Usability on Screen Readers for mobile phones: A critical vision

Jorge Sepúlveda
Department of Computer Science and Engineering
IST/Technical University of Lisbon
jorge.miguel.sepulveda@gmail.com

Abstract

Nowadays there are thousands of mobile phones all over the world and almost everyone on civilized countries has at least one. These devices are present in our daily lives both for personal and professional purposes and let us perform a lot of tasks beyond communicating. That’s true for almost everyone except for blind users.

Blind users still use their mobile devices mainly for communication through phone calls and short text messages. Most of the remaining functionalities are not accessible to this kind of users because blind people face a lot of problems using their devices. To deal with these problems the most popular solution is the mobile screen reader.

Our study is an evaluation about the behavior of this solution, pointing out its flaws and the spots where it can be bettered. As a result of this evaluation we produced a few guidelines to use during the process of designing mobile user interfaces for visually impaired.

Keywords
Usability, Mobile, Visually Impaired, Screen Reader, User Interface Design Guidelines.

1. Introduction

Due to technological developments of recent years the ability of processors has increased significantly but its size and cost become increasingly smaller and cheaper. This also happened with memory, substantially increasing its capacity and reducing its physical size and its market price. These technological advances enabled the evolution of mobile devices and its capabilities.

In today's society the speed at which information travels is much greater than a decade ago. Such changes mean that today it’s increasingly important to be updated with the latest information and be able to access it anytime and anywhere.

These factors make the mobile phones an increasingly essential and indispensable device for the everyday life. To confirm this trend we can check on the strong growth that occurred in the penetration rate of mobile devices in Portuguese market. According to ANACOM [1] figures, from the first quarter of 2010, there’s a penetration rate of 149% of mobile phones in Portugal which is higher than the European average which stands around 122%. This data means that there are on average more than one mobile device per capita.

With all this technological evolution the vast majority of mobile devices are no longer mere mobile phones but small sized personal computers offering an extended range of features and being no longer simple communication devices to make phone calls and send text messages.

This new generation of mobile devices offers lots of features to its users, whether for personal or professional purpose. There are features that allow users to manage an address book, send and receive emails, surf on the internet, take photographs or manage personal tasks, just to name a few of the most common.

Despite all this technological development and the offered features in mobile devices, there are still some part of the population that experience serious difficulties in using most of these features. A big part of the population that deals everyday with these problems is the blind people.

Regarding these problems, our work aims to contribute to their solution. Our contribution is a study about the current solutions for the use of mobile devices by visually impaired, a study about the interaction between visually impaired and their mobile devices, an Heuristic Evaluation of a screen reader, Guidelines for the Design of user interfaces for the visually impaired, and some practical solutions for the encountered problems.

2. Related Work

In order to try to decrease the gap between visually impaired users and mobile devices, there are a range of approaches aimed at resolving or addressing some of these flaws. The solutions can be divided mainly into three broad categories: custom made hardware, accessories and hardware adaptations to existing mobile devices and software-based solutions.

Given this division will be presenting a list of the most relevant solutions currently on the market or under development. This will not be an exhaustive list of all existing solutions but a list of the ones that have the most relevant aspects. We’ll also provide a brief analysis on the advantages and disadvantages of each approach in order to identify gaps that are not currently met or problems for which current solutions are inefficient.

2.1 Custom made hardware

There are a wide range custom made hardware solutions based on Braille alphabet communication. Initially appeared a set of solutions based Braille keyboards and Braille Cell displays that
were mainly for word processing, having some of them the capability to manage contacts and perform some other minor tasks.

Later on, some of these devices enabled some of mobile phones capabilities but most of them where only able to be used through a real mobile phone. These devices connected with the mobile phones, through a docking station or through infrared communication.

These devices were designed ultimately not succeed due to their high price and their low portability. These devices were very large and weighted too much for a portable device. Furthermore the dependence of other devices to perform some of the mobile phone tasks compromised even more its portability.

Later on two new and totally different proposals arose. One of them was a low cost braille based mobile phone but it only allowed phone calls. The other proposal was also based on braille. It was a device developed for sending and receiving text messages. Both this devices did not prevail because their capabilities were much reduced.

The custom made hardware solution did not prevail because it has lots of constraints that compromised its success. It’s price was one of the problems, the weight and dimensions of their devices were also a problem and the set of functionalities that they offered was very reduced.

2.2 Accessories and hardware adaptations to existing mobile devices

In this category there were a few interesting solutions like the Sypole [2] or 3GM [3]. These solutions rely on a thin plastic film that is placed on the screen of a PDA and simulates various buttons on the touch screen. This film simulates different behaviors depending on the approach in question.

