accentuated or total visual impairment, Internet access is possible through screen reading software, applications associated with voice synthesizer software, which permits users to navigate the internet, read and send emails and connect with other people through social networks, including Instagram. Consequently, interfaces must be designed to, when accessed by assistive technologies, provide easy interactions, capable of being detected and interpreted correctly.

### **2.3 Inclusive Instagram: Accessible content**

The WCAG (W3C, 2008) is organized in principles, guidelines and testable success criteria. Each guideline and success criterion has its own specific techniques to evaluate whether they have been met. In it, it is described that textual information for any non-textual content, such as graphical information, in

the case of the visually impaired, or sound information, and the case of the hearing impaired, must be provided.

Since Instagram's Web version can be used by people with seeing disabilities, not only accessibility to the web site's entire structure must be ensured, it is also necessary to give special attention as to how these users can understand the images and videos published, like them and comment on them. In this case, one of WGAC's guidelines of the Principle of Perception must be observed, as it is directly linked to accessibility issues for the understanding of images: Guideline 1.1 Text Alternatives: Provide text alternatives for all non-textual content so that it can be presented in different ways, according to the necessity of the users, for example: enlarged characters, braille, speech, symbols or a simpler language.

One of the ways of providing textual information for any non-textual content, such as graphical information, in the case of the visually impaired, is through the "alt" attribute, which provides a textual equivalent for the images; it adds an alternative text to the image which is read by the screen reader, thus providing meaning to the image (Ferreira and Nunes, 2008).

### **2.4 Accessible content – Audio Description**

Audio description is a process which consists of transforming visual images into words, which are then spoken during the silent intervals of audiovisual programs or live performances (Cintas, 2005). The audio-visual translation technique encompasses describing images in words, transmitting feelings, through intersemiotic translations, that is, translations which consist of the conversion of one system of symbols into another, translation of verbal text into a nonverbal text, such as dance, painting, music, etc. In the view of the Ministry of Communications of the Brazilian Federal Government (Ministério das Comunicações, 2006), audio description is conceptualized as "a narration, in Portuguese, integrated to the original sound of the audiovisual piece, containing descriptions of sounds and visual elements and any additional information which is relevant to allow a better understanding of them by people with visual or intellectual disabilities".

Audio description is an accessibility resource which allows people with visual disabilities to view and comprehend photography, videos, films, theatrical plays, TV show, exhibits, musicals, among

others (Queiroz, 2012). Through this resource, people with seeing disabilities are able to understand the scenarios, costumes, facial expressions, body language and several actions which are not presented during the spoken words of an artistic exhibit.

Audio description was initially done informally, by people who accompanied those with seeing disabilities to shows; these would narrate to them what they could not listen to in the script of the show or when people with seeing disabilities asked questions, they answered doubts, during a movie, theatrical plays or other types of shows (Queiroz, 2012).

This description technique originated in the United States, in the 1970s, from the ideas developed by Gregory Frazier in his thesis "Master of Arts" from the University of San Francisco - USA, where, for the first time, the term "audio description" was used. As this resource was being publicized, it gained space in media, through the Japanese TV network, NTV, which began to transmit its programming in audio description in 1983. Then, Gaberta (Ofcom, 2010) a TV network in Catalonia, Spain, also did so. The Cannes Film Festival also joined the idea in 1989. In Brazil, the first audio-described movie screenings took place during the São Paulo International Short Film Festival, during the 2006 and 2007 editions (Silva, 2009).

The audio description can also be used for the comprehension of videos and pictures (Queiroz, 2012), it can become a tool to facilitate the interaction of the visually impaired on the Web. As the use of screen readers demands and requires a wide use of key combinations and a steep learning curve, most people with low sight prefer not using it whenever possible. Therefore, an alternative resource, like an audio description, which presents the details of an image in an audio file, can be an alternative for these people. Likewise, an audio description can be useful for blind users, since the figure's description wouldn't be provided in a synthetic manner by a reader, but through a pre-recorded audio, which, if done correctly, can relay richer information than a synthesizer for someone with a seeing disability (mild or total).

