



Designing Multimodal Mobile Interaction for a Text Messaging Application for Visually Impaired Users

Carlos Duarte^{1*}, Simon Desart², David Costa¹ and Bruno Dumas²

¹ LASIGE, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, ² PReCISE Research Centre, University of Namur, Namur, Belgium

While mobile devices have experienced important accessibility advances in the past years, people with visual impairments still face important barriers, especially in specific contexts when both their hands are not free to hold the mobile device, like when walking outside. By resorting to a multimodal combination of body based gestures and voice, we aim to achieve full hands and vision free interaction with mobile devices. In this article, we describe this vision and present the design of a prototype, inspired by that vision, of a text messaging application. The article also presents a user study where the suitability of the proposed approach was assessed, and a performance comparison between our prototype and existing SMS applications was conducted. Study participants received positively the prototype, which also supported better performance in tasks that involved text editing.

Keywords: non-visual interaction, body-based interaction, gestures, voice, multimodal interaction, accessibility

OPEN ACCESS

Edited by:

Francesco Ferrise,
Politecnico di Milano, Italy

Reviewed by:

Farhan Mohamed,
Universiti Teknologi Malaysia,
Malaysia
Gualtiero Volpe,
Università di Genova, Italy

*Correspondence:

Carlos Duarte
caduarte@fc.ul.pt

Specialty section:

This article was submitted to
Human-Media Interaction,
a section of the journal
Frontiers in ICT

Received: 05 September 2017

Accepted: 20 November 2017

Published: 06 December 2017

Citation:

Duarte C, Desart S, Costa D and
Dumas B (2017) Designing
Multimodal Mobile Interaction for a
Text Messaging Application for
Visually Impaired Users.
Front. ICT 4:26.
doi: 10.3389/fict.2017.00026

1. INTRODUCTION

Today's mobile devices are flat touchscreens where you interact through touch and gestures. While this seems simple and easy for the average sighted person, persons with visual impairments have increased difficulties to do standard operations with their mobile phones. However, valid accessibility solutions have been developed during the last decade. Current solutions consist in screen exploration techniques for navigation and text input. While they enable easier interaction for the blind, those techniques can still turn out to be very cumbersome and slow.

Novel technologies and research projects in input sensing have opened up new ways to interact with computers and other devices. Body interaction explores taps on skin, mid-air gestures or natural movements in order to trigger actions in interactive devices. Moreover, performing actions that involve our own body does not require any form of visual attention. These techniques, combined with our proprioceptive capabilities, are interesting alternatives for interacting with, for example, mobile devices.

Still, the range of operations that are done today in mobile devices introduces requirements that will not be dealt with body based interaction techniques. For example, text input is a common operation in mobile devices. For text input, other alternatives are, probably, more efficient than body based ones. To be able to address the multiple requirements, a multimodal solution combining different input modalities is more efficient.

This work addresses accessibility issues affecting persons with visual impairments when interacting with their smartphones. We introduce a new multimodal approach, based on on-body interaction and speech, aiming to improve the accessibility of smartphone's user interfaces. To demonstrate its applicability, we present the design of a prototype of a text messaging application

following the proposed approach. We conducted a user study with 9 participants with visual impairments, where our prototype was compared with the current way text messages are entered by the study's participants. Not only did we receive positive feedback from the study's participants, we found that with our prototype they were able to perform different operations with a text messaging application quicker, even though they were not experienced with the proposed approach.

In the following section, we present previous research on mobile accessibility, with a particular focus on text entry, and on body based interaction. Section 3 presents our approach to the design of accessible interaction in mobile devices. The user study is presented and its findings discussed in the next section. Finally, section 5 concludes the article.

2. RELATED WORK

In a 2014 study, Ye et al. (2014) stated that 85% of visually impaired interviewed, own a smartphone. This is an interesting fact since touchscreens offer more accessibility barriers than non-smartphones. However, not only smartphones offer features that are not available in other phones, but visually impaired people also want to be trendy, even if they have to face interaction difficulties. Hence the importance of offering accessibility features and alternatives that accompany novel technologies.

2.1. Mobile Accessibility

A lot of research has been done in the last years to ease the interaction of visually impaired users with their smartphones, achieving better results than traditional methods. For instance, (Kane et al., 2008) presented an alternative solution to the standard features offered by mobile device's developers. The authors implemented Slide Rule, an interaction technique to improve the accessibility of multitouch screens when used by visual impaired users based on four main gestures.

Guerreiro et al. (2008a,b) developed NavTouch, a technique based on directional strokes to navigate through the alphabet in a way that decreased the cognitive load of memorizing the position of where the characters are located on the screen.

Bonner et al. (2010) presented a novel solution for touchscreen's accessibility issues when used by visually impaired people. No-look Notes is an eyes-free text entry system that uses multitouch input and auditory feedback. It offers a two-step access to the 26 characters of the alphabet with a small number of simple gestures and an 8-segment pie UI that removes the precise targeting required, for example, by QWERTY keyboards layouts.

Southern et al. (2012) proposed BrailleTouch, a six-key chord braille keyboard for touchscreens. This technique was designed to use the smartphone's screen faced away from the user and held by both hands. The screen was divided in 6 parts in the same way that BrailleType did and other control keys, such as space, backspace, or enter, were implemented through flick gestures.

In addition to text input *via* Braille Cells proposed in BrailleTouch, Holibraille (Nicolau et al., 2015) offers multitouch capabilities to perform text editing (e.g., navigation through words, text selection). Two finger gestures combined with non-dominant hand gestures allow the user to edit and navigate in the

text. The same authors also developed a novel correction system for these types of input (Nicolau et al., 2014).

2.2. Body Interaction

Recent technological advances and research projects about wearable technology and sensing devices have opened up new possibilities for using our body as an interaction platform. The always present skin, when combined with our proprioceptive capabilities, plus the control we can exert over our limbs are a sound alternative for interacting with personal devices, such as smartphones. This is an area with a great potential for improving the accessibility of interactive systems for multiple population groups, which is demonstrated by some recent works on this topic.

Virtual Shelves (Li et al., 2009) uses the proprioceptive capabilities to support eyes-free interaction by assigning spatial regions centered around the user's body to applications shortcuts. To measure the different angles (θ and ϕ planes, i.e., up-down movement and left-right movement) between the body and the arm holding the cellphone, the system uses an accelerometer and a gyroscope.