Sypole uses the film to simulate a keyboard. This solutions also has a voice synthesizer that provide the necessary feedback on the activities performed by the user, helping him to navigate through the menus for device applications.

In the case of 3GM the film is used to simulate a set of nine buttons. Four of these buttons map the alphabet or the numbers in a quadripartite way, the remaining are used for navigation, the introduction of the "space" character, to validate content and use help. Navigation between the letters or digits of each key is made by successive touches on that key traversing the letters mapped there.

There’s also EdgeWrite [4] that was conceived for motor-disabled users because of its lack of precision that did not allowed them to write correctly. This solution proposes an adaption to a PDA that limits the writing to a small square made of acrylic. This confined interaction area could be a good solution for visually impaired because they would have a physical reference to know where they were writing.

These kinds of solutions have some interesting aspects but they only contribute to the resolution of a few isolated problems and also force the user to have more than just one physical device. Visually impaired users want to have the same common devices that the sighted users do.

2.3 Software-based solutions

There are a few software-based solutions word mentioning. Regarding the problem of text input with T9, these solutions are the NavTap [5], NavTouch [6] and BrailleTap[5]. These solutions propose a totally approach for text input.

The NavTap text input method aims to reduce the cognitive load inherent in memorizing the positions of characters on the keyboard in order to allow the user to navigate to the desired one. In this solution the navigation is made through a matrix that contains the letters of the alphabet arranged in sequence with a vowel positioned at the beginning of each line followed by other letters of the alphabet until the next vowel.

The navigation through the character array is made by using four keys on the keyboard (2, 4, 6, 8) that are arranged to form a cursor around the center key (5), indicating the four possible directions for navigation. There’s a button reserved for the introduction of spaces in the text, another to delete the last introduced character another to listen to the suggestions made by an auto-completion method.

Figure 1 – Text input method used on NavTap and NavTouch and pointer used on NavTap.

NavTouch has the same operating principle as the NavTap, the main differences between the NavTap and NavTouch are in the way navigation is done through the alphabet of letters available for selection. NavTouch uses a touch screen allowing the user to navigate through the character array of characters directional gestures made on the device display.

BrailleTap, has a different approach using the device keyboard to input Braille text. It’s mapped a common matrix of six cells (2, 3, 5, 6, 8, and 9) on the device keyboard. This matrix represents the six points of a Braille character. This mapping allows you to press a certain key in order to fill or empty each of the points to build the desired Braille character as shown in the Figure 2.

Figure 2 – BrailleTap keyboard.

These are tree really interesting solutions for the text input problem. The results during the test phase of these solutions were
very promising. Probably it could be the best way of dealing with text input by visually impaired on mobile devices.

Despite the quality of these solutions they only deal with the problem of text input. There’s a more comprehensive that presents solutions for a wider range of problems present in mobile phone interaction. This solution is the screen reader.

It consists of a software application that is installed on a mobile phone and renders the information displayed on the screen through synthesized speech output generated using text-to-speech (TTS) technology and routed through the device’s speaker or a headset.

The screen reader is the most common solution for blinded people’s problems with mobile devices and the one that blind users like most. Because the screen readers are the most popular solution and because it’s the one that is most close to a better solution it was the one on which we focused our study.

### 3. Understanding Mobile Screen Readers

#### 3.1 A Typical Screen Reader

A screen reader is an application that runs over the operating system of a mobile phone and transforms the information showed on the mobile display into auditory feedback.

As the name says, this kind of application just reads the information displayed on the screen and plays it through auditory feedback. There is no interpretation of the type of content that is shown on the screen and the information is transferred from one channel to another without any concern on adaptation of the information to reality of these distinct channels of communication.

The auditory feedback is done through a speech synthesizer (TTS - Text-to-Speech) whose function is to reproduce human speech in an artificial way.

The screen reader has the great advantage of allowing a blind user to get additional feedback from a mobile device in addition to the auditory feedback that comes by default with the operating system of the phone, which is clearly inadequate for their special needs.

The additional auditory feedback allows the blind user, to perform a series of tasks that previously was extremely difficult or virtually impossible to accomplish. Therefore, this solution is regarded as a major contribution by the vast majority of these users.

However despite of the ability of allowing an increasing on the number of performed tasks and interactions between blind users and their mobile devices it’s not good enough to allow a totally efficient and effective interaction between user and device and does not allow the user to use all the features available on their phones.

#### 3.2 User Studies

We conducted a study with seven blind users that performed a series of tasks in order to achieve some information about their behavior regarding the use of mobile devices.