In order to obtain universal access applications, it is fundamental to observe and analyze the difficulties and abilities of users with limitations, as they guide the mental model used throughout their interactions with the system. This evaluation enables harmonious interaction and, at the same time, guarantees comprehensible and navigable content (Queiroz, 2012). The participation of users with limitations

assists in the understanding of how they interact on the Web and use assistive technologies (Abou-Zahra et al., 2008). Through the observation of interaction strategies from different users in distinct contexts and utilizing several assistive technologies, difficulties faced can be identified (Melo, 2007), incorporating the experiences of these groups as users of the system (Slatin and Rush, 2003).

In this way, the evaluation of users interacting with Instagram will allow the identification of the barriers they face and a better assessment of their experiences while interacting with two distinct technologies, the screen reader or audio description.

## 2.5 Related Works

The use of audio description has been addressed in many academic papers, in which authors analyzed the benefits of this tool for those with seeing disabilities. Santos (Santos, 2016), for example, has contributed, in the field of translation studies, for the development of the researches and has addressed the use of audio description as mediation in museums. By means of a case study in the Indigenous Peoples Memorial, the author proposes reflections on the implications of the use of audio description in the experience of a person with a seeing disability in a museum, aiming to provide access to visual pieces and create, through the means of verbal language, conditions for the inclusion of this public.

Villela (Villela and Losnak, 2016) in his turn, used audio description to depict pictures of the Military Dictatorship period, helping to keep the memory of remarkable facts alive for the Brazilian society, including for people with visual impairment. The objective of this work was the creation of an accessible photo-documentary about the fifty years of the Military Dictatorship in Brazil, presenting the most significant moments of this period for people with seeing disabilities through the use of audio description. Additionally, a script was created for the presentation of previously selected and audio describe pictures, focusing on historical scenarios and characters, depicting people's physical characteristics in a very objective manner.

Other works are concentrated in proposing and developing new resources which allow the visually impaired to more effectively access websites on the Internet, such as virtual social networks. At the end of 2015, Facebook announced that it was working on an artificial intelligence based object recognition tool in order to help blind users have an idea of the pictures people shared on Facebook (Dickey, 2015). The

solution consists in processing the image and generating a new alternative text which could then be described for the user through a screen reader. The engineers involved in the project believe that, despite not completely describing the images with all their details, the level of engagement of the visually impaired could increase.

As in this paper, the Facebook initiative described above aims to provide the visually impaired with new opportunities of participation in social networks, describing the images published by the users.

The main difference is in the reach and assertiveness of each method: the use of audio description that has been recorded by the content's own author could present richer details and allow a greater understanding by those with seeing disabilities.

### **3.3 RESEARCH DEVELOPMENT**

This research's method was qualitative-observational, (Cresswell, 2009; Denzin and Lincoln, 2003) based on a case study in the "União de Cegos do Brasil" (Brazilian Union of the Blind) institute and involved a public composed of both young and elderly participants. This study aims to address the following research question: "How to evaluate and compare two forms of describing images published on Instagram: one utilizing the descriptive text read by a screen reader and another utilizing audio description recorded by the image's own authors?". The development of this work was done by two researchers and organized in four stages: a) test preparation; b) participating user's profile selection; c) execution of the tests and d) analysis of the results.

a) Test Preparation: We took into consideration features regarding free screen readers for computers running the Windows operating system.

For the realization of the tests, the NVDA (NonVisual Desktop Access) was chosen for being familiar to the participants and also for having more than forty-three language options, including Portuguese, and for being able to be used from a USB drive, with no need of installing this reader in a computer (NVDA, 2014).

The tests were done in the União dos Cegos institute, a Federal, State, and Municipal public utility institution, founded in 1924, whose mission is to ensure that a person with seeing disabilities is able to reach its potential as a full citizen.

To capture the opinion of the visually impaired participants, two questionnaires were formulated. The first questionnaire (pre-test questionnaire) addressed questions about the profile of each participant, such as educational level, type of disability, profession, age, gender and computer using habits. Regarding computer use, the following questions were asked: Do you have any experience with screen readers? Do you use computers to access the Internet?; Do you know any social networks?; Do you have an account on any social network?; With what frequency do you use social networks? If you did not use any social network, what would be the reason?

