BloNo: A New Mobile Text-Entry Interface for the Visually Impaired

Paulo Lagoá, Pedro Santana, Tiago Guerreiro, Daniel Gonçalves, and Joaquim Jorge

INESC-ID/Technical Superior Institute, Lisbon Technical University
Av. Rovisco Pais, 1049-001 Lisboa
{plagoa,psantana,tjvg,djvg,jaj}@immi.inesc.pt

Abstract. We present a new mobile text-entry method that relies on alphabet navigation and dismisses memorizing, offering visually impaired individuals an easy writing mechanism. Current mobile text-entry interfaces are not suitable for blind users and special braille devices are too heavy, large and cumbersome to be used in a mobile context. With the enormous growth of mobile communications and applications it was urgent to offer visually impaired individuals the ability to operate this kind of devices. Evaluation studies were carried and validated the navigation method as a new mobile text-entry interface for the target population.

Keywords: Mobile, Visually Impaired, Text-Entry, Accessibility, SMS.

1 Introduction

Nowadays, the mobile phone is an essential tool for most people. Tasks like making calls, using the Short Message Service (SMS), managing contacts and using the agenda are executed on a daily basis. SMS, in particular, have had an enormous growth in the past years: in 2000, 17 billion SMS messages were sent; in 2001, the number was up to 250 billion and 500 billion SMS messages in 2004. More than one trillion text messages were sent in 2005 [1]. This growth is related with the low costs to the user and social factors – it doesn't disturb the receiver, it provides diffusion and it doesn't demand an immediate answer. The number of services accessible by SMS has also been growing in the past years, including alerts, quizzes and advertising. However, for visually impaired individuals this service is very difficult or impossible to use. This is also true for the majority of applications available on a cell phone. Nowadays, mobile device potentialities are diverse and cell phones are commonly used as clocks, notebooks, and agendas, among others. As for SMS, the visually impaired are also deprived from using these services. Traditional mobile devices are not equipped with keyboards adapted to those users' needs nor do they provide any kind of feedback for their actions. Special hardware devices, although making possible for blind users to use mobile services, are too heavy and cumbersome to be carriable and especially, be used on a mobile context. Currently, visually impaired individuals have limited access to mobile devices, particularly, text-entry-based applications.

Fig. 1. Regular Mobile Phone Keypad

We present an interface that relies on alphabet navigation and dismisses memorizing, offering visually impaired individuals an easy writing mechanism. Moreover, menu navigation was also implemented and several practical tasks were evaluated with the target population.

2   Background

In the last few years, we have witnessed an enormous growth on mobile communications. Furthermore, with the increase of mobile device's memory and processing capabilities, cell phones became real mobile computers, gathering functionalities from digital cameras, games to management and organization applications. Although these devices are very similar to personal desktop computers their input and output capabilities are much more limited. Researchers leaned over this problem over the past few years trying to overcome mobile devices limitations and increase user's performance as well as easing mobile usage.

2.1   Mobile Devices Text-Entry Methods

Text introduction in mobile phones is done through a 12 key keypad (Fig. 1). Most keypads follow the ITU E.161 standard, featuring number keys 0-9 and two additional keys (* and #). The letters are spread through keys 2-9 alphabetically ordered, forming groups of three or four letters. The space character is usually assigned to the 0 key, but it depends on the mobile phone model. Several methods have been developed to introduce text in these devices and can be divided in two main categories: multi-press systems and predictive systems.

Multi-Press systems
Multi-press systems categorize all systems that require one or more key presses to enter a character. These systems don't use word or phrase prediction to help the user. The most known and used multi-press system is multi-tap which requires one or more button presses to obtain a certain character. A disambiguation problem arises when

Fig. 2. MessageEase: English keyboard configuration

the user wants to enter two or more consecutive characters present in the same key with two possible solutions: key press timeout, an acceptance button.

MessagEase [2] allows text introduction with two key presses. The system is based on the letter frequencies of certain language. According to these frequencies, the letters are induced into the keyboard. For the English language, the distribution is as on Figure. 2. To introduce such letters, the user presses the corresponding key twice (for example, to enter 'a', key 1 is pressed twice). To enter less frequent letters, the usage is as follows: to enter 'u', for example, the user presses the 5 key and then the 2 key (the 5 key works as a directional key). Less-Tap technique uses language knowledge to rearrange the letters within each button. This rearrangement is obtained from the letter frequencies for certain language. This way, the most frequent letter (from each group associated with the key) requires only one key press [3]. Two-key Input requires exactly two key presses for each letter. The first press selects the group of letters (for example, '4' selects GHI). The second press selects the letter from the group (for example, '2' selects 'h'). This approach solves the disambiguation problem present in other methods.