The information gathered during this study allowed us to have some practical information that complemented the other information collected at other stages of this study. The information that we collected was grouped by categories because of its relation between the different problems.

### Navigation

The way the menus of the phones are structured is too complex for those who cannot see and they need to recreate a mental structure to orient themselves within them. The menus are designed for mobile phone users who can afford to get visual feedback of the device and for which there exists the same cognitive load as for a blind user, as a sighted user can know the exact location where you are looking through a contrary to what happens to a blind user.

TALKS convert what was idealized to be visual feedback into auditory feedback. This solution turns out to be peaceful and effective in some contexts where the complexity of the content is reduced and where this mapping becomes more effective. However, if the menu structure and navigation is too complex, this solution is worse than expected by users, and does not truly solve their problem.

The indications that TALKS do along the path does not make it easy to understand the system status and the hierarchical structure of the different menus. This means that users often feel lost when navigating within the menus, particularly in areas that used less frequently in a daily basis. As a result the users do not feel confident to explore features that ignore and to explore areas of mobile phones which are not used.

This problem is something that is not unique to blind people since it is commonly found in public with greater difficulties in using new technologies. This type of public has a greater tendency to feel "lost" while using a mobile phone. One example where this situation exists is often among the elderly population.

However despite of many difficulties that the sighted people can feel, they are able to get the information through multiple looks into the screen, becoming aware of the status of the device and its context in.

For the blind people, using TALKS, the situation is somewhat different and more worrying, as there is no room for gradual adjustment to the status of the system and context is easily lost. Unlike sighted users, that can make their uptake of information, increasing its complexity and "digesting" it at their pace, the blind are unable to do so.

The information is transmitted to them in sequence and total (at once) without allowing them to distinguish different types of information, without establishing any hierarchy relationship and mixing the context of available actions and other distinguishing information.

### Help and Context

Users often arrive to a menu entry and fail to see its contents. The same is also true within the menus, once the user reaches a certain functionality. It becomes difficult to realize the objective of a particular feature only by their name because, according to the manufacturer and model of phone, there are different features with the same name and above all different names for the same functionality, depending on the mobile phone.

In general it is difficult to realize what’s the extent of functionalities available on a particular mobile phone, what each one them really does and where to find them.
Within a given functionality, e.g., the calculator, it is difficult for a user to know which buttons do what, for example knowing where the plus sign or the comma are situated. In this example, the calculator is a feature that people even know quite well their function, remaining just the difficulties in finding a particular button.

There are additional difficulties however, when a user comes to a feature where he wants to set time and date and the system allows the possibility to choose time zone, date format, time format, current time, current date, etc..

For the user, it is often difficult to understand the context of the situation where he is. This situation is very common within the forms, where for example at the introduction of a contact, the user only has the notion that there are many fields as the filling goes. You just know that there is a text field for the surname after you complete the name text field. The user may even have already placed the information on the wrong field that he will only know it at the end. Another example of situation like this is when the user type in the content of a message into the recipient text field and only understands it afterwards.

This way of organizing form fields is very error prone. This makes users feel more uncomfortable with the use of these features and thus makes them use it less.

Auditory Feedback Control

One of our findings during the observations was the fact that users do not feel a real control over the audio feedback. Many times the users wanted to change the playback speed of audio feedback while performing a particular task, and were unable to do so. This situation was particularly worrying while reading text messages.

The lack of control over how they control audio feedback contrasts with the free form as sighted users make use of these devices and how to perform their tasks. A sighted user reads the information present on the phone screen in order to suit his abilities and needs.

TALKS plays messages at the speed that is configured on the phone and at once turning it very difficult for the user to understand the whole message remaining with just an idea of what its content. There was a way of changing the feedback speed of TALKS but it affected the speed of all the TTS software, and was not something momentary, forcing the user to change it again later to have the desired speed in the rest of the features.

There was one user who could use a feature of TALKS, unknown by most, allowing him to read a written message, sentence by sentence. However this was only possible after a first quick reading of the message, automatically done by the system. The fact that this user was amblyopic cannot be overlooked, because he has an extra sense that a blind user does not have since he has a residual capacity of vision.

Text Input

Another problem often pointed out during the study phase was the text input. In almost all occasions, whenever it was necessary to introduce a piece of text, either on simply assigning a name to a contact or writing a message there were difficulties of most users.

The problem was both with the introduction of special characters, accents or punctuation and with misspelled words. On one hand was difficult for some users know where they were to find certain characters or punctuation marks, on the other hand, was also common for users to write misspelled words without knowing about this situation.