The second questionnaire (post-test questionnaire) encompassed questions related to the comprehension of the audio descriptions of the shown images, such as: How would you grade your understanding of the image of the 1st test? How would you grade your understanding of the image of the second test? Based on the two image description tests, which was the best for your understanding? To answer the questions about the comprehension of the images in both tests, it was necessary for the participant to attribute a grade in a scale from 0 to 10.

For the realization of the audio description tests, two images were selected from a public Internet base. Despite being different, the images should present a similar context and theme, restricting only the manner in which they would be described, so that the participants would influence the results. In this way, two images were selected (figure 1 and figure 2), representing a family composed by father, mother and two children in a moment of leisure. These images are referenced as "image 1" and "image 2".

The redaction of the descriptive texts of each image followed the same style and format. The full text of each image is reproduced below.



Figure 1: Image described by the screen reader (image 1).

Descriptive text of Figure 1: “A family is playing on a green lawn. The father and his son 8 years-old are standing. The son is flying at colorful kite and the father proudly watches it fly. The mother is sitting on the grass with her younger daughter on her lap, watching father and son play. The day is lovely, with very blue skies, no clouds, and they all seem very happy”.



Figure 2: Image described through audio description (image 2).

Descriptive text of figure 2: “A family is walking along the sand on a beautiful beach. The mother carries her youngest son on her shoulders. The father is just behind and plays with his other son, throwing the boys up in the air to catch him soon after. The

seawater is quite blue with a few small white waves. On the horizon, there's a mountain with green trees and a few houses. They're all smiling and seem quite happy”.

At first, one of the limitations of this research could be the fact that the audio descriptions were produced by the researchers themselves; they were not made by professionals specialized in audio description. But the choice to make them in a personalized way was deliberate, as, if applied to Instagram, they would be generated by the users themselves. The recording of the audio description was made by a collaborator. She read the descriptive text in Figure 2, in natural speech, respecting the pauses predicted on the text's punctuation.

Then, two local copies of the Instagram website's HTML files were made, maintaining all their layout and visual identity. One of the copies was modified to present “image 1” and its descriptive text, as if it were a normal user publication, and the other copy received “image 2”, but inserting below it, a button that, when pressed, would play an audio with the recording of their descriptive text made by human collaborator.

b) Profile definition and selection of participants: In order to participate in the research, the participant would have to be an adult over 35 years old, and have severe visual impairment. It was decided not to pre-screen the participants as to avoid their commenting to one another, which could influence the results of the research. The participants were invited according to their availability in the activities and programming at the União dos Cegos do Brasil institute.

The answers from the pre-test questionnaire revealed that 50% of the blind declared to have completed High School. Of those with low vision, 33% declared to have higher education and 70% declared to have finished Elementary School. As for the disability, 60% of the respondents had a type of severe visual impairment (total or low vision) that consists of the total lack of visual perception of any type of light. 80% of the participants reported to have made use of a computer for Internet access. Concerning screen readers, 67% of participants utilize this type of software and the other 33% are aware of it. All of them reiterated that they have heard of social networks such as: Facebook, Instagram, Twitter, Whatsapp. However, 33% have never accessed any of these social networks. 67% of participants access social networks twice a week, on average, and are mostly female. Regarding the profession of these participants, they were the most varied (pensioner, retirees, early education teachers and the medical area).

For the sake of maintaining the anonymity of these participants, we sought to preserve their names, which were coded as: P1, P2, P3, P4, P5 and P6. The profession, age, type of disability and computer use of the participants are illustrated in Table 1.

Table 1: Code, Profession, Age, Visual Disability, Computer Use of Participants.

Code	Professional	Age	Visual Disability	Computer Use
P1	Pensioner	35	Total	No
P2	Retired	45	Total	Yes
P3	Doctor	59	Low Vision	Yes
P4	Retired	67	Total	No
P5	Retired	51	Total	Yes
P6	Teacher	68	Low Vision	Yes

c) Execution of the Tests: All of the research's details, its objective and mainly the benefits that could be expected from this work were explained to the institute's coordinators by the researchers. The coordinators requested that the duration of the test did not exceed three hours in total so as not to compromise the participant's programmed activities for the day. A room was made available for the research team, containing two laptops (one for the execution of the tests, another for support), sound speakers and headphones.