In Li et al. (2010), an extension of Virtual Shelves (Li et al., 2009) is presented. This interaction technique leverages proprioception to access application shortcuts. However, it was not meant for visually impaired people to use. Therefore, the authors performed a study to measure the directional accuracy of visual impaired persons and adapt the system for them. Also, the original work was not intended for mobile devices while the scope of this work was to enable mobile interaction (e.g., launch applications) when walking on the street.

Oakley and O'Modhrain (2005) developed a motion based vibrotactile interface for mobile devices. The authors use 3-axial acceleration sensing to directly control list positions, instead of using this sensor to control the rate of scrolling or directional movement. The goal is to link or associate certain specific orientations to specific list items.

Dementyev and Paradiso (2014) developed WristFlex, a system that makes use of pressure sensors to detect pulse gestures and is capable of distinguishing subtle finger pinch gestures with accuracy over 80%.

Matthies et al. (2015) describe a novel interaction technique, Botential, that makes use of sEMG and capacitive sensing, placed on a wristband, to detect the different body parts that it is in contact with.

Other recent approaches, such as Skinput (Harrison et al., 2010) and PUB (Lin et al., 2011) use the skin as a means of interaction. These projects use bio-acoustics and ultrasonic signals, respectively, to locate finger taps on body.

The latest efforts by Makino et al. (2013) go beyond just touching and tapping actions by using a new technique with photo reflective sensors, enabling pinching, pulling, twisting and caressing.

Dezfuli et al. (2012) present a palm-based imaginary interface to control the TV. While the average effectiveness of the palm touches was around 96.8%, to track the hand movements and gestures a Kinect was used, making this not suitable for mobile environments.

2.3. Social Acceptability

These novel interaction techniques were also subject of several social acceptability studies. The willingness to perform these gestures will largely be dictated by how appropriate those actions look and feel when performed in public. Profita et al. (2013) and Rico and Brewster (2009, 2010) findings show that there is a significant relation between audience or location and the willingness to perform certain gestures. This factor must be taken into account when building such systems. Findings also show that users are fonder of subtle movements, movements that are similar to what already exists in current technology (e.g., gestures similar to the ones in touch devices), and movements similar to the ones used in our everyday lives and enjoyable movements. On the other hand, participants in these studies stated that uncommon, large or noticeable movements would look weird in public settings.

In a work previous to these social acceptability studies, Costanza et al. (2007) already used surface electromyography (sEMG) signals to capture subtle gestures, ones that are difficult for observers to perceive when someone is performing them.

Williamson et al. (2011) developed a multimodal application with gesture recognition and audio feedback (speech and sounds). Despite being based on subtle gestures, the authors observed that some gestures were considered unacceptable in certain settings by the participants in the study. Consequently, the study participants developed new ways of performing the same gesture. To address this issue, such systems must be flexible and develop correction mechanisms in the recognition process. Additionally, the authors found that the willingness to perform gestures in a public setting does not depend only on the type of audience but also if they are a sustained spectator (e.g., other passenger on a bus) or a transitory spectator (e.g., a person walking in the street).

3. DESIGNING MOBILE ACCESSIBLE INTERACTION

Although there is an increasing number of projects emerging in the body interaction field, its potential for improving mobile accessibility is being underestimated so far. We propose to explore

body based interaction techniques, combined with speech in a multimodal system (Dumas et al., 2013), toward this goal.

When exploring solutions designed for the visually impaired population, one must make sure that interaction is reliable, robust and adaptive to the user and the different contexts of use. In the proposed design space, it is necessary to understand how the human body can be used as a mean of interaction considering, in particular, the characteristics of this user group. Albeit vision is considered the primary spatial reference, there is no consensus (Jones, 1975; Thinus-Blanc and Gaunet, 1997) if the level of spatial awareness within the body range is affected negatively or positively by vision impairments. The ability of accurately tapping in very specific parts of the body (e.g., dominant index finger tapping on the distal phalanx of the ring finger in the other hand) can also be affected by how long the person is blind or by the training received to develop proprioceptive skills. Moreover, as people with visual impairments do not have full notion of their surroundings, some gestures may be considered uncomfortable to perform in public settings (e.g., pointing may hit another person).

Thus, the overall design process will require several steps (Figure 1). First, we have to study how body based techniques can improve the different activities performed in mobile devices. For that, it is necessary to understand their limitations. UI (user interface) navigation and text input on mobile devices are examples of activities we will focus on. We will investigate if a body-based interaction technique is appropriate for single or multiple activities. For instance, is a technique suitable only for application navigation (e.g., sliding a finger over the forearm to move through interactive elements) and target selection (e.g., performing a mid-air gesture to select an item), or can it be also used for content input?

Second, we will explore where the novel interaction techniques can replace existing techniques, where they cannot, and where they can complement those. Thus it is important to understand what existing traditional input modalities can be replaced or complemented by this novel interaction technique, such that the device retains the same level of functionality.

Finally, we want to consider how body interaction can be used as an intelligent interaction mechanism. Given that perceptual, motor and cognitive capabilities vary from person to person, and

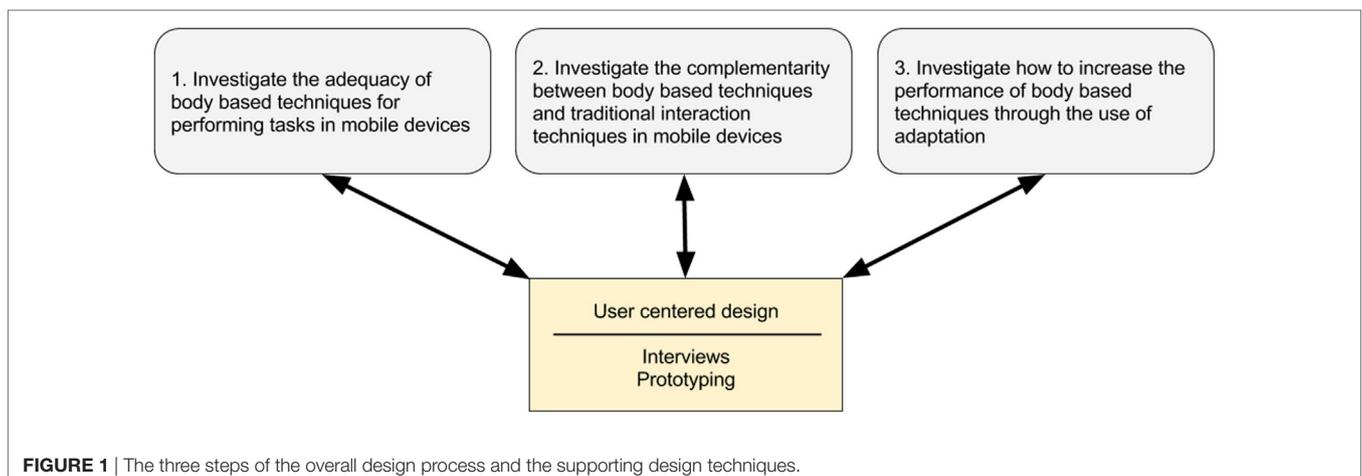


FIGURE 1 | The three steps of the overall design process and the supporting design techniques.

in the case of visually impaired persons many factors can affect those capabilities (e.g., born blind, disease that affected vision, etc.), we want to study adaptive mechanisms that are able to map the user's capabilities to the potentially complex location of skin based interactive points and gestures.