The system default text input, multi-tap, is not intended to be used by blind users. The need to keep in mind the mapping of all characters on the keyboard, it is very painful and difficult to achieve it. The alternative is to listen to the details given by the TALKS audio feedback but the information that is provided is insufficient, making it often difficult to discern the key that we select and if it has already been introduced or not.

3.3 Heuristic Evaluation

At this stage of our investigation we conducted a heuristic evaluation, carried out by experts in usability. This evaluation had as its main objective to allow us to have more critical view of the usability and functioning of the screen reader and simultaneously expand our horizons in order to have some different perspectives about the application.

At this evaluation we identified some usability flaws that caused most of problems previously detected and that severely conditioned the user experience over the mobile phone with screen reader. The following is a summary of the main flaws detected, making a parallel with the heuristics of Nielsen.

Home Screen

When a user arrives or returns to the home screen, the screen reader reads all the information present on the screen. This screen contains information concerning the operator, current date and time and what actions are available to accomplish through the options available in the lower corners of the screen.

This information should only be available when requested by the user. The user does not control or exercise free will as the Nielsen heuristic advises. Providing all this information every time we stand in the screen goes against the recommendation of aesthetic and minimalist design where the user is always confronted with irrelevant information.

Menus and Navigation

When a user browses the menus, TALKS, only informs the placement of a particular option in the options list or menu at the first time or when we turn back. If someone navigates the menu options, the screen reader, let us show for example that we are at the option 4 out of 12.

Once again the state of the system is not visible, because we do not know all the information needed to assess the state where we are. Moreover, this behavior does not satisfy the heuristic that indicates that the system should have consistency and adherence to standards and should not have different behavior for identical situations.

The absence of additional information on the menu takes, for example, that we can give a complete circuit around the menu without realizing that fact. To do this we would have to remember that had already passed through a particular option instead of easily recognizing it by the indication of the index of the list of options, for a given menu.

Within the menus is also often to lose all sense of context because we the user not have the notion of path. The user is only informed
of the name of a certain menu for the first time he goes into it, this information is never transmitted again.

When navigating through the options of a particular menu, sometimes the user is given an indication that there is a submenu. This information is in direct competition with the contents and further complicates the task of navigation.

The design should be minimalist, with information about the submenu only available if the user requests it.

When the screen reader reads the options associated with the lower corners of the screen refers to the keys 1 and 2. More than once the reviewers interpreted these keys as alphanumeric keys on available on the keypad of the phone.

In fact these keys are the ones on the upper corners of the keyboard, which could probably be considered as key option one and two or more properly left key and right key option.

This behavior goes against the heuristics which recommend that the system should speak the language of the user. This situation helps the user make more mistakes.

Abbreviations, descriptions and graphic signs

The whole system is full of abbreviations that make no sense in oral communication, as not meeting the heuristic indicates that the system should speak the language of the user.

There are many unfamiliar names for certain things and even worse is that many times there are different designations to refer to exactly the same thing. The system mentions, for example, "msg", "short message", "sms", "written messages" or simply "messages" to refer the exact same thing. Besides from not speaking the user language it is not consistent.

Another feature also very present throughout the system is the use of graphic signs such as punctuation or parentheses that only make sense in a written communication. Such inscriptions are not to be transported to the audio channel. In this case we have a situation where the system does not speak the language of the user. For example the system reads "open parenthesis one bar seven close parenthesis " when it should read something like "one out of seven".

Forms

Whenever a user is faced with a form to fill with content, TALKS, reads him the words "text" as being to enter text, though sometimes the form is not intended to writing text but for digits. This goes against the recommendation to avoid bias and shows that the system does not speak the language of the user.

When a user comes to a form where they are needed to fill various fields, the system provides too much unnecessary information and does not provide information that enables the user to know the content type that you enter at a particular field in a form and which fields are present in the form where you are.

In this case the system does not offer a minimalist design and does not provide help and documentation to allow the user to assist in its task of filling out the form. The method of exit and re-enter also works but is not efficient as suggested by the heuristic that advises flexibility and efficiency in the system.

In some forms, the user, only obtains information about the content after exiting the field and returning to it again. There should be a simple and direct way for the user to know the content they just introduced to avoid errors and so the user can always have present the state of the system.

Text Input

This is one aspect that has more flaws in the process of interaction between user and device. The usability of text input is poor and can become a real headache for users.

The method of standard text entry, multi-tap, is very difficult to use by blind users, since the time of acceptance of each key is very short. The acceptance time is identical to the time that the synthesizer takes to provide information about the selected character. This means that, often, you end up writing a certain character when you only wanted to know whether the selected character was intended to proceed to the next character.