As they arrived at the institution, the participants were directed in pairs to the test room by one of the institutions coordinator. This type of approach (in pairs) was a request of the institutions coordinators for the realization of the tests. In the test room, the participants received the initial information from both researches so they would understand clearly what would be done and what was expected from each one of them. Besides ensuring that all participants would receive the same information in a standardized fashion, the objective of this initial explanation was to reassure the participants and make them more comfortable during the test run.

The first activity was the application of the pre-test questionnaire. The questions, whose objective was to collect each participant's profile information, were read by one of the researchers and the answers written down on printed forms.

In the second activity, each participant listened to the text description of "Image 1" by the NVDA screen reader. Then, before answering or making any comments, the participant listened to the text description of "Image 2" through the execution of an audio description which had been previously recorded by a human collaborator. The choice of the description being made first by the screen reader was due to the fact that 67% of the participants were already familiarized with it and that it had already been extensively studied in the literature. At any moment the participants were informed about which description method was being used.

Lastly, after listening to the image descriptions, the post-test questionnaire was applied. Once again, the

researcher read the questions to the participants and wrote down the answers on printed forms.

During the test run, the main impressions, difficulties and reactions from the participants were registered, and are described below.

The participant P1 reported that after listening to the audio description of image 2, she had the sensation of being part of the scene, since the intonation put by the human voice was very real, as if the family described in the image was by her side; If she had to opt between the two descriptions, she would opt for the audio description. It was noticed that this participant had no greater difficulties in carrying out this test. As far as emotional reactions are concerned, it can be said that she was very secure, determined and alert in carrying out the test and, at the end, she still said: "it's already over!"

For participant P2, who wasn't used to screen readers, there was no difficulty in doing the test, even while being tense, determined and attentive. At the end of the test, she said that application developers should be more concerned about building tools with audio description, aiming to include the visually impaired who are largely forgotten by that professionals, thus ratifying her choice for audio description.

Participant P3 was the most enthusiastic about taking part in the tests. Before starting, she mentioned that she loved screen readers and audio description and asked: which movie are you going to show us? Even though she had been oriented on how the tests would be performed, that she would not be shown a movie, she did not lose her good mood and determination when she realized it was not a movie. So, good mood was another defining trait of this participant. There were no noticeable difficulties in the handling of the equipment during the tests, as she felt very secure with them. According to her, the audio description was so real, clear, enriched in details, as, for example, the sound of the waves, that, if she could, she would like to be able to play with the couple's children. She concluded her participation by saying: "audio description can be seen as a way of the visually impaired getting to know a world which can't be seen or explored by many".

Volunteer P4 was somewhat tense in the expectation of what would happen, but was determined to conclude the tests. For this voluntary there was no difference between the descriptions of images 1 and 2 made by the screen reader and the human voice, even though she was not used to working with screen readers. When asked about which description she would prefer, she said she would opt for the audio description, as it more closely depicts the reality of the facts to the spectator.

Impressions and reactions of participant P5 drew the most attention, as it was the participant who had lost almost all sight five years before (he became blind at age 46) due to complications of a glaucoma caused by diabetes mellitus. He had graduated in programming and was very determined, enthusiastic and attentive during the tests; he distinguished himself from the others by his professional experience in handling computer equipment and social networks and, therefore, did not have any difficulties during

the tests. Regarding the description of the images, he said the audio description was far superior when compared to the one made by the screen reader. He reiterated that more effort and investment should be made so that software developers could build more tools that use audio description. According to him, if there were more investments in audio description tools, the social inclusion of the visually impaired would be better promoted.

The last participant, P6, was tense but attentive to instructions and handling of the equipment. When asked about the best image description, after the realization of the test, she answered that she would prefer the screen reader's description, even though she was not used to working with them. However, in her opinion, the description of the images utilizing the audio description tool could stimulate the use of social networks.