To explore these interaction techniques, a set of prototypes will be iteratively designed (Hinman, 2012), in collaboration with end user representatives, i.e., it will be developed based on a user centered design approach (Norman and Draper, 1986; Svanaes and Seland, 2004; Kangas and Kinnunen, 2005). Depending on the requirements, these prototypes can target interaction or technological validation. In the latter group, prototypes can range from acoustic sensors capable of capturing and defining a large set of positions on the hand or forearm; light sensitive sensors that capture skin displacement when sliding for instance the finger on the forearm; to sensors sensitive to muscle movements which can capture gestures or movements. For the former group, interactive techniques will be tested with multiple fidelity prototypes, sometimes resorting to Wizard of Oz techniques, whenever the technology is not at an adequate readiness level (Molin, 2004).

Our hypothesis is that by combining body based gestural and voice input we are able to increase the performance of users with visual impairments when interacting with mobile devices. One possible implementation is assigning gestures to specific application commands or shortcuts to certain applications. Through this, we aim to reduce the time a user spends navigating the smartphone UI until the desired application or contextual command is reached. UI navigation can represent a significant percentage of the interaction time for the visually impaired, thus we focus on a feature with a large potential to be useful. While it is true that the proposed solution may increase the users' cognitive load, we intend to take advantage of being able to memorize a set of gestures without overwhelming them. With the current accessibility solutions, users also end up memorizing the number of navigation steps required to reach the desired option to speed the navigation process.

In what concerns text input, we propose to use voice recognition in substitution of the current traditional methods (e.g., QWERTY keyboard display). The goal is to rely on the increasingly robust and accurate voice recognition technology to decrease the time and effort that visually impaired people have to put into writing a text. Nevertheless, in the future we will consider body based solutions for text input, like those based on, for instance, the Lorm alphabet (Gollner et al., 2012; Caporusso et al., 2014). By offering multiple modalities we will address the need for text input in contexts where speech recognition could not be socially or technically feasible.

Combining these two forms of input, body-based and voice, we are able to offer an eyes and hands-free input interaction with the mobile device, making it suitable to interact in multiple contexts. In the next section we describe the design of a prototype of an application for text messaging that takes advantage of these combined input modalities.

3.1. Case Study: SMS Application

Sending text messages is one of the most basic and most used features in every mobile device. This section presents a first

prototype of an SMS application based on non-traditional interaction methods: body-based gestures and voice recognition.

Inspired by the interaction approach presented above, a multimodal SMS application for the Android platform was developed, offering the standard SMS service features:

1. The user can enter a message.
2. The user has the possibility to review and edit the message if necessary.
3. The user can find and select the message addressee.
4. The user can send the entered message to the selected addressee.

Following our approach, by not relying on a visual interface, the user should be able to access every functionality without the need to navigate through a list of commands. To meet this goal, different actions are available to the user, with each action triggered by a specific gesture. This approach has the added benefit of allowing users to command the mobile application without having to hold the smartphone on their hands.

Existing gesture recognizers based on sEMG sensors, like the one presented in Costa and Duarte (2015), could prove useful for this scenario. However, they are not capable yet to reliably address a per-user on-the-fly customizable set of gestures. Therefore, for our prototype we decided to use the Wizard-of-Oz technique in the user studies. Gestures made by study's participants are interpreted by a human that inputs the command in the application. Voice commands are recognized using the speech recognizer available in the Android API.

The possible actions are triggered by gestures, with commands being a combination of gestures and voice input:

1. Add a message: after recognizing the "add a message" gesture, the speech recognizer is turned on and the user is notified by a sound. The user then dictates the message.
2. Review the message: after recognizing the "review message" gesture, the text to speech engine is turned on and the message is read word by word.
3. Edit a specific word in the message: during the message review, the user can stop the speech engine (by performing a gesture) whenever she wants to edit a word. Whenever that gesture is recognized, the application speaks the chosen word and turns on the speech recognizer. The user then dictates the new word. After that, the system informs the user that she will have to choose between five proposals. The application reads the proposals and the user stops it (by making a gesture) whenever she wants to select the desired proposal. The application speaks the selected proposal and the change is carried out in the message.
4. Enter a token to search for a contact: after recognizing the "enter contact" gesture, the speech recognizer is turned on and the user says the name of the contact or a word that begins with the same letter of the contact's name. Then the application speaks how many contacts match this search. The user can add other tokens to refine the search, or proceed to hear the names returned by the search.

5. List the contact that match the current search: the application reads all the contacts that match the search token previously entered and the user can stop the application to select the desired contact. The application orally notifies the user of the chosen contact.
6. Send message: after recognizing the gesture the message is sent and the user is notified.

The design decision of presenting five proposals during text editing is to counter possible misrecognitions of the speech recognizer. The same motivation was behind the decision to consider only the first letter when entering a search token. Special attention was given to the message editing process to make it more accessible. Indeed, Azenkot and Lee's survey showed that people with visual impairment spend 80% of their time editing when using speech recognition which can be frustrating (Azenkot and Lee, 2013).

4. USER STUDY

The conducted user study had two main goals. The first was to understand and study the behavior of visually impaired persons when interacting with the proposed approach, i.e., observe what type of gestures would persons with visual impairments perform for certain tasks and if they would feel comfortable to do so in different contexts. The second goal of the study was to compare, performance wise, the developed prototype with the applications that participants currently use to send text messages.

4.1. Methodology

To address the study's goals, we collected both quantitative and qualitative data through questionnaires, interviews and task observation.

