LêRótulos: A Mobile Application Based on Text Recognition in Images to Assist Visually Impaired People

Juliana Damasio Oliveira[(✉)], Olimar Teixeira Borges[(✉)],
Vanessa Stangherlin Machado Paixão-Cortes[(✉)], Marcia de Borba Campos[(✉)],
and Rafael Mendes Damasceno[(✉)]

School of Technology, Pontifical Catholic University of Rio Grande do Sul (PUCRS),
Porto Alegre, Brazil
{juliana.damasio,olimar.borges,vanessa.stangherlin,
rafael.damasceno}@acad.pucrs.br, marciabcampos@hotmail.com

Abstract. The autonomy of the visual impaired person can be evaluated in day to day activities like recognizing objects, identifying textual information, among others. This paper features the OCR technology-based LêRótulos application, with the objective of helping visually impaired users to identify textual object information that is captured by the camera of an smartphone. The design of the prototype followed guidelines and recommendations for usability and accessibility, aiming for greater user autonomy. There was an evaluation with specialists and end users, in real situations of use. The results indicated that the application has good usability and meets accessibility criteria for blind and low vision users. However, some improvements were indicated. Related work is presented, the LêRótulos design process, the results of usability and accessibility assessments, and lessons learned for the development of assistive technology aimed at visually impaired users.

Keywords: Accessibility · Assistive technology · Evaluation
Mobile devices · OCR · Usability · Visually impaired people

1 Introduction

A person's autonomy with visual impairment (VI) can be evaluated in daily activities, such as recognizing objects, identifying textual information, among others. Interaction design is about designing interactive products to support the way people communicate and interact in their daily lives, whether at home or work [12]. In these circumstances, it is essential that interactive products are developed to overcome barriers faced by VI in their daily tasks. Shilkro et al. [15,16] state that blind users are interested in reading text fragments, such as restaurant menus, screen texts, business cards and canned labels. It is worth mentioning that these simple tasks, such as reading text fragments, can be a significant challenge to be overcome by a person with VI [11].

Assistive technology (AT) emerged as a way to help people with VI. In addition to screen readers, lenses and electronic magnifiers, braille printers, sticks with obstacle sensors, there are applications available on smartphones that become essential allies since it is possible to focus on a single device with different resources, reducing costs and portability [6]. However, despite these several advantages derived from mobile devices, people with VI still face difficulties on using them due to the lack of integration with applications and screen readers, problems related to handling, use requirements and device's physical characteristics, which tends to have fewer and fewer physical buttons.

Additionally, there are applications that propose to read small text fragments, through the recognition of photographs taken from these objects, e.g., Be my eyes[1], Taptapsee[2], Abbyy[3], Knfbreader[4]. However, these applications are usually complex to use, paid and/or are in English, which becomes a barrier for the Brazilian public.

In order to address this issue, we present and discuss the development and evaluation of an Android application, which recognizes texts from images captured by the smartphone's camera, intended for the Brazilian public. Called LêRótulos, this application uses OCR technology (Optical Character Recognition) available from Microsoft Cognitive Services[5] and screen reader Talkback[6], native to Android platform. Our prototype design followed guidelines and recommendations for user's usability and accessibility aiming for greater user autonomy.

The contributions for this paper are: (i) LêRótulos application design, with an accessible and usability interface for identifying texts in objects, (ii) usability studies to evaluate the application's use, (iii) that LêRótulos can be used in real-world situations and demonstrate some of the foremost problems that people with VI faces while using text-reading applications, and (iv) lessons learned with LêRótulos creation process and evaluation.

2   Related Work

Bigham et al. [2] argue the use of the VizWiz[7] application, which allows users who are blind to capture images from the environment, send them and receive information about it in real time. For this, there is a network of collaborators, which is formed by Web workers and services (software of object recognition, e-mail and Twitter), for example. The study describes that the network of collaborators has increased as there are more questions to be answered, while there

---

[1] http://bemyeyes.com/.

[2] http://taptapseeapp.com/.

[3] http://www.abbyy.com/textgrabber.

[4] http://www.knfbreader.com.

[5] https://azure.microsoft.com/pt-br/services/cognitive-services/computer-vision/.

[6] https://play.google.com/store/apps/details?id=com.google.android.marvin. talkback.

[7] http://www.vizwiz.org/.

is a small financial return for those who collaborate. This app is available on Android and iOS systems.

Jayant et al. [7] described the use of EasySnap, which assists a blind person to take pictures. In this way, the application provides real-time feedback on the image quality the camera is aiming at, also considering informations like frame adjustment, zoom level and lighting. This app is available on the iOS system.