The time established by the research method for the realization of each participant's tests was 15 minutes. On average, the tests lasted approximately 10 minutes, which included: the objective of the research; the profiling of the participants; the understanding of the method utilized in the two image descriptions; the questionnaires (pre-test and post-test); and the listening of the descriptions.

At the end of the data collection, the information was consolidated and analysed.

Research limitations: One of the limitations of this research was the fact that only images were analyzed. No work was done regarding videos. One of the accessibility recommendations determines that all real time (live) or pre-recorded audio and/or video content, must be made available through alternative content which presents transcribed or described information.

#### 4.4 EVALUATION OF TEST RESULTS

For the purpose of result analysis, it is possible to divide participants in two groups according to their level of experience with screen reading software. Out of the six participants, three had some experience with screen readers and the other three didn't have any contact with this type of software. Of the participants who had some experience with screen readers, P1, P3 and P5 are highlighted. P1 revealed that she has been using this type of software for over 10 years. P3, who had lost her vision when she was young, reiterated that she has been using the screen reader for over 15 years. P3, who has not been able to see for over 5 years, began using this application after losing her sight.

The first two questions of the questionnaire had as an objective make a direct comparison between the two methods of image description used in the test. Figure 3 illustrates the grades attributed by the participants regarding

the description made by the screen reader, on a scale from 0 to 10.

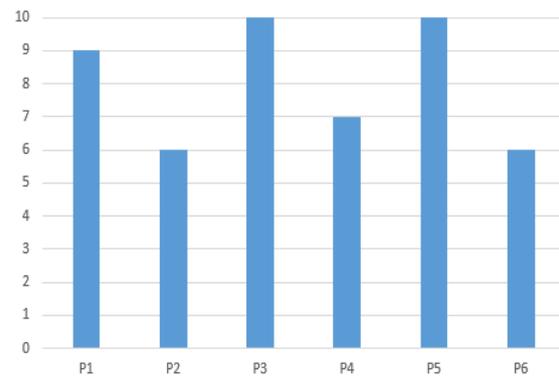


Figure 3: Grades attributed by the participants regarding the description made by the screen reader.

The lowest grades for the description were given by the participants of the group that didn't have any prior experience with screen reading software. When restricted to this group, the average grade for the description falls to 6.3. According to the participant's own comments, the frequent use of screen reading software increases the comprehension level of what is listened to during computer use. This justifies the higher grade given by the group which had experience with screen readers. In this group's opinion, the average grade for the screen reader's description was 9.7.

Figure 4 shows the grades given by the volunteers in regards to the audio description, on a scale from 0 to 10.

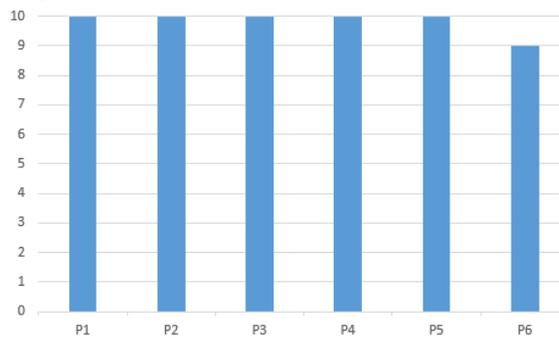


Figure 4: Grades attributed by the participants to the audio description.

When asked about which of the two methods favored a better understanding of the images, only one of the participants opted for the screen reader's description. Every other participant thought image comprehension was better through the use of audio

description, which represents 83% of total participants.

It is worth mentioning that, in the group which had prior experience with the screen reader, two participants had given grade 10 to both the screen reader's description as well as the audio description, that is, they had classified both methods the same way. However, if they had to choose between one of methods, they would choose the audio description.

The last question of the post-test questionnaire had as an objective to understand whether, in the opinion of participants, it the audio description to describe imagens would encourage the visually impaired to use social networks. All the participants answered yes to this question, resulting in a 100% approval rating.

The graph of Figure 5 illustrates the participants' perceptions of emotions, difficulties, and impressions during the tests, where enthusiasm, determination, attention, tension, and safety regarding the use of the screen reader were recorded.