For blind individuals, multi-press systems are rather difficult to use as the user must memorize the letters associated with every key and must always be very careful with the amount of presses to achieve a given letter. As feedback is inexistent, the user can't have a notion on the message evolution. Multi-press frequency-based approaches are even worst that multi-tap as the alphabet sequence is lost and therefore memorizing needs increase.

Predictive Systems

Predictive input technologies use language knowledge to predict what text the user is going to enter. Letter anticipators predict the next letter based on the prefix entered by the user. Word completers predict the suffix of the introduced word or the remainder or even the entire phrase, featuring suggestions to the user. Hybrid systems combine the letter anticipators and the word completers [4].

T9 (Text on 9 keys) [5] is the most widely used predictive system on mobile phones. The system compares the sequence of keystrokes to words in a dictionary, in order to determine the intended word. If the user ends the word introduction and the displayed word isn't the desired, a "Next" button should be pressed, and all the words

sharing the same key sequence are shown. If the word doesn't appear, the user should switch to multi-tap mode to entry the desired word. Using LetterWise, a text introduction method created by MacKenzie et al. [6], the user presses the key that features the desired letter. If the letter doesn't appear on the display screen, the user presses a "Next" key until the letter shows up. The prediction is done letter by letter, according to letter prefixes probabilities. The WordWise[7] method is similar to the T9, although in WordWise, when certain word isn't in the dictionary, LetterWise is the method in use to introduce the word (instead of Multi-tap). iTap, developed by Motorola, is very similar to T9: as the user writes, it compares the sequence of keystrokes to words in a dictionary, in order to determine the intended word. However, iTap also completes the words, featuring a list of possible suffixes of the word entered, and this is the main difference between the systems. eZiTap system, from Zi Corporation, combines multi-tap text introduction with word prediction, allowing an expert multi-tap user to improve performance without requiring any learning effort. As the user enters the desired word, a list of words sharing the same entered suffix is displayed. A "Next" key is used to navigate among the list.

Predictive systems with no visual feedback are prohibitive for blind people as the user isn't aware of the actual evolution of the system and the current message state. For example, using the T9 system, there isn't a relation between the key press and the letter appearing at the screen (it is affected by the predictive system frequencies) and it is not certain that, even when the correct keys are pressed, the word appears.

2.2 Visually Impaired Focused Approaches

Alternative devices were developed to overcome the difficulties arising from visual impairments. Typically, these products' goal is to serve as a Personal Digital Assistant (PDA) providing functionalities like Contact Management, Calculator, Notes, Clock or SMS (sending and receiving).

Fig. 3. Braillino

They normally allow connection to a desktop computer or a cell phone, acting as an interaction bridge between the visually impaired individual and the device.

Braillino, from Handy Tech Elektronik GmbH, is an electronic notebook adapted to the visually impaired that allows the connection to regular mobile devices through a docking station or bluetooth. It is also possible to connect it to a personal computer (Fig.3). It is based on a braille keyboard for input. It uses a software package called Taks&Braille that permits access to mobile phone funcionalities, offering braille output support and voice feedback (synthetic speech).

Alva Mobile Phone Organizer combines the funcionalities of a PDA and tri-band GSM cell phone allowing the users to make and receive calls, send and receive messages, take notes and manage contacts. Considering user interaction, it contains a twenty cells braille screen as well as speech synthesis and eight input braille keys.

There are several devices similar to Braillino or Alva Mobile Phone Organizer and they all share the same flaws: the large size and weight and the prohibitive costs when compared with regular mobile devices. Although it is true that blind users can use braille-Based devices to accomplish their goals, it is also true that these devices are too heavy and large to be carriable and used while on-the-move. Also, considering a usual scenario where a blind user handles a cane with one hand, it is impossible to operate this kind of devices (Fig. 4).

In a totally different scope are the screen readers, solutions that can be used in a regular mobile device, giving the users feedback on screen evolution and replacing visual feedback. Nuance Talks is an example of this type of assistive technology for the visually impaired. Although screen readers make possible for a blind user to use a mobile device, they still require for the user to memorize letter's placement. Considering performance issues, predictive systems are still very difficult to use.

Fig. 4. Alva Mobile Phone Organizer

3   Task Analysis

To identify the users' needs and capabilities as well as their common use of mobile phones we interviewed 8 visually impaired individuals. The interviews were conducted at Raquel and Martin Sain Foundation, a learning center for blind people, to six male and 2 female individuals, with ages between 21 and 64 years.

Considering mobile phone use, it was curious to verify that all users possess a cell phone and use it once or more a day to make calls. Only two of the users reported

having problems to dial a number due to the lack of feedback and small keypad size. Normally, dialing isn't problematic and therefore this task is not the main concern of our work.