Saleous et al. [14] research developed the software Read2Me[8], which uses OCR-based technology and text to audio conversion through Text-to-Speech (TTS). Two prototypes are presented: RPi-based Platform and Android Application. The first uses a Raspberry Pi 2 Model B (RPi) microcomputer and a camera, which can be attached to a pair of glasses, for example. Thus, the camera module, which is in the RPi, captures the image and executes the OCR that is in a service in the cloud, and then executes the TTS. The other prototype was developed as an Android application. In the comparison between the prototypes, users stated that it was easier to use the RPi, however, the accuracy of the smartphone's camera was better than that used in the RBi.

Shilkrot et al. [15] presents FingerReader, an index-finger wearable device, which makes real-time reading of printed texts as the user swipes the text. Thus, the device makes a local sequential reading, of linear and non-linear texts, from a close-up camera view. Performs on Mac and Windows machine.

The Be My Eyes application connects by video call the user with VI to a sighted volunteers network, who are able to describe what is being captured by the smartphone's camera. To access the network, the user and volunteer need to be registered on the platform. The app is available for Android and IOS.

The TapTapSee application has the most similarities with LêRótulos. Among the strengths, it has the ease of use, identification of different types of objects based on images and possibility to share the recognized text. However, it does not yet have an interface and sound system in Portuguese. Although TalkBack text is read in the configured language, all menus and device usage guidelines do not have customization for other languages. Another point to note is that the camera was customized without the inclusion of autofocus, which could make it difficult to read OCR.

3   LêRótulos Application

LêRótulos application aims to convert textual information from images captured by the camera's smartphone to audio description. The application development was based on interactive design process proposed by Preece et al. [12], which has 4 basic activities: establish requirements, (Re) design, build an interactive version and evaluate. These activities should complement each other and repeat themselves, until the end product becomes available to users. The following is a description of what was done in each step.

---

[8] https://read2me.online/.

3.1   Establish Requirements and (Re) Design

The application requirements were identified through usability goals based on [13] and accessibility. These goals were essential for the development of the application and for the evaluation with later users:

– Metas de usabilidade
  • *Be easy to remember how to use:* it should be easy to remember how to use the system. The application should be well organized, intuitive and the sequence of steps required for label recognition should be easy so the user will not forget how to perform them. Questions: Does the user make too many errors when using the system? Is a previous training phase necessary? What types of interface support are provided to help users remember how to perform tasks? Can the user perform the activities easily? Is it easy to remember how to use the application? Is the user able to use the application without needing help?
  • *Be efficient in use:* the application should allow to recognize the object label through the camera's phone. Questions: Is the user able to use efficiently the application and quickly recognizing objects? Does the user find suitable the number of clicks needed to detect an object? Is the application efficient for users to achieve their goals?
  • *Be safe to use:* the application should disable buttons that are not needed. In addition, it must protect the user from dangerous and undesirable situations. Questions: How was the occurrence of false positives in the text recognition? Does the application disable unnecessary buttons?
  • *Be useful:* the application must have commands that allow to identify object labels. In addition, it must provide the necessary functionality so that users can do what they need or want. Questions: Is it better to recognize texts through the application than through third-party help or reading Braille labels?
  • *User satisfaction:* the application should promote a good user experience for users. Questions: Does the user feel good while using the app? Does the user feel confident when using the application?
– Accessibility goals
  • *Accessibility:* the application should be well integrated with accessibility resources. Questions: Was the application well integrated with accessibility features? Has the user encountered any barriers using the application?

3.2   Build an Interactive Version

LêRótulos was developed for the Android platform, chosen for being an open source platform that brings cheaper and innovative products to customers and better development platforms for programmers [9]. The official Android website[9] has a developer area where it explains, which are the best practices in the use

---

[9] https://developer.android.com/guide/topics/ui/accessibility/apps.html.

of accessibility for both native and implemented components. These tips have been observed in LéRótulos development.

For text recognition we used an API called Computer Vision provided by Microsoft[10].This API provides image analysis services to obtain information about the visual content of an image using OCR technology, which is used to extract text from images, in a way that allows the manipulation of these texts in digital form. To use this API, one must generate a key that allows 5 thousand transactions monthly or choose to pay the service and have unlimited access.

From the usability and accessibility goals, and the mentioned technologies, the LêRótulos application was developed as it can be visualized in the Fig. 1. The operation is simple and can basically be used as follows:


Fig. 1. LêRótulos screens. (a) Initial screen with instructions for use. (b) Photography of an object. (c) Recognized photo data.