From the data in this figure, it was verified that 31% of the participants were determined, that is, they were convinced that they could carry out the tests. The visually impaired who had attention during the experiment represent 25%. The facial expressions regarding tension could be observed in 19% of the participants. The perception of enthusiasm depicted 12% of users. And the safety during the tests was perceived in 13% of the volunteers.

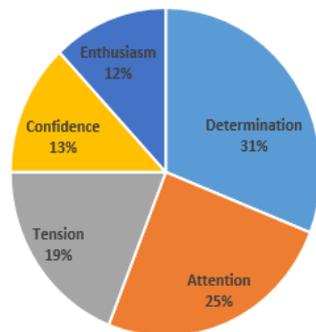


Figure 5: Perception of participant's emotions.

With respect to the question "do you know any social networks?", the participants mentioned they knew the following social networks: Facebook, Instagram and WhatsApp. Without exception, they mentioned their knowledge of Facebook. Figure 6 illustrates the percentage of the participant's knowledge of these social networks.

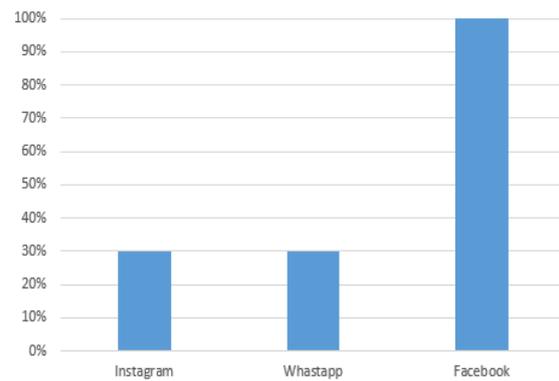


Figure 6: Knowledge of social networks.

When analyzing the use of social networks it was found that not all participants access them. Comparing the participants in relation to the use of this entertainment channel on the seven days of the week, it was noticed that only two of the participants accessed this type of channel. Participants P1 and P4 were not considered as they do not access any social network.

Regarding the voice of the screen reader, the majority of the participants were of the opinion that it was a very computerized, synthesized voice, and that for a better understanding and clarity of speech, the screen reader would have to be well configured, as voice quality is determined by its similarity to the human voice.

As it was mentioned in the previous section, the two participants that dealt with screen readers almost daily, were the ones who most questioned the choice for the synthesized voice for the realization of the test, as in their opinion, the voice was not very appropriate for use in the image description.

Participants P1 and P4 declared that they were considered to be people of "low-income", who depended on the technological resources available at the União dos Cegos institute for internet access and, consequentially, social networks. As such, they need to commute from their residences to the institute, aiming at socialization and, therefore, digital inclusion. Participant P1, who is a pensioner and depends on the government's financial resources for her livelihood, said: "A computer could be a Christmas gift". The retired participant, P4, reiterated that she does not have the financial conditions to buy a computer, and is not able to connect to social networks. For these participants, access to social networks would open new forms of interaction and communication, gradually decreasing their digital exclusion, as well as enriching their studies.

## 5. 5 CONCLUSIONS

The objective of this paper was to evaluate and compare two forms of describing images on Instagram, one through the reading of an image's descriptive text read by the screen reader and another through an audio description recorded by the picture's own author, which is heard through the execution of an audio file. Through the realization of tests involving a group of people with seeing disabilities, four with total impairment and two with low sight, it was possible to obtain important information about the participant's preferred method of image description and if the inclusion of audio description resources on Instagram could encourage the participation of people with visual impairment.

The analysis of the data collected during tests shows that the use of audio description allowed better image comprehension. The fact that the audio description of an image is narrated by a human (the speech of the screen reader is created by a sound synthesizer, which sounds somewhat artificial), was fundamental for the understanding, resulting in no difficulty of comprehension by the participants. Even among participants who already had previous experience with screen readers, the audio description was chosen as the best option. All the participants stated that having the possibility to listen to an audio description of an image that has been recorded by its own author (giving a greater personal focus to the content), would increase the participation of the visually impaired on Instagram, which, as it is completely image and video based, is currently barely inclusive for this public.