This study was carried out in accordance with the recommendations of Regulamento da Comissão de Ética para a Recolha e Protecção de Dados de Ciências with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Comissão de Ética para a Recolha e Protecção de Dados de Ciências.

The first step was to complete a short characterization questionnaire divided in two sections: personal information and mobile phone usage habits. A final question was asked in order to perceive the acceptability of doing midair gestures in a public setting.

In the next step, the performance of the prototype and the standard text message application used by the participants were compared. Participants were divided into two groups following a counterbalancing design. One group begun by using their standard messaging application. The other group begun by using the prototype.

Before using the prototype, all participants were asked to define gestures for the available commands:

- Start application;
- write message;
- review message;
- edit specific word during the reading;
- confirm a proposal;
- enter a search token to find the contacts;
- list the contents that match the search token;
- confirm one of the contacts while reviewing the list;
- send the message.

Participants were told they had no constraints about the gestures they could select, i.e., they could do mid-air gestures, gestures on their body, or whatever they found appropriate. The test moderator exemplified gestures of the different categories before asking the participants to define gestures for the requested commands. While the participants were defining the gestures, no guidance was provided besides explaining what the command was supposed to perform. All participants defined and performed all gestures while seated on a chair.

Being a novel form of interaction, before completing the tasks in the prototype participants were given a training task to get acquainted with the flow of the application and recall all the gestures they defined moments before.

Participants were then asked to perform the same set of tasks in both the prototype and their usual messaging application. To compare their performance, a quantitative evaluation was designed based on time to perform different tasks on both applications. We provided a smartphone with our prototype for one of the conditions. For the other condition, participants used their own smartphones and their preferred messaging application. We requested access to the participants' smartphone before starting the tasks to collect data about their list of contacts in order to make sure that all tasks had similar difficulty. We did not store this data. It was only used for preparing the tasks to be executed. Three different tasks with increasing difficulty levels were defined:

1. *Easy* task: the participant is asked to enter a text message (the message was: "Hello, can you call me back as soon as possible"). The participant is requested to review the message entered. No edition of the message is requested. The addressee is easily findable (maximum 3 results in the prototype condition and a contact figuring at the beginning or the end of the contact list for the standard application in the participants' smartphone).
2. *Normal* task: the participant is asked to enter a text message (the message was: "I am waiting for you at the entrance of the store"). The participant is requested to review the message entered. The edition of 1 word of the message is asked (replace "waiting" with "looking"). The addressee is moderately easy to find (maximum 5 results in the prototype condition and a contact figuring in the first or last quarter of the contact list for the standard application in the participants' smartphone).
3. *Hard* task: the participant is asked to enter a text message (the message was: "Can you show me where Bruno is?"). The participant is requested to review the message entered. The edition of 2 words of the message is asked (replace "show" with "tell" and "is" with "went"). The addressee is more difficult to find (more than 5 results in the prototype condition and a contact figuring in the middle the contact list with a name beginning with a frequent letter for the standard application in the participants' smartphone).

To ensure that the text entering and the contact finding tasks done on the participant's smartphone and the prototype had similar levels of difficulty, we (1) asked them to enter the same message in both devices; (2) artificially controlled the position of the contact in the list of contacts in the prototype and, as aforementioned, previously reviewed the participant's contact list and selected one contact in a similar position.

The first application used in the trial and the order of the three tasks were both counterbalanced. Each task was timed by an observer, registering the time each task and sub task (e.g., write message, review, etc.) took to complete and the errors made.

An evaluation concerning the delay between the reading of two words during message review or contact listing was also performed. Participants were asked to choose between 3 possible delays (1, 1.5, and 2 s) the one preferred. The order of presentation of the delays was randomized.

The final step consisted in filling two System Usability Scale (SUS) (Brooke, 1996) based satisfaction questionnaires: one for the standard application and one for the prototype.

Given the time needed to complete all tasks, approximately 1 h and 45 min as found in pretrials, it was decided to split the tasks by application into two sessions of 1 h each in two different days.

4.2. Results

In this section, we present the results found in the user study.

4.2.1. Participants

The group of participants was composed by 9 persons (3 females). Their ages ranged from 24 to 56 years old ($M = 32.78$, $SD = 9.62$). Eight of the participants lost their vision between 14 and 31 years old and one has partial vision since birth.

Regarding mobile phone consumption habits, 6 participants stated to own a smartphone for less than 2 years. Five persons own an Android phone and four an iPhone. When asked to assess their expertise in the use of the built in assistive technology (Talkback or VoiceOver), 3 considered themselves experts, 3 proficient, 2 average, and one a beginner. In the specific case of using messaging application, 2 considered themselves to be experts, 4 proficient, 2 average users, and 1 considered himself as a beginner. Nearly all participants send and receive messages everyday (8 out of 9).

When asked if when walking outside, they stop to answer an incoming call or an SMS or they ignore it, 6 participants stated to answer while walking. Regarding the acceptability of performing midair gestures in any context, public or not, 4 participants stated they would do it. The other 5 participants stated that it would depend on the gestures and location.

4.2.2. Gestures Definition

Participants were asked to define a gesture for each one of the 9 used commands. In total, 81 gestures were registered. Despite the high number of gestures, few were repeated among participants. Nevertheless, we could observe a trend for 4 commands:

1. **Adding text.** Movement similar to writing with the fingers on leg (**Figure 2A**).
2. **Select a proposal.** Index finger touch (**Figure 2B**).



FIGURE 2 | Most common gestures for: **(A)** writing a message, **(B)** confirming a selection, and **(C)** sending a message.

3. **Selecting a contact during the list review.** Index finger touch (**Figure 2B**).
4. **Send message.** Swipe Right (**Figure 2C**).

Other interesting findings were observed in the gesture definition phase. For instance, all participants concluded that using the same gesture for confirmation and selection purposes would be the best solution, decreasing the number of gestures to memorize. Regarding the gesture's visibility, 3 out of 9 participants defined subtle gestures, with most made beside the participant's pockets or on an available surface.