– When the application opens, a screen with instructions appears and are narrated to the user as shown: "Welcome to LêRótulos. You are in the application to recognize text from objects. To exit, press the Home key on your phone. To get started, double-tap the screen, position the subject in front of the camera, and take the picture." (Fig. 1a), and to open the device's default camera the user only need to press this screen.
– The user must position the camera in the direction of the object at a distance of approximately 20 cm and take the picture (Fig. 1b).
– Depending to the device's camera, the user may need to select the photo confirmation button for the application to start recognizing.

---

[10] https://www.microsoft.com/cognitive-services.

- In the recognition screen, the time for finalizing recognition may be influenced by internet speed.
- If the device loses its connectivity to the Internet, a message will appear informing the connection lost.
- If the image does not have text, or if it has been unreadable, a message appears stating that the text was not recognized.
- As long as there is no recognition, the user can continue shooting.
- When the caption of the object is recognized, it is spoken to the user narrated to the user (Fig. 1c).

(a)   (b)
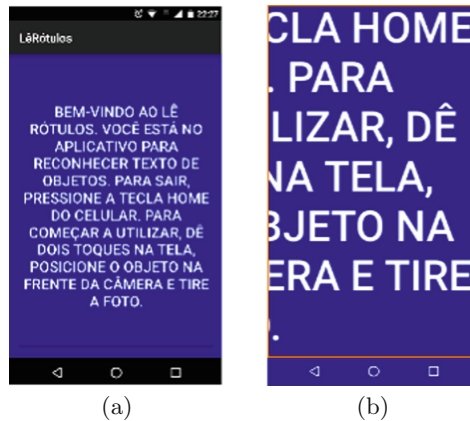
Fig. 2. Resource for low vision. (a) Contrast and colors patterns for low vision. (b) Screen magnification (Zoom).

To facilitate its use by people with low vision, it has a graphic interface, which follows recommendations of Kulpa et al. [8]. Also, the border color of the button responsible for opening the camera for capturing the photo, and scrolling text in the box with the text identified were enabled. In addition, the "Magnifier" function can be enabled on the Smartphone (Fig. 2).

3.3   Evaluated

Two evaluations were used to evaluate LêRótulos: evaluation by inspection with Human-Computer Interaction (HCI) specialists (Study 1) and evaluation with end users by observation of use and questionnaires (Study 2), described in the continuity.

Study 1 - Evaluation by Inspection. LêRótulos was evaluated by HCI and application experts, who used the inspection evaluation method, called heuristic evaluation (HE). This method is based on the 10 heuristics of Nielsen [10] to evaluate usability problems. Table 1 details the profile of these evaluators.

Table 1. Specialist profile

Specialist	Time experience	Experience in heuristic evaluation
E1	4 years	1 year and 6 months
E2	2 years and 5 months	1 year
E3	10 years	10 years

Results. We identified 19 usability problems, with some related to more than one heuristic. There are no problems associated with the "Recognition rather than memorization" heuristic. Table 2 reports the number of errors pointed out by the evaluators for each Nielsen's heuristic. Table 3 check the amount of problems encountered for each severity.

Table 2. Heuristic errors

Nielsen's 10 heuristics	Amount of errors
Visibility of system status	4
Match between system and real world	4
Consistency and standards	5
User control and freedom	3
Error prevention	7
Recognition rather than recall	0
Flexibility and efficiency of use	4
Aesthetic and minimalistic design	2
Help user s recognizes, diagnose and recover from errors	2
Help and documentation	1

We chose to present the results as [4,5]. In this way, the main identified problems grouped by violated heuristics will be presented.

1. **Visibility of system status:** the application displays sound information on the home screen, instructions, screen capture (photo) and results. However, while the app is recognizing the photo text, no feedback is given to the user. The suggestion would be to include a beep to inform that the image is being processed. Additionally, the user must be informed that he can use the camera of the phone, in the capture of the photo, in both portrait and landscape mode. It was also pointed out a violation in the use of the back button of the camera, which, instead of returning to the previous screen, remained in the image text recognition screen. Thus, even if the image was "approved", the application continued to issue the information to await recognition.
2. **Match between system and the real world:** in general, the terms and vocabularies used in the Labels were considered to be familiar. However, in

Table 3. Severity errors

Severity	Amount of errors
0 - there is no consensus on the usability problem	0
1 - cosmetic problem	1
2 - minor problem	6
3 - important usability problem to fix	10
4 - Usability catastrophe - imperative to correct!	2
Amount	19

the results screen was being issued the label "oral box", which is not intuitive. Refers to the box in which the text that is recognized at the end of the photo processing is found. Experts suggested changing to "image reading" or "text resulting from capture". Also, in the recognition screen, the back button was being read as "navigate up" button, which is not related to its function. The suggestion was to change to "back". Other buttons were unmarked and tagged.