As it has been demonstrated throughout this article, audio description has shown itself to be an excellent tool for the inclusion of the visually impaired, permitting greater access and participation in cultural and leisure activities and education. Furthermore, accessibility standards for Internet web sites help developers make them accessible, ensuring access to all, including people who have some type of the visual disability. The results of this paper show that the use of audio description, allied to the fulfillment of accessibility requirements, can be decisive for these people's access to image based social networks, such as Instagram.

As in the description made by screen readers – Which depend on the production of text or subtitle that explains the image - the use of audio description on Instagram would also depend on the collaboration of users who publish pictures, as they would be

responsible for recording the audio description of their own images.

In future research, besides evaluating possible accessibility limitations in Instagram's WEB version, it would be important to study the modifications and new functionalities that would be necessary to be able to implement the correct use of audio description on Instagram.

In the application for smartphones, for example, new functionalities could be created which allow users to record the audio description in a quick and simple fashion. Currently, the publication of images on Instagram is done exclusively through smartphones, which already offer hardware and software tools for audio recording. As such, a user could publish a picture on Instagram and, shortly after, record the audio description with their own voice on their own devices.

As for Instagram's WEB version (referred to in this paper as an opportunity for the visually impaired to access this network), modifications should be made in order to offer new audio description resources. In this case, the focus would be on offering users forms of searching and identifying images which have audio description and allowing users to listen to them. Additionally, it would be interesting to create a new form of interaction in which the visually impaired user could send a request to the author of an image so that he would record an audio description, in case it had not been made yet. Besides being a form of increasing the volume of audio described images, this resource still establishes a new form of contact between the visually impaired and other users on Instagram.

New studies will be done to efficiently plan and define the set of changes on Instagram's systems, materializing the benefits of the use of audio description for the visually impaired pointed out in this research.

Another aspect that could be explored is the interest of volunteers in participating in social networks using audio description.

## REFERENCES

- Abou-Zahra, S., Bjarno, H., Duchateau, S., Restrepo, E., Henry, S., McGee, L., Pouncey, I., Rush, S., Sutton, J. and Wassmer, S. (2008) Evaluating Websites for accessibility: Overview ◦ web accessibility initiative ◦ W3C. Available at: <http://www.w3.org/WAI/eval/Overview.html> (Accessed: 15 December 2015).