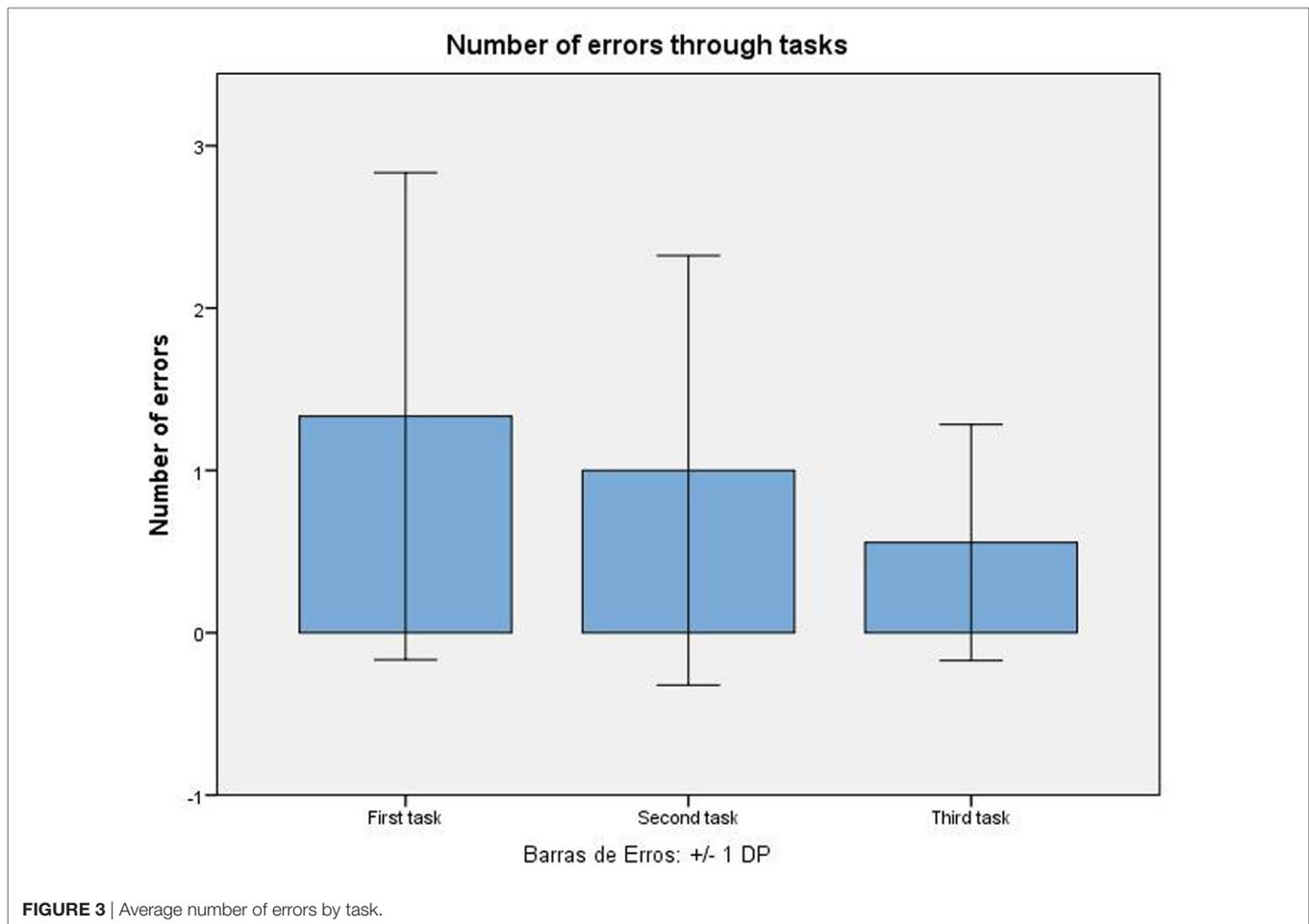
Additionally, we observed that several gestures meant different actions depending of the participant:

1. **Thumbs up.** One participant considered this the gesture to start writing the message while another participant chose this gesture to send the SMS.
2. **Swipe up with two fingers.** This gesture was defined to review the message but also to listen to the contacts list.
3. **Drawing the letter C.** One participant chose this gesture to enter a search token to find a contact while other participant did the same gesture to listen to the contacts list that matched with the token.
4. **Index finger double touch.** This gesture was defined to start writing the message and also to listen to the list of contacts that matched the search token.
5. **Point with index finger.** One participant chose this gesture to open the application while other chose it to start writing the message.

4.2.3. Learning Effect

While performing the tasks on the prototype, 26 gesture-related errors were registered. We considered errors: (1) forgetting what gesture to do to trigger an action; (2) performing the wrong gesture; and (3) selecting a wrong list item due to taking too long to recall the gesture. On average 0.96 errors occurred per participant ($SD = 1.20$).

It is interesting to analyze if those errors diminished over time. We registered the number of errors during the first task ($M = 1.33$, $SD = 1.5$), second task ($M = 1.00$, $SD = 1.32$), and third task ($M = 0.56$, $SD = 0.73$) as depicted in **Figure 3**. Observing the mean values, data suggests that participants progressively memorized the gestures needed to complete the tasks. For this reason, a statistical analysis was conducted. In order to check for any significant differences between task errors, we began



by assessing the normality of the data. The results for all tasks showed that normality was not verified ($p_1 = 0.077$; $p_2 = 0.005$; $p_3 = 0.008$). Consequently, we conducted a non-parametric Friedman test, comparing the errors made in each task, which revealed that there was not a statistically significant difference ($\chi^2(2) = 1.826$; $p = 0.401$).

4.2.4. Reading Speed Preferences for Selection Tasks

Participants were asked to rate their preference of 3 different reading speeds on two contexts: selecting a word from the message; and selecting a contact from the list of contacts. The purpose was to understand how fast they are able to react and perform the gesture to select a word or contact.