3. **Consistency and standards:** some button-related violations have been detected. The camera button was not coming back and the label was not being identified correctly. Also, in the image recognition screen, it was informed "To take a new photo, double-tap the take photo button that is located in the upper left corner or touch the back button of the mobile". Again, the back button of the application was returning to the previous screen and did not stay on the same screen as the application.

4. **User control and freedom:** was not found a way to pause the execution of the informative text of the initial screen, being that the user needs to listen to the whole dialogue. If the user touches this screen again, the text restarts and there is no control over this feature. It was suggested that the user be able to control the progress of the text presentation, being able to pause, restart or finish its execution. Taking into account the problems already presented on the camera back button, the user should also be able to choose whether to return to the home screen, take a new photo, or wait until processing is complete.

5. **Flexibility and efficiency of use:** there is repeated information on the home screen and the photo recognition screen, which can make it tedious to the user with more experience in using the Label. Another suggestion was to be able to capture images with a click anywhere on the device screen. The application allows the photo to be obtained only with the standard camera button. Another point mentioned is when the correct recognition of the characters of the photograph does not occur. The system could inform and ask the user to rephrase the photo without having to change the screen, increasing the efficiency in the use of the application.

6. **Aesthetic and minimalist design:** the dialogs should contain necessary and relevant information, with an access point for more information if the user wishes to obtain them. Repeated use instructions have also been identified.
7. **Error prevention:** there were cases of false positives and the suggestion was that LêRótulos informs when an image was not clear. Another suggestion was to include a filter or dim the image. As for the messages in the photo recognition screen, the following instruction is given "To listen to the message again, move your finger in the lower half of the screen", but does not inform the type of slide and in which direction the movement should be performed. When the finger move right or left, all objects on the screen are reset, not the message. Additionally, errors may occur due to a missing photo history, such as if the user accidentally clicks to make a new photo, the previous one is not saved. Since it is difficult to repeat the same photo, it was suggested that it could keep the previous photo for consultation.
8. **Recognize, diagnose, and recover from errors:** there was a suggestion that when the characters were not correctly identified, a clear and informative error message could be presented to help the user understand what had happened and repair the problem.
9. **Help and documentation:** it was suggested to include a faster help option in addition to the instructions that are reported when the application is initialized.

The results of the heuristic evaluation allowed to identify problems in the interface and in some non-precise ways of using the Labels. This evaluation was complemented by the evaluation with real users, described in Study 2, and will be considered in the next version of the application.

Study 2 - Evaluation with End Users. In order to evaluate the accessibility and usability of our application and to verify if LêRótulos assists the user in the identification of text objects, Study 2 was carried out with the target audience of the application. This evaluation involved 6 main steps, which were based on [17]:

Table 4. Sample

Id	Age group	Gender	Visual impairment	Platform used
P1	18–28	Male	Totally blind (congenital)	Android
P2	51–61	Male	Low vision (acquired)	Android
P3	29–39	Male	Totally blind (congenital)	Android and IOS
P4	40–50	Male	Totally blind (acquired)	Android and IOS
P5	18–28	Female	Totally blind (congenital)	IOS
P6	29–39	Female	Low vision (acquired)	Android

1. Definition of the target audience and selection of participants: as a selection criterion, participants should already be smartphone and screen readers users. Thus, 6 participants participated, being 4 people blind and 2 with low vision, who were recruited through friends nominations (snow ball sampling technique [20]). The profile of the users can be verified in the Table 4.
2. Definition of the platform to be used: a Motorola RAZR i device with Android 4.0 operating system and native Talkback reader enabled, depending on the availability of the resource in the research group's lab. However, participants P3 and P4 preferred to use their own handsets, with Talkback enabled and with their usage preferences. LêRótulos has been installed on these devices.
3. Definition of usability and accessibility evaluation methods: the evaluation occurred with real users. The level of experience in using the *TalkBack* screen reader has been checked and demonstrated how the LêRótulos worked. Afterwards, tasks were performed to be performed with the use of the LêRótulos. To evaluate the usability of the application was used the System Usability Scale (SUS) [3], which has 10 closed questions with a 5-point Likert scale with a range from Strongly Agree to Strongly Disagree. In order to evaluate questions related to application accessibility and satisfaction of use, a questionnaire containing 13 questions was elaborated. These issues were based on the usability and accessibility goals used in the application design and in [18]. There are 8 open issues related to accessibility and 5 questions related to use satisfaction. Of these, 4 were open questions.
4. Preparation of the test: 4 tasks were developed to test the accessibility and usability of the Labels (Fig. 3).
   – Task 1 - Classification of objects of the same size: six packs of instant noodles containing three different flavors were supplied: 2 meat, 2 chicken and 2 tomatoes. The task was to identify the flavor of each of them and to group them according to the flavor (Fig. 3a).
   – Task 2 - Identification of objects of equal size: Three packages of identical drugs were made available in size and shape and needed to identify the name of the drug and chemical compound present in each (Fig. 3b).
   – Task 3 - Identification of objects of different sizes: similar to task 2, however, using three drug packages with different sizes and text fonts. The participant needed to identify the name and compound of the drug (Fig. 3c).
   – Task 4 - Identification of text of business cards: three business cards with different texts and fonts were made available. The objective was to identify the text of these cards (Fig. 3d).
5. Evaluation: it was done individually and in the place of preference of each participant. Initially, a questionnaire was filled out that contained 11 questions to identify the profile and register the user experience with smartphones. After completing the tasks, the participants answered the SUS questionnaires and the questionnaire of accessibility and satisfaction of use. There was video recording, audio and photos of the tasks being performed, with the informed consent of the participants. The execution time and errors committed in each task were recorded.