Half of the users (4) use the phone's contact list and they all share the same technique: memorize the contact's position in the list. Those who don't use the contact list justify it with lack of feedback from their phones. Although this result is somewhat surprising (we expected the contact list not to be used at all), the method is rather limited and lacks scalability as the users will only remember some important contact list positions.

Half of the users stated to send text messages (SMS) from which half stated to do it on a daily basis. All users that send SMS use Talks Screen Reader to get voice feedback but only one uses a predictive system (T9). The users generally agree that menu navegation is difficult and cumbersome and that the feedback is limited taking them to recurrently make mistakes.

4   Our Approach

Regular mobile device text-entry methods are suitable for visually capable individuals and seek to improve user's performance. Hence, it is possible for someone, with no experience, who doesn't remember the location of a letter, to easily look and recognize the key where that letter is. Those methods imply that the fingers *dance* through the keyboard, choosing letters and special characters, among ten or more keys. Once again, we easily overcome this issue appealing to vision. A blind user can't do that. The mark present at key '5' gives blind users the notion of the keypad layout but not feedback on the selected letter and, although users can make an effort to memorize a letter's placement, feedback is essential. Even SMS experts need to occasionally look at the words being written to ensure message correction. Moreover, expertise is acquired by using the method and receiving the feedback. Only after an extensive and successful use of the writing mechanisms the users can get used to them and, in some cases, no longer need constant visual feedback. Screen readers (such as Talks) overcome some of the issues as they offer the user feedback on the screen progress. However, keypad feedback is still inexistent which often leads to mistakes and sometimes giving up. Although users make sense on the message progress, they still have to know where to press to get the desired letter/action.

We can only offer visually impaired individuals mobile device accessibility if those devices can be easily available and usable. Therefore, based on user needs, capabilities, and available devices we decided that our method should be compatible with regular mobile phones and, therefore, with the regular 12 key keyboard layout requiring no-extra hardware (i.e. expensive and heavy braille keyboards).

When looking at the available solutions it is also clear that text-entry methods are not suitable for visually impaired individuals, even when using a screen reader, as memorization is still required. Therefore we designed our system not to require memorization or difficult adaptation stages. It is based on a new navigation-based text-entry method and speech synthesis feedback. Besides text-Entry based applications, BloNo provides the necessary feedback to navigate through menus and access applications.
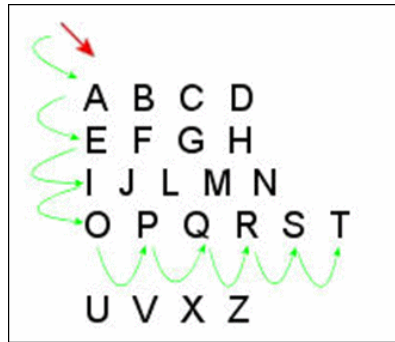
914   P. Lagoá et al.

Fig. 5. Navigation method using two directions

4.1   Vowel Navigation Method

To accomplish our goals, a new text-entry system was developed. The alphabet was divided into five parts, each starting with a different vowel, as these are easy to remember unlike other arbitrarily chosen letters. Those parts were then used to create a character matrix (Fig. 5). We use the keyboard keys as a joystick to navigate that matrix. Keys '2' and '8' allow the user to navigate vertically through the vowels, while keys '4' and '6' will allow navigation through the rest of the alphabet horizontally (Fig. 5). Both the horizontal and the vertical navigations are cyclical, which means that the user can go, for instance, from the letter 'z' to the letter 'a', and from the vowel 'u' to 'a' (Fig. 6).

Using all directions will allow the user to reduce the number of key presses needed to reach to a certain letter. Comparing the approaches followed in Fig.5 and Fig. 6 we can state that, by using all the directions, the user needs only two strokes to reach the letter 't', instead of nine strokes (using only two directions). Also, the user receives voice feedback before accepting any letter (automatically after a configurable timeout), therefore reducing entry mistakes. The method requires little memorizing and even with low alphabet mental mapping the users can still, in a worst case scenario, navigate forward until they get to the desired letter (there are no wrong buttons, just shorter paths). This is a major issue for the visually impaired as they can totally rely on audio feedback before accepting any letter, increasing the text-entry task success and the motivation to improve the writing skills.

4.2   Automatic Word Finishing – AWF

To optimize user performance while introducing text, we developed a system that suggests (on user demand, by pressing the '9' key) a list of possible words that start with the prefix the user has entered. The user accepts the word by pressing the '5' key once. Punctuation characters can be selected by continuously pressing the '5' key.
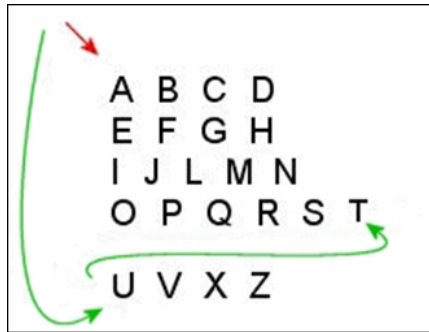
Fig. 6. Navigation method using four directions

4.3 Task-Aware Dictionaries

To enhance AWF performance we developed task-aware dictionaries. These are focused on the specific task being performed by the user and their goal is to reduce the spectrum of words suggested by the AWF. As an example, when entering a name on the contact list, the only words the user will need are names and surnames. In the same way, when entering an address, the user will only need street names.