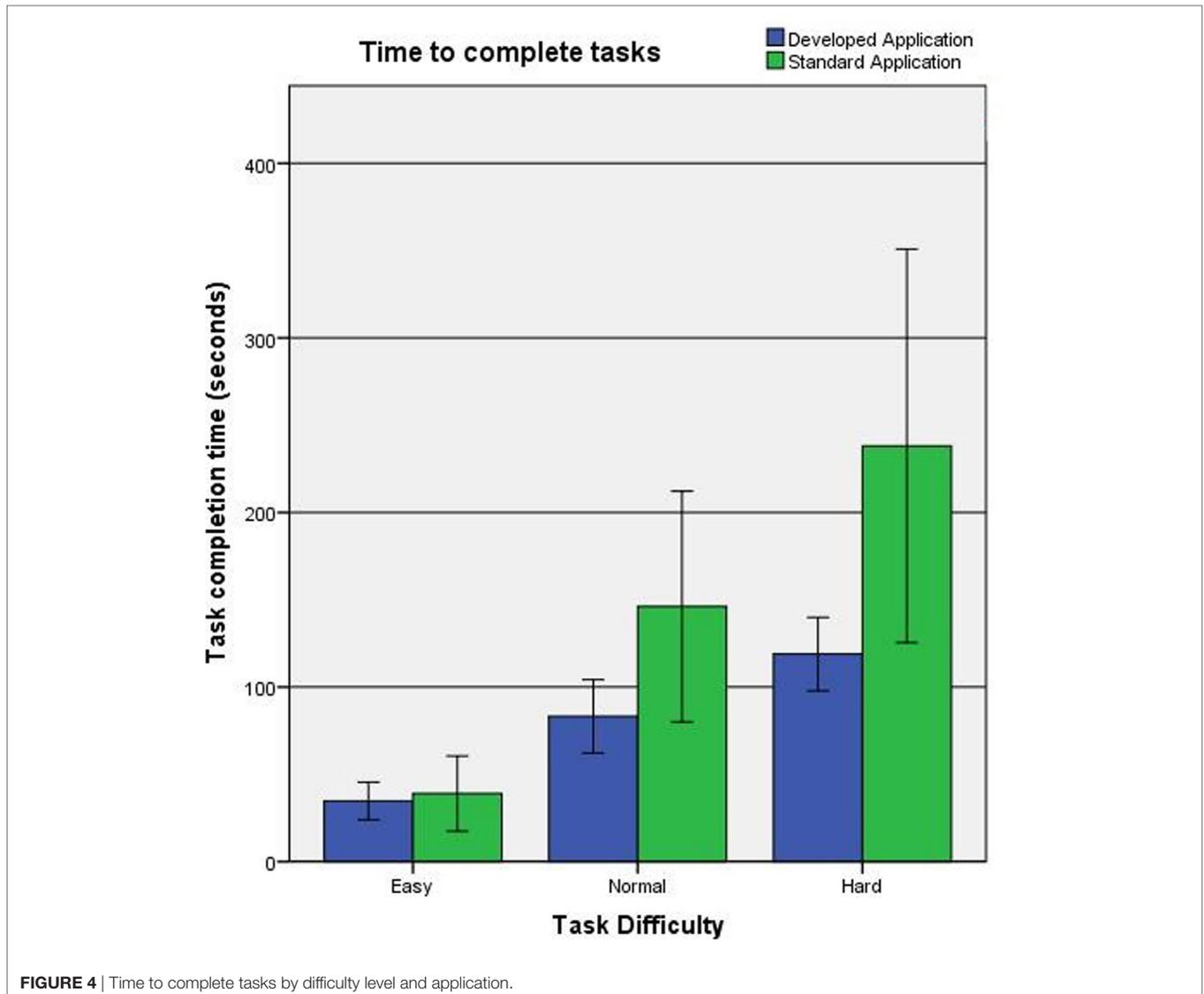
Findings show that despite 22% of the participants fail to select the right word while using the fastest speed, over 55% preferred this option. Those who failed stated they could do it with more training. On the other hand, when selecting a contact from the list, there was not a preferred reading speed choice. Four participants preferred the fastest speed and other 4 preferred the middle one. However, 33% of the participants failed when selecting a contact in the fastest speed mode.

4.2.5. Task Execution Times

Figure 4 presents the time taken to complete the three tasks with both applications. When using the developed prototype, participants took, on average, 34.78 s ($SD = 10.77$) to complete the easy task. For the normal task, participants took, on average, 83.22 s ($SD = 20.94$) to finish it. Finally, for the hard task, participants needed 118.89 s ($SD = 21.03$), on average, to complete it. When using the standard SMS application, participants took 39 s ($SD = 21.54$) to complete the easy task, on average. Participants completed the normal task in 146.11 s ($SD = 66.08$), on average. The hard task took, on average, 238.11 s ($SD = 112.74$) to finish.

By observing the data, two trends seem to emerge. More difficult tasks took more time to complete, as expected, and tasks were completed quicker when made with the prototype. To validate these observations we need to look for the effects of the independent variables (task difficulty and application) on the dependent variable (time to complete the task). We began by assessing the normality of the data. The results showed that the data follows a normal distribution except for one group (easy task on standard application).

Because of its robustness regarding violation of normality, a parametric two-way repeated measures ANOVA test was



still performed. Results show that there is a significant interaction between the two independent variables ($F(2,16) = 8.024$, $p = 0.004$). In order to understand this interaction we need to determine the difference between the groups at each level of each factor.

The first factor to be analyzed was the application used. We computed a simple main effect test to observe the differences between the two applications for each level of task difficulty. Results showed that for the easy task we could not find any significant difference between the two applications ($p = 0.664$). However, we found statistical differences in the normal ($p = 0.026$) and hard ($p = 0.012$) tasks. We can conclude that our prototype allowed the study participants to complete the tasks quicker than the standard SMS application in the tasks that required message editing, i.e., the normal and hard tasks. In the task that simply required text input and the selection of a contact in the beginning of the list of contacts there were no differences between the applications, even though, on average, our prototype was still more than 4 s quicker than the standard application.

The second factor analyzed was the task difficulty. We ran another simple main effect test to assess the differences between task difficulty within each application. Observing the pairwise comparisons, the results show a significant difference for easy and normal tasks ($p < 0.001$), easy and hard tasks ($p < 0.001$), and normal and hard tasks ($p = 0.001$) for the prototype. The same results occurred in the standard application. Findings show that there is also a significant difference between easy and normal ($p = 0.001$) tasks, easy and hard tasks ($p < 0.001$), and normal and hard tasks ($p = 0.012$). Regarding task difficulty, we can conclude that more difficult tasks take more time to complete, independently of the application used, which is an expected result.

4.2.6. SUS Analysis

Both applications rated high on the SUS scale. The prototype averaged 74.44 points, while the standard SMS applications rated, on average, 74.72 points. To check for any statistically significant differences regarding the SUS scores for both the standard applications and the developed prototype, we first assessed the normality

of the data. The results of Shapiro–Wilk test showed data with a normal distribution for both conditions. A paired t-test showed that there was not a significant effect ($t(8) = -0.0322, p = 0.975$) on the SUS score when comparing the results between the standard SMS applications ($M = 74.72, SD = 20.78$) and the prototype ($M = 74.44, SD = 10.66$).