This means that blind users are almost obliged to memorize the layout of the keyboard characters, knowing that the key is associated with each character and how much time is necessary to press certain key to achieving the desired key. Besides being forced to memorize this information are also required to have the necessary dexterity to be able to recall all of this information every time they wish to introduce some character. Varying the time of acceptance of characters between 1 to 2 seconds, we are talking here about a very high level of skill, especially when taking into account novice users..

This situation becomes even more complicated when users want to enter special characters, which are used less frequently and therefore are placed in more advanced positions in the list of characters associated with each letter.

Besides this problem of high severity, there is an additional subject that makes it difficult to correct text input. When we introduce a character not given us any message informing that the character was introduced. For example, when trying to find a "will" can make a full turn to the list of characters available through two key phone, turning the letter "a", or simply introduce a new character “a” because the information is exactly the same.

These types of errors usually are only noticeable when the screen reader reads a word, after the introduction of a space. In this instance the error may be noticeable and is phonetically incorrect, otherwise the user will never notice it.

All the issues described here violate with high severity some of usability heuristics. The system state is often unclear, not knowing where and what the user is typing. The user is not exercising free will, he has to enter characters according to the pace of acceptance of characters set by multi-tap which is not thought to auditory feedback.

This method of entering text, promotes the occurrence of errors even as a way of learning. Often users are obliged to make mistakes to understand the action they have done and the system state. This form of text entry encourages memorizing, for it requires having a mental mapping of all characters available and what keys where they are. This system clearly does not help the user recognize, diagnose and recover from mistakes. Finally, the provision of help and documentation is clearly reduced or absent depending on situations.

As can be seen on the ten recommendations made by the heuristics of Nielsen, this method of text entry violates at least six of them. To add to these problems described in the text entry,
there is another related action to delete a character, which is further worsen this scenario.

When a user tries to delete a letter, presses the delete key and the screen reader reads the character to the right of the cursor and not the character to the left, that'll really delete. This is even more worrying when a user wants to delete a letter in the middle of a word.

If the user wants to delete the letter "a" at the word "boat", to retreat with the cursor, TALKS, will give him the following feedback as he goes backwards "space" (if you have a space in front of the word) "o", "c", "t", "a". At this point, pressing the delete key will delete the letter "b" but the user will convinced that he is deleting the letter "a". The user is only aware of this when he moves the pointer to the end of the word and it is read again. Otherwise the error will remain undetected.

In this situation, TALKS, does not respect the consistency and adherence to standards, because it displays exactly the same behavior in opposite situations. Clearly does not prevent errors, on the contrary fosters its existence. If the user wishes to avoid errors have to resort to the principle of memory instead of the recommended recognition. After introducing the character, the screen reader, does not help the user to detect the error that he had just committed

Making phone calls

In the initial screen, when we begin to enter a phone number to make a call, TALKS, reads all the available information on the screen, mixing the information concerning the operator, date and time, with the number that we just mark it concatenating the description of the actions associated with the buttons in the lower corners of the screen. This situation only occurs for the first digit that is dialed but worrisome enough to be highlighted.

This situation violates the heuristic which states that the system state must always remain visible. The user remains in doubt if he is still the main screen or you already are dialing a number, this means that the user is tempted to dial the first number again because he did not know whether correctly or not. This encourages the violation of the heuristic error that recommends avoiding mistakes.

After the user dials the phone number you want to make a call, and press the green button to get it done, the TALKS fails to inform the user that the call is being held or what number is being dialed.

You just realize that the call is being held when you hear the ringer. Once again we have the present state of the system which causes the user can cause serious errors as to make a call without wanting to actually do it or calling the wrong number, and realize this fact only after someone answers on the other side of line.

Receiving Phone Calls

When you receive a phone call on your phone, you only get informed that the phone is receiving a call through the telephone ring, by vibration or by both together. If the user wants to know who the caller is becomes very complicated to know. There is no indication from the phone to say what actions they can perform when it receives a telephone call.

This situation violates the heuristic advises that one should give the user help and documentation. This type of situation also creates errors because you do not know how you can answer or reject the call without actually trying to do it.

There is, however, a key that identifies the caller but in reality it also serves another purpose: it is the silence button. This button when pressed identifies the name of who is making the call if it is part of the user's contact list, or displayed the phone number of contact. However this information is only available once. You may not hear the sender's data more than once and from the moment that makes you no longer have any feedback about the status of the call, the phone stops ringing and / or vibrate.

This is a clear breach of heuristic rules, not allowing the information on the state of the system when it stops playing and / or vibrating, the user does not control or exercise free will, because if you want to hear the sender's information for more than once cannot do it and if you want to know who is contacting and subsequently remove the sound on the phone cannot do it.