5   Evaluation

Tests were made with five totally blind users assessing the new text entry interface as well as the techniques used to improve it, in a set of tasks including message writing and contact manipulation (Fig. 7). The tests were carried at Raquel and Martin Sain Foundation in a quiet room with no disturbances, with each student individually. All the five users presented 100% visual deficiency.

The test session start with a brief system presentation, followed by an ambientation phase. To support the evaluation method presentation, we printed a set of instructions and the description of the tasks in a braille manual. The evaluation consisted in six different tasks.

The first task consisted in writing, with no acceleration methods (i.e., AWK), the portuguese sentence – *Perdi o autocarro e não vou chegar a horas*. The text entry interface was well understood by all users with a mean of 2 words/minute ratio and a 1.4 error rate, after a brief 5 minutes training session. In the second task the users were asked to write the same sentence, but this time being able to use the AWF predicting system. The AWF mechanism reduced the number of keystrokes per character (KSPC) in 13.9%, the time to accomplish the tasks in 9.9% and the error rate to 1.

Adding a contact was tested in 3 different conditions: without AWF, with AWF and with AWF and task-aware dictionaries (names only). AWF mechanism improved the total time in 32.4% and task-aware dictionaries improved it by 14.5%, with a total time reduction of 42.2%.

Finally, the users were asked to send an SMS to a defined name in the contact list. This task gathered several steps: navigating menus, writing the text, choosing the

contact and sending the message. Both the message and the contact in the contact list were pre-defined by us: the message was *sim* and the name was *Paulo Lagoá*. All the users successfully completed the task with a mean time of 104 seconds.

In general, all users were satisfied with the system and showed great improvements in their performance across the evaluation process (increased the number of words/minute and reduced KSPC, from the first to the final task). It is also important to notice that, although all the users had a cell phone, only one was capable of writing and sending messages. With the presented system all the users completed the tasks with success navigating through menus, writing and sending SMS and managing contacts without apparent difficulties and with good learning rates.

Fig. 7. Blind user testing the system

6   Conclusions and Future Work

While today mobile phones are a common tool used in everyday life, they strongly rely on visual feedback for their correct operation. As such, visually impaired users have trouble using them, contributing to frustration and social exclusion. Existing alternatives are not practical or depend on memorization, often leading to mistakes. We developed a new text-entry interface that requires no memorizing and no extra hardware making possible for any visually impaired individual to use a regular mobile device, accomplishing even the most difficult tasks like writing a message or managing contacts. We evaluated the interface with the target users and validated the approach, showing that users input words with an acceptable rate improving their performance quickly with almost no previous experience. As future work, we will study other text-entry interfaces as well as shortcut mechanisms that ease menu navigation, like accelerometers, RFID tags or voice recognition.

References

1. GSMA's Hubbing Programme to Streamline Global Delivery of SMS Messages, GSM Association Press Release (2006)
2. Nesbat, S.B.: A System for Fast, Full-Text Entry for Small Electronic Devices. In: Proceeding of the Fifth International Conference on Multimodal Interfaces, ICMI 2003, Vancouver (November 5-7, 2003)
3. Pavlovych, A., Stuerzlinger, W.: Less-Tap: A Fast and Easy-to-learn Text Input Technique for Phones. Graphics Interface (2003)
4. Smith, J.: Thumbscript: Designing a general solution to the problem of text input in small devices - WHITE PAPER, Thumbscript Development (2000) www.thumbscript.com
5. Darragh, J.J., Witten, I.H., James, M.L.: The Reactive Keyboard: A Predictive Typing Aid, IEEE Computer, pp. 41–49 (1990)
6. Grover, D.L., King, M.T., Kuschler, C.A. Patent No. US5818437, Reduced keyboard disambiguating computer, Tegic Communications, Inc., Seattle, WA (1998) http://www.t9.com/
7. MacKenzie, I.S., Kober, H., Smith, D., Jones, T., Skepner, E.: LetterWise: Prefix-based disambiguation for mobile text input. In: Proc. UIST 2001, pp. 111–120. ACM, New York (2001)
8. MacKenzie, I.S., Kober, H., Gutowitz, H., Jones, T., Skepner, E.: Linguistically Optimised Text Entry on a Mobile Phone (2001)