- Agha, G. (2008) Computing in pervasive cyberspace. *Proceedings of the ACM Communications of the ACM*, 51, 1.
- Cintas, Jorge Díaz (2005), Audiovisual translation today: a question of accessibility for all. *Translating Today*, London, n. 4, p. 3-5, July 2005.
- Cresswell, J.W. (2009), *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 3rd Edition. Thousand Oaks: SAGE Publications.
- Denzin, N.K. and Lincoln, Y.S. (2003) *The landscape of qualitative research: Theories and issues*. Thousand Oaks, Ca. Sage Publications, Inc.
- Dickey, M.R. (2015) Facebook's working on A tool to help the blind 'See' images. Available at: <http://techcrunch.com/2015/10/13/facebooks-working-on-a-tool-to-help-the-blind-see-images> (Accessed: 14 May 2016).
- Ferreira, S. B.L e Nunes, R (2008). *e-Usabilidade*. Rio de Janeiro: LTC Editora.
- IBGE (2010). Sala de imprensa | notícias. Available at: [http://www.ibge.gov.br/home/presidencia/noticias/noticia\\_visualiza.php?id\\_noticia=2125&id\\_pagina=1](http://www.ibge.gov.br/home/presidencia/noticias/noticia_visualiza.php?id_noticia=2125&id_pagina=1) (Accessed: 14 May 2016).
- Lévy, P. (1999). *Cibercultura*, São Paulo, Editora 34. Coleção Trans.
- Melo, A. (2007) *Design Inclusivo de Sistemas de Informação na Web*. Doctoral Thesis, Universidade Estadual de Campinas, Instituto de Computação, Campinas.
- Ministério das Comunicações. Portaria nº 310, de 27 de junho de 2006. (2006) Available at: <http://www.anatel.gov.br/legislacao/normas-do-mc/442-portaria-310> (Accessed: 15 October 2015).
- NBR 9050. (1994). NBR 9050 Associação Brasileira de Normas Técnicas. *Acessibilidade de Pessoas Portadoras de Deficiências a Edificações, Espaço, Mobiliário*. Rio de Janeiro: ABNT.
- Nicholl, A. (2001). *O Ambiente que Promove a Inclusão: Conceitos de Acessibilidade e Usabilidade*. Assentamentos Humanos Magazine, 3, 2.
- NVDA (2014). *Manual do Utilizador do NVDA 2014.3*. Available at: <http://www.nvda.pt/files/html/userGuide.html> (Accessed: 20 November 2015).
- Ofcom. *Guidance on standards for audio description*. (2010). Available at: [http://www.ofcom.org.uk/static/archive/itc/itc\\_publications/codes\\_guidance/audio\\_description/introduction.aspx](http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/audio_description/introduction.aspx) (Accessed: 18 November 2015)
- Paschoal, Mariana. *INSTAGRAM: O APLICATIVO QUE REVOLUCIONOU O MUNDO DA FOTOGRAFIA* (2015). Available at: <http://blog.emania.com.br/instagram-revolucionou-o-mundo-da-fotografia> (Accessed: 20 November 2015)
- Piovesan, S. D., Wagner, R. and Rodrigues, L. (2013), *Acessibilidade em redes sociais: em busca da inclusão digital no Facebook*. *Informática na educação: teoria & prática*, 2013.
- Quadros, Yves. *6 redes sociais para prestar atenção em 2015* (2015). Available at: <http://www.8020mkt.com.br/6-redes-sociais-para-prestar-atencao-em-2015> (Accessed: 28 November 2015).
- Queiroz, M. A. (2012). *Bengala Legal*. Available at: <http://www.bengalalegal.com> (Accessed: 16 May 2016).
- Recuero, Raquel (2014). *Contribuições da Análise de Redes Sociais para o estudo das redes sociais na Internet: o caso da hashtag# Tamojuntodilma e# CalaabocaDilma*. *Fronteiras-estudos midiáticos* 16.2: 60-77.
- Ribeiro, Igor. *Instagram: 29 milhões de usuários no Brasil* (2015). Available at: <http://www.meioemensagem.com.br/home/midia/2015/11/09/instagram-chega-a-29-milhoes-de-usuarios-no-brasil.html> (Accessed: 28 November 2015)
- Santos, L. D. S. (2016). *Audiodescrição em museus: a experiência em acessibilidade no memorial dos povos indígenas*.
- SECOM. *Secretaria de Comunicação Social da Presidência da República. Pesquisa Brasileira de Mídia 2015* (2015).
- Silva, M. (2009). *Com os olhos do coração: estudo acerca do áudio descrição de desenhos animados para o público infantil*. 218f. Dissertação (Mestrado em Letras e Linguística) –Universidade Federal da Bahia, Salvador.
- Slatin, J., Rush, S. (2003) *Maximum Accessibility: Making Your Web Site More Usable for Everyone*. Massachusetts: Addison-Wesley.
- Stivanin, T. (2015) *Fato em Foco - em 2015, Instagram se consolida como umas das redes sociais mais populares do Brasil*. Available at: <http://br.rfi.fr/geral/20150105-em-2015-instagram-se-consolida-como-umas-redes-sociais-mais-populares-do-brasil> (Accessed: 28 November 2015).
- Tavares, Wellington, e Ana Paula Paes de Paula (2015). *Movimentos Sociais em Redes Sociais Virtuais: Possibilidades de Organização de Ações Coletivas no Ciberespaço*.
- Villela, L. M., & Losnak, C. J. (2016). *Abrindo os olhos sobre a Ditadura Militar: audiodescrição como recurso de manutenção da memória brasileira*. *Cadernos de Tradução*, 46-65.
- W3C (2008) *Web content accessibility guidelines (WCAG) 2.0*. Available at: <http://www.w3.org/TR/2008/REC-WCAG20-20081211> (Accessed: 28 November 2015).