If you receive a call and not answer then the system reports that we have a missed call to see, however, does not provide any instruction on the options that I can do with that information. This is one aspect that is clearly recommended by the heuristic of availability of help and documentation.

Sending messages

When sending messages, the process is somewhat confusing due to the junction of the process of sending the actual message to the writing of it. In this respect there should be a minimalist design that would allow to have a form where you can only write the message content and then proceed to write another form for the recipient.

After selecting the recipient's post should be at a click away but it's not what happens. After entering the recipient, the user must go through a series of screens to send the message. In this respect we need a little more flexibility and above all efficiency.

Message Reading

There is no control over the reading of the message. When a user opens a message it is automatically read, only being possible to re-read it again in full. You cannot only read a certain phrase or word. At this point the user is not exercising free will. This lack of control over the reading of a message causes the user usually get an idea of the message content but do not know its contents accurately. In this case we are talking about the reversal of the principle of recall over recognition.

By entering the details of a message to access the content if I am again informed of the sender and the date of receipt. This information is already available in the summary of the message, the list of messages within the mailbox where incoming messages or information available in the event of receiving a new message. Thus, it makes no sense to repeat information we already have shuffling it with the message content.

In this situation the design should be minimal in order to provide us only the relevant information can be found the other information later.

3.4 Guidelines

After an analysis of all information previously collected, we compiled a set of design guidelines that we believe can contribute positively to the designing process of interface design for mobile
devices taking into account the special needs of blind and visually impaired users.

This design guidelines aim to reflect a series of conceptual principles of interface development, gathering information from practices such as Nielsen usability heuristics and information resulting from research conducted in this study. All information gathered was adapted to the special needs of this population in particular.

**System Status and Context**

The system status should always be known by the users. They should know exactly where they are and what the current state of the system is. This information should be revealed every time there’s a change in the system. The user should be aware of system information in a succinct and staggered way so that he can understand the differences between different types of information that are transmitted to him, regarding that relevant information should not compete with secondary information.

There should also be simple and practical way the user could query the state of the system and understand its context, every time he feels wishes.

Users should be aware of what actions they can perform in a given state (screen) of the system. These indications may be transmitted after the status change and must also be able to be consulted later when the user feels that particular need.

**Use of correct language and adapted to transmitting channel**

The language used throughout the system must be consistent, always using the same terminology and not using variations of terms for the same thing. The terms used throughout the system should be the closest possible to the real world and should avoid the use of technical terms so that the users may always feel identified with the language used.

Should be avoided all the specific features of written communication that do not have a direct match with oral communication. It should not be used, abbreviations and other insignificant graphic signs such as oral parentheses. It should be taken into account the use of sound icons in order to allow the user to easily identify the context in which he is.

**Full control over auditory feedback**

The user should have total control over the audio feedback of his device. It should allow users to change the speed and volume of auditory feedback in a quick and agile mode regardless of the context in which they are.

The user should be able to read the content of a particular screen according to his needs. It should be possible to switch between a sequential reading of text, sentence by sentence, word by word or letter by letter reading. The user should be able to easily choose which portion of text he wants to be reproduced and could move forward, backwards, up or down arbitrarily.

**Shortcuts and alternative navigation**

It should be possible for users to access any function without having navigation as a barrier. There should be “shortcuts” that allow quick access to any feature or an alternative method of navigation that provides easy access to all functionalities present in the system without having to go through traditional navigation over menus.

**Contextual Help**

The user should be provided with a mechanism that allows him a way to understand the range of functionalities that has at its disposal on the phone. Through this method he should get a list of features allowing him to obtain a representation of the information architecture of the system in order to get an overview of the system.

Each feature should also be accompanied with a brief description that the user could consult in order to get a general idea of its functioning. Besides the description of each particular action it should also allow, the user, to query the operating instructions in order to easily learn how to work with that particular functionality or to have a quick access to the operating mode in case he has forgotten it.

This feature has particular importance when filling out forms, where the user will normally only be aware of the type of content to populate fields after the completion of the information. It should be possible to have a brief description of which fields are present in each form and what type of information should introduced.

**Aid on text entry**

There should be a method that would help the user in text entry in particular at the entry of special characters. The method currently used is very inefficient and enhancer of a lot of mistakes. There should be a text entry method designed for blind and whose operation should be considered based on a hearing interface instead of a visual one regarding that they would not see it.

This method should also have a mechanism to minimize the occurrence of spelling errors that is one of major problems in text entry for blind people.