4.3. Discussion

One of our goals was to understand how participants would react to a different form of interaction with their smartphones, namely one that involved performing gestures but not on their smartphones' touchscreen. Our findings are in accordance with previous work by Profita et al. (2013) and (Rico and Brewster, 2009, 2010). Our subjects are conscious of their actions in a public setting and some state they would feel embarrassed performing such gestures. However, if those gestures could be performed more subtly they are willing to use them if that would bring interaction advantages.

Regarding the definition of gestures, the most significant conclusion is that every confirmation command should be triggered by the same gesture, independently of the context. This supports a consistent interaction and reduces the number of gesture to memorize. Another significant observation is that we could hardly find repeated gestures for the same action across participants. Even though the sample of participants recruited is small, the diversity of proposed gestures leads to the conclusion that, for this kind of interaction, users need to be able to define their own gestures, thus the interface should support customization. Still, one participant suggested to replicate the already known touch gestures, for the actions where this solution could be applied. This would reduce the cognitive load and the need to learn new gestures.

The majority of participants preferred quicker reading velocities despite not being able to select an item in time, arguing that with enough training they would be able to perform correctly. This opens up an opportunity to assist users, through a correction algorithm that would adapt the selection as a function of the users' reaction time.

When comparing participants' performance in their usual text messaging application and a prototype developed following our proposed approach, we were able to find advantages in our prototype. While there were no significant differences when the task required no editing of the message, we found that our prototype supported quicker message correction than the standard solutions available. This results from the careful design of the message navigation and correction mechanisms. This prototype was designed to take advantage of the possibility to have commands issued by the user, through on-body gestures, while aurally inspecting the entered message. Additionally, since those commands need not be entered *via* the smartphone's touchscreen, users are able to instantly input different commands, instead of having to navigate a list of commands, which is the current paradigm.

Regarding the brief usability assessment, through the SUS questionnaire, findings showed no statistically significant difference between the standard SMS application and the developed prototype. Still, both achieved a high SUS score of 74 points.

According to the SUS scoring system,¹ results higher than 68 are above average therefore considered usable. Additionally, 5 participants stated that *“this application is much easier than the one that already exists for the blind”* and two other participants said it to be *“very functional upon learning the gestures.”* Moreover, 8 participants stated that the *“keyboard is too small which can lead to errors while writing the message”* when discussing current applications. However, two participants raised concerns about using an online voice recognizer, stating the limitations of being required to be always connected to the Internet because of the lack of performance of current offline versions.

The quantitative findings of this study have limited generalization capability due to the small number of participants. Nevertheless, the qualitative findings are important from a formative perspective, offering initial support for the future design and development of applications relying on body interaction techniques, especially for, but not limited to, visually impaired users. Another limitation of the proposed approach is the reliance on sEMG sensors and their low accuracy for gesture recognition if we wish to be able to recognize subtle gestures. One possibility to overcome this limitation is using sensors based on different signals, as long as they remain as inconspicuous as possible. One such sensor can be the FlickTek,² which can be used in a bracelet, and works based on the biometric signals generated by the movement of tendons to recognize finger gestures.

5. CONCLUSION

Despite including accessibility options, both Android and iOS mobile systems still lack decent support for some types of activities and contexts of usage. For instance, writing text is still a time-consuming task for users who suffer from some kind of visual impairment. Besides, to write a message, users currently have to use both hands. This is particularly important when considering this user group since they need to hold a cane or a guide dog in one hand while walking.

Our approach envisions to make the interaction with mobile devices 100% hands free by exploring the advantages of multimodal techniques, specially the use of mid-air and on-body gestures combined with voice recognition. While the ultimate challenge is to design an interaction model for the whole system, this article presented an initial case study focused on an application for the sending of text messages. A user study was conducted and results showed positive feedback from the participants regarding the use of body based interaction. We also found the proposed design to have a better performance when text editing was required, when compared with current available text messaging applications. Other findings raised the need for socially acceptable or conspicuous gestures, and the usefulness of customized and adaptive interaction techniques.

Reserved for the future is the integration of a fully functional gesture recognizer which takes into account the feedback obtained during this first experiment (e.g., subtle gestures, possibility to

¹<http://www.measuringu.com/sus.php>

²<http://www.flicktek.com/>

customize gestures). Moreover, we intend to expand and generalize this novel interaction method into the whole operating system, making it available as an accessible service so that all applications can benefit from it.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Regulamento da Comissão de Ética para a Recolha e Protecção de Dados de Ciências with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Comissão de Ética para a Recolha e Protecção de Dados de Ciências.

REFERENCES

- Azenkot, S., and Lee, N. B. (2013). "Exploring the use of speech input by blind people on mobile devices," in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '13* (New York, NY: ACM), 1–8.
- Bonner, M. N., Brudvik, J. T., Abowd, G. D., and Edwards, W. K. (2010). "No-look notes: accessible eyes-free multi-touch text entry," in *Proceedings of the 8th International Conference on Pervasive Computing, Pervasive'10* (Berlin, Heidelberg: Springer-Verlag), 409–426.
- Brooke, J. (1996). "Chapter SUS: a quick and dirty usability scale," in *Usability Evaluation in Industry*, eds P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester (London: Taylor and Francis), 189–194.
- Caporusso, N., Trizio, M., and Perrone, G. (2014). "Chapter pervasive assistive technology for the deaf-blind need, emergency and assistance through the sense of touch," in *Pervasive Health: State-of-the-Art and Beyond*, eds A. Holzinger, M. Ziefle, and C. Röcker (London: Springer), 289–316.
- Costa, D., and Duarte, C. (2015). "From one to many users and contexts: a classifier for hand and arm gestures," in *Proceedings of the 20th International Conference on Intelligent User Interfaces, IUI '15* (New York, NY: ACM), 115–120.
- Costanza, E., Inverso, S. A., Allen, R., and Maes, P. (2007). "Intimate interfaces in action: assessing the usability and subtlety of EMG-based motionless gestures," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07* (New York, NY: ACM), 819–828.
- Demyntev, A., and Paradiso, J. A. (2014). "Wristflex: low-power gesture input with wrist-worn pressure sensors," in *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, UIST '14* (New York, NY: ACM), 161–166.
- Dezfuli, N., Khalilbeigi, M., Huber, J., Müller, F., and Mühlhäuser, M. (2012). "Palmrc: imaginary palm-based remote control for eyes-free television interaction," in *Proceedings of the 10th European Conference on Interactive Tv and Video, EuroITV '12* (New York, NY: ACM), 27–34.
- Dumas, B., Solórzano, M., and Signer, B. (2013). "Design guidelines for adaptive multimodal mobile input solutions," in *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '13* (New York, NY: ACM), 285–294.
- Gollner, U., Bieling, T., and Joost, G. (2012). "Mobile lorm glove: introducing a communication device for deaf-blind people," in *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction, TEI '12* (New York, NY: ACM), 127–130.
- Guerreiro, T., Lagoa, P., Nicolau, H., Gonçalves, D., and Jorge, J. (2008a). From tapping to touching: making touch screens accessible to blind users. *IEEE MultiMedia* 15, 48–50. doi:10.1109/MMUL.2008.88
- Guerreiro, T., Lagoa, P., Nicolau, H., Santana, P., and Jorge, J. (2008b). "Mobile text-entry models for people with disabilities," in *Proceedings of the 15th European Conference on Cognitive Ergonomics: The Ergonomics of Cool Interaction, ECCE '08* (New York, NY: ACM), 39.