6. Results analysis: the videos, the evaluators notes related to the observations of use and the participants answers to the evaluation questionnaires were analyzed. The length of trial sessions varied depending on the level of user experience on smartphone usage, Android operating system, and TalkBack reader. The results are described below.

(a)

(b)   (c)   (d)
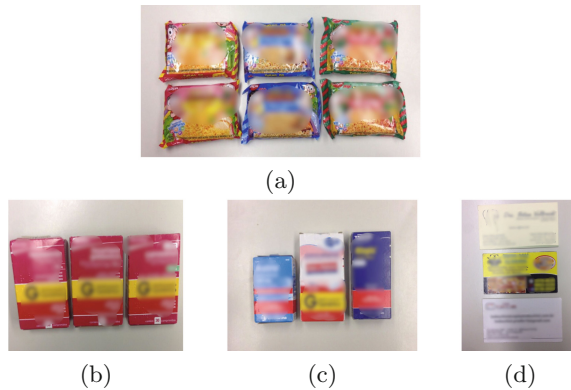
Fig. 3. Tasks (a) Task 1, (b) Task 2, (c) Task 3, and (d) Task 4

Results. The profile questionnaire allowed to verify the experience of the participants with the use of Smartphones. Among the uses were: phone calls, social networks, e-mail, clock, contacts, text messages, weather, Youtube, games and calendar. Participant P4 has already used sound recorder and musical instrument tuner. Additionally to these applications, users were asked about the use of more specific applications where they reported using audiobooks, ballot readers, barcode readers, homebank applications, color identifiers. Participants P4 and P5 had already tried other applications to read texts in objects. P4 was quoted as Be my eyes, Docscanner, TestGraber and Prizmo applications. Already the participant P5 quoted the KNFB Reader. Participant P6 reported that it uses the digital loupe feature instead of screen reader. User P3 has already used the Talks screen reader of the Symbian platform and the user P4 the screen reader ShanePlus.

After completing the profile questionnaire, the participants started the Tasks. The execution time of the tasks was counted from when the participant started the task until its completion (Table 5). The total participation of each user lasted on average 1 h and 30 min, including the time of the training and to respond to the questionnaires. One of the requirements for the application to work is that the user is connected to the Internet. In the case of the participant P4, access to the internet was made by mobile cellular network (3G), unlike the other participants, who used Wi-Fi connection. This reflected in the execution time of the tasks, because when the connection failed, it was not possible to recognize the text.

Table 5. Execution time per participant

Id	Task 1[a]	Task 2[a]	Task 3[a]	Task 4[a]	Amount of time[a]
P1	09:22	05:03	03:48	03:57	21,30
P2	19:39	03:34	02:18	10:26	35,17
P3	11:31	02:59	02:08	07:44	23,42
P4	11:13	15:51	03:32	15:43	45,39
P5	08:15	08:05	03:58	04:13	23,91
P6	08:16	04:24	01:45	03:09	16,94
Amount	67,36	38,76	16,09	43,92	166,13
Average run time	11,22	6,46	02,68	7,32	27,68

[a]in minutes

Table 6. Amount of errors per participant

Id	Task 1	Task 2	Task 3	Task 4
P1	6	0	0	2
P2	8	0	0	1
P3	17	1	0	8
P4	2	5	0	–
P5	4	2	1	3
P6	5	1	0	0

The Table 6 informs the number of times that each participant had to repeat the image capture, which is being described as a read error, or only as an error. Despite the errors, the tests demonstrated that the application is effective in identifying the labels of the tested objects, because all the users were able to complete the tasks successfully. Considering the time for carrying out the activities and the number of errors, it can be concluded that the use of LêRótulos was more efficient in Task 3, in which packs of medicines of different sizes were compared. During the execution of Task 4 by participant P4, there was a problem in the recording that prevented the counting of errors during the activity.