**4. Proposed Solutions**

The vast majority of encountered problems along the completion of our study were normally related to bad design principles of user interface design. The interface design should not contemplate something that was met to be transmitted through visual communication into something that will have to work through auditory communication.

This transformation is done without any adaptation or conversion for needs of specific communication channel. Note that many of these problems stem from its own operating system and applications from the phone and not the screen reader itself, but it is clear that the simple replacement of verbatim information provided by its auditory representation is insufficient and inadequate.

Based on the produced design guidelines we developed some prototypes that aim to address some of the problems encountered during our study.

These prototypes are aimed to propose a new approach for solving certain types of problems and to present alternatives to users so they can broaden their knowledge based on potential alternatives to the functioning of screen reader systems they use in their daily life.
These solutions were tested with users allowing us to compare the performance of these solutions with the performance of TALKS for the same set of tasks. We propose to the users to carry out certain tasks on both devices and measured their performance for comparison.

**Application Search**

One problem that we detected during the study was that people feel a lot of trouble finding an application that they do not regularly use and/or an application that they do not know exactly where is located within the hierarchical structure of the phone menus.

In this sense we came with an idea of creating a solution that should allow application search across the entire phone, so that the user should not feel the need to bear the organization of the menu in his mind whenever he wants to search for a specific application.

Another problem we have encountered during the study was that, in different phones the same features could have different names. This was some users could not easily find what they were looking for because they not always remembered the name the wanted feature. To mitigate this problem we decided to add a list of associated terms to each feature.

This worked with a text form where the user should insert the wanted text and then the system would search it for him and return him a list of possible matches. This list was identical to the other lists present in the system and user was allowed to navigate through them looking for his wanted application.

**Message Search**

Another problem we have encountered during our study was the existence of some difficulties in navigating lists of information especially in cases where these lists had a lot of records. This would typically happen as in a message box with lots of messages.

For a user to be able to find a particular message in his message box, he would have to go through all the messages one by one until he found the desired one.

The fact of that a user could know the recipient or part of the content of the wanted message was not accounted of. The user could not filter message list using this information.

In this context, emerged the idea of creating a search method for messages. This search method was very similar to the application search. However as a contextual search it allowed a different treatment for messages searching for different types of information. While at the search application ate the system takes advantage of the existence of associated terms to the application’s names, in this context the search should take into account the name of the recipient and/or the message content.

It was also thought to include the possibility to search by date, but was not implemented because it is information of a different nature involving other types of search form. In addition to search by date as some would want to search for a specific range of dates which would make the search even more complex. We have chosen not to implement this functionality because we wanted flexibility and simplicity but simultaneously an easy method of learning.

An alternative solution to the use of dates in the search field would be the sorting the search results by date. This ordering could be performed in ascending or descending way. This ordinance could also be valid for both dates and message recipient’s names. This ordinance was not implemented not add complexity to the process. However, it is a concept that could be exploited for the use of more experienced users.

**Cross system solutions**

Throughout the system we introduce the operation of two special keys, which always contained the same meaning, being adapted to the context. The “*” key allowed contextual search on each screen and the “#” enabled contextual help, providing information regarding the system status, context, a brief description of present features and its mode of operation.

We also altered the option list disposal throughout the system always keeping the information of the selected options along the user’s navigation. It was very helpful for the user to always have this information present because it allowed him to know current selected option, whether he had already got through the whole option list and was returning to beginning or if there were still many options to reach the end of a list at some given point.

We also decided to introduce a feature that would not only deal with spelling errors but mainly with phonetic errors. All text entry passed through this filter where the system cleaned all the special characters allowing it compare the search phrase with the system names that had also passed through this filtering. This change was made for research purposes and it allowed that a search for “joao” or “joio” returned the same results.

**5. Evaluation**

The aforementioned prototypes try to overcome some of the barriers identified in the preliminary studies performed with the target population. To be able to assess their values and if they really improve a blind person’s mobile interaction effectiveness, we had to go back to the users and evaluate the impact of the prototypes.

**5.1 Procedure**

The evaluation was performed in the Raquel and Martin Sain Foundation, a rehabilitation and formation center for the blind, with five blind people with ages comprehended between 47 and 57 years. All of them were mobile device experienced users and used a screen reader for at least 5 years.

In this evaluation, we selected two tasks, repeated several times by all users, using both the traditional and the proposed methods. This enabled us to look individually at each approach and to compare them. On the traditional end, the users performed the tasks with a Nokia N73 with Nuance Talks, their day-to-day screen reader. To evaluate our approach, we have deployed our prototypes in a HTC S310 with Windows Mobile 5.0. The screen reader user was Loquendo 6.0, male Portuguese voice.