AUTHOR CONTRIBUTIONS

CD and BD conceptualized the study. CD, SD, and DC designed the study and analyzed the data. SD implemented the prototype. SD and DC conducted the study and collected the data. CD and DC drafted the manuscript. SD and BD provided revisions. All authors approved the final version of the manuscript for submission.

FUNDING

This research was supported by Fundação para a Ciência e Tecnologia (FCT) through a PhD grant SFRH/BD/111355/2015.

- Harrison, C., Tan, D., and Morris, D. (2010). "Skinput: appropriating the body as an input surface," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10* (New York, NY: ACM), 453–462.
- Hinman, R. (2012). *The Mobile Frontier: A Guide for Designing Mobile Experiences* (New York, NY: Rosenfeld Media).
- Jones, B. (1975). Spatial perception in the blind. *Br. J. Psychol.* 66, 461–472. doi:10.1111/j.2044-8295.1975.tb01481.x
- Kane, S. K., Bigham, J. P., and Wobbrock, J. O. (2008). "Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques," in *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility, Assets '08* (New York, NY: ACM), 73–80.
- Kangas, E., and Kinnunen, T. (2005). Applying user-centered design to mobile application development. *Commun. ACM* 48, 55–59. doi:10.1145/1070838.1070866
- Li, F. C. Y., Dearman, D., and Truong, K. N. (2009). "Virtual shelves: interactions with orientation aware devices," in *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology, UIST '09* (New York, NY: ACM), 125–128.
- Li, F. C. Y., Dearman, D., and Truong, K. N. (2010). "Leveraging proprioception to make mobile phones more accessible to users with visual impairments," in *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility, ASSETS '10* (New York, NY: ACM), 187–194.
- Lin, S.-Y., Su, C.-H., Cheng, K.-Y., Liang, R.-H., Kuo, T.-H., and Chen, B.-Y. (2011). "Pub – point upon body: exploring eyes-free interaction and methods on an arm," in *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11* (New York, NY: ACM), 481–488.
- Makino, Y., Sugiura, Y., Ogata, M., and Inami, M. (2013). "Tangential force sensing system on forearm," in *Proceedings of the 4th Augmented Human International Conference, AH '13* (New York, NY: ACM), 29–34.
- Matthies, D. J. C., Perrault, S. T., Urban, B., and Zhao, S. (2015). "Botential: localizing on-body gestures by measuring electrical signatures on the human skin," in *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '15* (New York, NY: ACM), 207–216.
- Molin, L. (2004). "Wizard-of-oz prototyping for co-operative interaction design of graphical user interfaces," in *Proceedings of the Third Nordic Conference on Human-computer Interaction, NordiCHI '04* (New York, NY: ACM), 425–428.
- Nicolau, H., Montague, K., Guerreiro, T., Guerreiro, J. A., and Hanson, V. L. (2014). "B#: chord-based correction for multi touch Braille input," in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, CHI '14* (New York, NY: ACM), 1705–1708.
- Nicolau, H., Montague, K., Guerreiro, T., Rodrigues, A., and Hanson, V. L. (2015). "Holibraille: multipoint vibrotactile feedback on mobile devices," in *Proceedings of the 12th Web for All Conference, W4A '15* (New York, NY: ACM), 30.
- Norman, D. A., and Draper, S. W. (1986). "User centered design," in *New Perspectives on Human-Computer Interaction*, Hillsdale, NJ: Laurence Erlbaum Associates Inc.
- Oakley, I., and O'Modhrain, S. (2005). "Tilt to scroll: evaluating a motion based vibrotactile mobile interface," in *Eurohaptics Conference, 2005 and Symposium*

- on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint* (Washington, DC: IEEE Computer Society), 40–49.
- Profita, H. P., Clawson, J., Gilliland, S., Zeagler, C., Starner, T., Budd, J., et al. (2013). “Don’t mind me touching my wrist: a case study of interacting with on-body technology in public,” in *Proceedings of the 2013 International Symposium on Wearable Computers, ISWC '13* (New York, NY: ACM), 89–96.
- Rico, J., and Brewster, S. (2009). “Gestures all around us: user differences in social acceptability perceptions of gesture based interfaces,” in *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '09* (New York, NY: ACM), 64.
- Rico, J., and Brewster, S. (2010). “Usable gestures for mobile interfaces: evaluating social acceptability,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10* (New York, NY: ACM), 887–896.
- Southern, C., Clawson, J., Frey, B., Abowd, G., and Romero, M. (2012). “An evaluation of Braille touch: mobile touch screen text entry for the visually impaired,” in *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI '12* (New York, NY: ACM), 317–326.
- Svanaes, D., and Seland, G. (2004). “Putting the users center stage: role playing and low-fi prototyping enable end users to design mobile systems,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04* (New York, NY: ACM), 479–486.
- Thinus-Blanc, F., and Gaunet, C. (1997). Representation of space in blind persons: vision as a spatial sense? *Psychol. Bull.* 121, 20–42. doi:10.1037/0033-2909.121.1.20
- Williamson, J. R., Crossan, A., and Brewster, S. (2011). “Multimodal mobile interactions: usability studies in real world settings,” in *Proceedings of the 13th International Conference on Multimodal Interfaces, ICMI '11* (New York, NY: ACM), 361–368.
- Ye, H., Malu, M., Oh, U., and Findlater, L. (2014). “Current and future mobile and wearable device use by people with visual impairments,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14* (New York, NY: ACM), 3123–3132.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Duarte, Desart, Costa and Dumas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.