The Fig. 4 illustrates two participants performing the tasks. At Fig. 4a, the participant P1, who is blind from birth, performing Task 1 and in Fig. 4b we have the participant P6, who has low vision performing Task 2.

After completing the tasks, the participants answered the questionnaire SUS and accessibility and satisfaction of use. The SUS questionnaire score was 90 points. According to the satisfaction scale of Bangor et al. [1], means that LêRótulos has been rated as excellent and with a high level of user satisfaction. Still, according to Tenório et al. [19], it is possible to recognize the usability principles indicated by Nielsen [10] in SUS issues. In this way, we have:

– Ease of learning: Questions 3, 4, 7 and 10 of the SUS questionnaire. One of the questions to be answered with the evaluation was to see if it is easy to learn

(a)   (b)
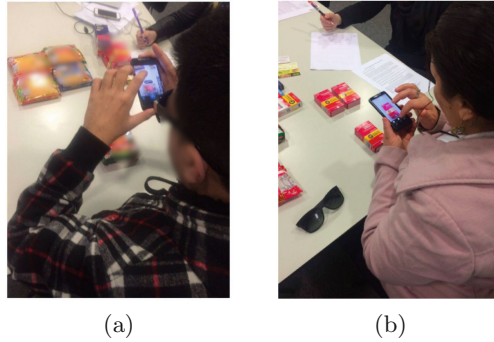
Fig. 4. Participants performing the tasks (a) user who are blind and (b) user with low vision.

how to use the application. All participants agreed that the application was easy to use (question 3) and they figured that people will learn how to use it quickly (question 7). With regard to learning, there was a suggestion that it was necessary to learn more about moving the phone and the screen reader to better use the application (question 10). P1 justified that he had never used the smartphone camera and needed to learn how to take pictures and the P5, who was a user only of the IOS platform, had not yet used Talkback. Only P1 reported that he would need the help of a person with technical knowledge to use the application (question 4).

- Efficiency: Questions 5, 6, 8 of the SUS questionnaire. All participants disagreed that the application was confusing to use (question 8). Only the participant P5 disagreed with the statement of question 5. The reason the participant described was that the application should use a camera of its own and not the standard camera of the smartphone because it considers it difficult to use. Three participants stated that they partially agree that the application presents a lot of inconsistency (question 6), since the text recognition had flaws that were described in the execution of the tasks. Additionally, P3 reported that failure in recognition is a common error that occurs with OCR:
  - *(...) I think only the recognition of letters, but this is a problem that has in all screen readers. It is not really a barrier of yours, but a barrier of materials.* (P3)
- Ease of memorization: Question 2 of the SUS questionnaire. All participants disagreed that the application was unnecessarily complex, suggesting that the system could be easily memorized and that users could easily remember how to use it.
- Minimization of errors: In the SUS questionnaire is represented by the same question 6 answered in the previous item Efficiency.
- Satisfaction: Questions 1, 4, 9 of the SUS questionnaire. All questions regarding usage satisfaction were ranked with full agreement by participants, which meant that participants were satisfied with their use.

In addition to the responses of the SUS questionnaire, the participants answered the instrument of accessibility and usability. In the accessibility questionnaire, participants report that the Talkback accessibility feature does not work properly with the Android camera. Depending on the model of the device, there are buttons that make it difficult to use, and there are applications where the button description labels are not correctly identified. Here are excerpts from testimonials:

– *I just think it needs someone to give the initial tricks of how to position the object in front of the camera.* (P1)
– *Only with the camera, it is difficult to find the button to take photo. My suggestion is to enter direct with the take photo button selected and only have this option on the camera."* (P2)
– *I found no barrier. Because the application is ok. The limitation is not the application but the camera. Does not apply to the application. Maybe I could take the picture straight away.* (P3)
– *The camera could have names on the buttons. (...) I had a bit of trouble with using the camera, but the problem was with Talkback not the application. I did not know the commands to take a photo.* (P4)

Participants with low vision (P2 and P6) answered questions about screen contrast. The participant P2 validated the contrast as appropriate, but preferred to use the audio to avoid straining the eyes. P6 stated that the contrast is adequate.