Before performing the tasks the users were able to experiment the device they would be using for about 5 minutes. The order of the prototypes was randomized to counteract experience effects.

The first task to be performed consisted in finding a particular application on the mobile phone. We asked the users to find the applications “Despertador” (Alarm), “Calculadora” (Calculator), and “Caixa de Entrada de Mensagens” (SMS Inbox). They had to perform the three tasks in both devices, using the traditional navigation mechanisms (and/or shortcuts) with the traditional
screen reading software, and using the Application Search mechanism in the HTC (our prototype).

The second task consisted in finding a particular message in the mobile device SMS inbox. All the users were asked to find three messages on both devices, using the traditional navigation method and the Message Search mechanism. The messages asked were as follows (in Portuguese):

- Mensagem sobre acidente na Ponte 25 de Abril
- Mensagem de cancelamento de voos da Sata nos açores
- Mensagem com resultado do jogo Portugal-Dinamarca

In both tasks, the researchers collected the time to accomplish the task, number of steps (screens) and errors. In the end of each session, the participants were presented with a questionnaire composed by 6 sentences for them to rate using a 5-point Likert scale.

5.2 Results

In the case of application search we found that this was a method that limited the occurrence of errors, causes few changes of context, since it only takes three steps to accomplish a given search. It also revealed to be a good alternative to the traditional method for cases in which the user looks for an application of which he does not know exactly where to find it or who do not know exactly its name.

Throughout these tests we also allowed the user to take advantage of the phonetic filtering. The use of this tool was transparent to the user as would just use his search and would see the results. Some users searched for the word “relog” to find the “relógio” application and the system returned their intended results.

Contextual help was also been throughout the system during the conduct of these tasks but it was not used given the simplicity of the tasks the users had to perform.

In the case message search case we also obtained very satisfactory results. In this solution we created a scenario where the message box had a lot of messages in order to prove that our solution was particularly useful when a user has a mailbox with lots of information and wants to find a specific message of which he does exactly remembers.

The performance verified through this solution proved to be much better than the traditional solution. The times were always lower than those regarding the traditional solution and our worst case scenario, was always better than the best scenario for normal navigation.

Both the solutions proposed proved to be very effective in avoiding mistakes. Only one error was detected on application search through our solution. This error occurred due to the fact that we performed a standardization of terms used on the system, having removed all the abbreviations of expressions and we missed out the inclusion of the older terms on the associated terms in order to avoid this kind of situation.

6. Conclusions

Mobile technology developments witnessed in the last few years has dictated an enormous success of all mobile device types. In particular, mobile phones are now essential items in our daily lives. More than mere communication devices, they are now also productivity and leisure tools.

The functionalities these gadgets offer resort mostly to interfaces strongly based on visual information, making them inaccessible to users with visual impairments. To overcome this barrier, the blind population is presented with screen reading software. This text-to-speech based tool translates the visual information in its auditory counterpart thus enabling a blind person to receive information on the screen status. However, a question remained: is the simple translation of text into audio enough to provide an effective, efficient and enjoyable user experience?

This thesis aimed to understand the values and flaws of mobile device screen readers following a user-centered approach. To this end, we worked closely with blind users first to understand the barriers and limitations they still face and, in a second phase, to iteratively design and evaluate possible solutions to overcome the major flaws identified. Indeed, the studies performed with experienced screen readers’ blind users showed that a simple text to audio replacement is not enough and, in some cases, is inadequate for an effective user experience. The information is enriched with visual elements other than text and simple features like the position onscreen are likely to have a meaningful effect on the interaction, one that is lost in translation. Examples of this problem, that occur system-wide, are relations between keys and onscreen elements, menu hierarchies, sequential forms, and other particularly visual metaphors. Both the study performed with the target users as a usability evaluation performed with usability experts enabled us to identify and understand the major problems of this approach. Building on these results, we proposed a set of guidelines for the design of mobile interfaces for blind users.

To validate our guidelines, we have instantiated them in a set of small and focused prototypes. They included Application Search, Message Search, Contextual Help, System-wide actions, and a phonetic correction mechanism. An evaluation performed with blind users presented our ideas as valuable and confirmed the underlying guidelines. Indeed, if the users are placed in the center of the design process and the interface is designed based on their capacities and needs, the success achieved is relatively enormous.

The guidelines presented go way beyond the prototypes evaluated. We believe they can and should be followed across all applications supporting user effectiveness. Future work includes evaluating other applications and scenarios with prototypes built based on the presented guidelines.

7. References