As far as audio information is concerned, participants report that the audio information was appropriate, well-crafted, creative, simple, and easy to understand. All participants agreed that they were able to manipulate sound information as to whether to repeat and stop the audios and that the application responded well to the Talkback reader. P3 made suggestions for improvement in audio information:

– *Maybe he could have a little help button that had all that information, or just take it out. Because people who have more advanced usability, the less they talk the better (...) I suggest a button on the screen that could activate. A button setting that becomes customizable. This is to see if you want the information or what you do not want. (...) <sound> information will be unnecessary for those more accustomed users. Low vision, for example, accustoms him to run the screen to read, and the blind man gets used to the information by listening once or twice (...) Usually applications have a help button.* (P3)

In the questions related to use satisfaction, suggestions for improvement were requested for application, and again, the camera issue was highlighted:

– *Just the question of the camera, my suggestion is that you enter direct with take photo button selected and you only have that option in the camera.* (P2)
– *Integration with the camera decreases the number of clicks. Greater integration with the camera.* (P3)

– *The camera might not request confirmation when taking the picture and go straight to recognition. The buttons on the camera are also labeled. And it would be interesting to copy the recognized text.* (P4)
– *It was very simple, very easy for me. Adapt the touch sensitivity a bit more. If I were to wear this, it would be perfect. My only difficulty was sensitivity.* (P5)
– *Let the application recognize people and larger objects as books.* (P1)

Also, with respect to the satisfaction of use, were cited:

– *Was able to take pictures and see the things I have. It is easy to use, to handle. It's not boring to deal with. (...) I found it very accessible, very calm to use.* (P1)
– *I found it excellent, makes it easy for the handicapped. It's very objective.* (P2)
– *Facility it provides to independence. We have not even tested this with money, but I think he'll read it too. I think you are to be congratulated both for the initiative and for the effectiveness of the project. It's a project that works, it's working well, so that's it.* (P3)
– *I found the recognition very good. The OCR used has a very good quality of recognition.* (P4)
– *It's a practical application, it does not have to come in many screens, there are some that are more complex, sometimes other applications have to keep fighting to recognize the text, and the amount paid is not worth it. That he has the option to repeat the reading of the text, and I think that the fact that he is good from day to day, I am sure that I will use direct. (...) He is very simple.* (P5)
– *That he is quick, that he is easy, that he is complete.* (P6)

Finally, the participants were asked if they would indicate the use of the application to other people, they all answered yes. Here are some testimonials:

– *Yes. And I also want to, because the application is very useful.* (P1)
– *For everyone, the application facilitates a lot. I even got confused when I took medicine because the boxes were the same.* (P2)
– *Would indicate. Because it has great utility to identify labels. Of all the applications I've tested in this regard, I think it's pretty cool.* (P4)

4   Lessons Learned

The development and evaluation process of LêRótulos has brought lessons learned, which are expected to assist other researchers in the area of assistive technology for visually impaired users. Among these, some are mentioned:

– Experienced participants suggested having an option to access the instructions for use instead of automatically making them available at each start of the system.

– The use of cellular resources, such as the camera, was not well accepted. When the LêRótulos application was started, it was decided to use the camera native to the cell phone. This decision was made believing that it would be easier for people with VI to use their own camera, which they had more familiarity with. In view of the problems encountered by specialists and people with VI, it is believed that this decision presented more problems than advantages. In this way, it is suggested to implement a camera, in which one can have friendly labels on the buttons, reduce the amount of these and customize their options and way of presenting the information on the screen.

– In the evaluation with people with VI it was possible to verify the interaction of the user with the cell phone. For example, it was possible to observe that instead of using the finger at each corner of the screen searching for information and identifying the buttons, most of the users touched only in the middle of the phone, traversing all the components of the screen. This highlights the importance of performing evaluations with the target audience of the application to understand the form of interaction and their needs.

– The Talkback speed used by participants with VI is superior to human speech. For this reason, each user must set the speed according to their preference.

– There was difficulty in the composition of the sample of the end users, even using the technique of snowball.

– The impact of choosing the location where the assessment will take place in controlled environments or natural environments should be checked, as the need for internet access or environmental characteristics (such as brightness) may interfere with the results of the evaluation.

5   Conclusion

The research aimed to introduce the application LêRótulos and evaluate if it meets its main objective, which is to help people who are visual impairment to recognize texts that are in images that are captured by the Android smartphone camera.

Inspection assessments were carried out with HCI and application specialists (Study 1), and evaluation with application end users, in real situations of use (Study 2). The results indicated that the application fulfills its function with good usability and accessibility, but indicated improvements for the same, which were described in the sections of this work. As future work, we intend to make the usability and accessibility corrections indicated in the evaluations as well as incorporate improvements, mainly, related to the camera.

Last, but not least, it should be noted that the P3, who has congenital blindness, used LêRótulos to assist him in the enrollment process at a university. The usage situation was that he needed to select different documents and there were no monitors to assist him. Through the application, you can recognize the documents, and separate the ones you needed to present. He thus exercised his right to access information and autonomy.

References

1. Bangor, A., Kortum, P., Miller, J.: Determining what individual sus scores mean: adding an adjective rating scale. J. Usability Stud. **4**(3), 114–123 (2009)
2. Bigham, J.P., Jayant, C., Ji, H., Little, G., Miller, A., Miller, R.C., Miller, R., Tatarowicz, A., White, B., White, S., Yeh, T.: Vizwiz: nearly real-time answers to visual questions. In: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology, UIST 2010, pp. 333–342. ACM, New York (2010). http://doi.acm.org/10.1145/1866029.1866080
3. Brooke, J., et al.: SUS-a quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
4. Cortes, W.R.P., Zanin, A., Soletti, L.V., Machado, V.S., Silveira, M.S., da Silva, P.H.L.: Zumbis vs sedentários: Quem irá vencer? avaliando a usabilidade do aplicativo zombie's, run! In: Companion Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems, IHC 2014, pp. 143–157. Sociedade Brasileira de Computação, Porto Alegre (2014). http://dl.acm.org/citation.cfm?id=2738165.2738206
5. Cunha, B.C.R., Machado Neto, O.J., Pimentel, M.D.G.C.: A heuristic evaluation of a mobile annotation tool. In: Proceedings of the 19th Brazilian Symposium on Multimedia and the Web, WebMedia 2013, pp. 89–92. ACM, New York (2013). http://doi.acm.org/10.1145/2526188.2526232
6. Damaceno, R.J.P., Braga, J.C., Chalco, J.P.M.: Mobile device accessibility for the visually impaired: problems mapping and empirical study of touch screen gestures. In: Proceedings of the 15th Brazilian Symposium on Human Factors in Computer Systems, IHC 2016, pp. 2:1–2:10. ACM, New York (2016). http://doi.acm.org/10.1145/3033701.3033703
7. Jayant, C., Ji, H., White, S., Bigham, J.P.: Supporting blind photography. In: The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS 2011, pp. 203–210. ACM, New York (2011). http://doi.acm.org/10.1145/2049536.2049573
8. Kulpa, C.C., Teixeira, F.G., da Silva, R.P.: Um modelo de cores na usabilidade das interfaces computacionais para os deficientes de baixa visão. Des. Tecnologia **1**(01), 66–78 (2010)
9. Maasalmi, E., Pitkänen, P.: Comparing Google's Android and Apple's iOS mobile software development environments (2008)
10. Nielsen, J.: Usability Engineering. Morgan Kaufmann Publishers Inc., San Francisco (1993)
11. Prates, D.: Acessibilidade Atitudinal. Gramma, Rio de Janeiro (2015)
12. Rogers, Y., Sharp, H., Preece, J.: Design de interação: além da interação humano-computador. Bookman (2013)
13. Rogers, Y., Sharp, H., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. Wiley, New York (2011)
14. Saleous, H., Shaikh, A., Gupta, R., Sagahyroon, A.: Read2me: a cloud-based reading aid for the visually impaired. In: 2016 International Conference on Industrial Informatics and Computer Systems (CIICS), pp. 1–6. IEEE (2016). https://doi.org/10.1109/ICCSII.2016.7462446

15. Shilkrot, R., Huber, J., Liu, C., Maes, P., Nanayakkara, S.C.: Fingerreader: a wearable device to support text reading on the go. In: CHI 2014 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2014, pp. 2359–2364. ACM, New York (2014). http://doi.acm.org/10.1145/2559206.2581220
16. Shilkrot, R., Huber, J., Meng Ee, W., Maes, P., Nanayakkara, S.C.: Fingerreader: a wearable device to explore printed text on the go. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, pp. 2363–2372. ACM, New York (2015). http://doi.acm.org/10.1145/2702123.2702421
17. da Silva, C.F., Ferreira, S.B.L., Ramos, J.F.M.: Whatsapp accessibility from the perspective of visually impaired people. In: Proceedings of the 15th Brazilian Symposium on Human Factors in Computer Systems, IHC 2016, pp. 11:1–11:10. ACM, New York (2016). http://doi.acm.org/10.1145/3033701.3033712
18. Sonza, A.P., et al.: Acessibilidade e tecnologia assistiva: pensando a inclusão sociodigital de pessoas com necessidades especiais. BBB, Bento Gonçalves (2013)
19. Tenório, J.M., Cohrs, F.M., Sdepanian, V.L., Pisa, I.T., de Fátima Marin, H.: Desenvolvimento e avaliação de um protocolo eletrônico para atendimento e monitoramento do paciente com doença celíaca. Revista de Informática Teórica e Aplicada **17**(2), 210–220 (2010)
20. Weiss, R.S.: Learning from Strangers: The Art and Method of Qualitative Interview Studies. Simon and Schuster, New York (1995)