

No-Look Notes: Accessible Eyes-Free Multi-Touch Text Entry

Matthew Bonner, Jeremy Brudvik, Gregory Abowd, W. Keith Edwards

GVU Center & School of Interactive Computing, Georgia Institute of Technology
85 5th Street NW Atlanta, GA 30308 USA

{matt.bonner, jbrudvik, abowd, keith}@gatech.edu

Abstract. Mobile devices with multi-touch capabilities are becoming increasingly common, largely due to the success of the Apple iPhone and iPod Touch. While there have been some advances in touchscreen accessibility for blind people, touchscreens remain inaccessible in many ways. Recent research has demonstrated that there is great potential in leveraging multi-touch capabilities to increase the accessibility of touchscreen applications for blind people. We have created No-Look Notes, an eyes-free text entry system that uses multi-touch input and audio output. No-Look Notes was implemented on Apple's iPhone platform. We have performed a within-subjects ($n = 10$) user study of both No-Look Notes and the text entry component of Apple's VoiceOver, the recently released official accessibility component on the iPhone. No-Look Notes significantly outperformed VoiceOver in terms of speed, accuracy and user preference.

Key words: accessibility; mobile device; multi-touch; touchscreen; text entry; eyes-free

1 Introduction

The development of touchscreens sensitive to multi-finger input has sparked a renaissance of interest in the technology's popularity. Devices like the Apple iPhone are part of a rush to take advantage of this wave of interest. Touchscreens also bring undeniable utility, reducing the need for peripherals like keyboards and lending themselves to collaborative use. Effective and flashy, touchable interfaces have appeared at the cash register of the supermarket, at the check-in line of the airport, and in the pockets of the masses.

Unfortunately, as discussed by McGookin *et al.* [1], the creation of accessible modifications and enhancements for touch-based devices is lagging behind the breakneck pace of mainstream development. Touchscreens pose an especially daunting challenge for blind or visually impaired users. In place of familiar devices like the keyboard, screen computing generally offers a uniform, featureless surface. The trend of emulating existing GUI's while using the finger as a mouse can translate existing problems into an even more difficult context. On top of this, mobile systems often feature unique or restructured interface layouts. The end result is a confounding environment for accessibility programs.

Eyes-free use by sighted users is also difficult. Touchscreens demand the user's visual attention, making it extremely difficult to carry out other tasks simultaneously. In some cases, social norms require discreet or covert use of a touchscreen. For example, a meeting participant could respond to an urgent text message or jot down a note without appearing inattentive. There are even situations where focusing on a touch screen is physically hazardous, such as sending a text message while driving.

On a system like the iPhone, users enter text for a myriad of tasks including text messaging, internet browsing and the use of third-party applications. Text entry proceeds with a virtual keyboard, creating a plethora of miniature, adjacent targets that are completely indistinguishable from one another without visual examination. Competitors like the Blackberry Storm by Research In Motion have followed suit with their own virtual keyboards. Text entry is thus a primary means of interacting with devices like the iPhone and Blackberry Storm, but it is extremely difficult to do without visual input.

Following this conviction, we developed No-Look Notes, an eyes-free gesture-based text entry system for multi-touch devices. No-Look Notes offers two-step access to the 26 characters with a small number of simple gestures that remove the precise targeting required by soft keyboards. Some basic text editing actions, such as backspace, are included as well. The combination of gestures and multi-touch capabilities results in a system that is both explorable and expeditious.

We begin this paper with a discussion of related work and a formative pilot study. We then present the design principles that follow from this synthesis, and describe the design of No-Look Notes. Next we report the results of a user study with visually impaired subjects, testing both No-Look Notes and the text entry component of Apple's new eyes-free accessibility tool, VoiceOver. We conclude with an analysis of our results and discuss future work.

2 RELATED WORK

Text entry on mobile devices is a well-studied research area. Mackenzie and Soukoreff give a thorough overview of text entry systems on mobile devices using pen-based and keyboard-based input [2]. Our system, No-Look Notes, focuses on a third paradigm of input: multi-touch text input. Eyes-free multi-touch text entry is an exceptionally young research topic, so our related work draws on both single and multi-touch text entry work, as well as the haptic augmentation of touchscreens.

2.1 Eyes-Free Single-Touch Text Entry on Touchscreens

Sánchez and Aguayo developed "Mobile Messenger for the Blind" [3], a messaging system for mobile devices that divided the screen into a 9-button virtual keyboard with multiple characters on each virtual key, much like a mobile phone keypad. To specify a character, a key must be tapped multiple times. The system used text-to-speech (TTS) for output.

While it is clearly difficult for blind people to perform positional input, such as using a computer mouse or finding targets on a touchscreen, Mobile Messenger for the Blind made this task easier by keeping the number of targets low, placing them at easy-to-reference locations based on the physical characteristics of the device (*e.g.*, the edges and the corners of the screen), and keeping the targets static.

Increasing the number of targets will not only make a target-based system more difficult to explore, but also more difficult to enter text on. Our system, No-Look Notes, completely avoids the accuracy issues of requiring users to precisely tap several times on a target by using a multi-touch input system. By using a circular layout, No-Look Notes may also more easily add new input ‘targets’ while maintaining a simple exploration strategy.

Numerous systems rely on gestures for text entry, sometimes in combination with targets. Tinwala and MacKenzie developed a system based on Graffiti strokes for eyes-free text entry [4]. While the system is targeted at non-visual use, Graffiti strokes require the user to trace complicated forms (*e.g.*, english letters) onto the screen. Ken Perlin’s Quikwrite [5], though not targeted at visually impaired users, attempts to speed Graffiti entry by replacing letter shapes with regions. Users enter text by dragging to a group of letters present on a particular screen region, then to a secondary subregion to enter a specific letter. As described by Plimmer *et al.* [6], writing and learning to write is an extremely challenging task for visually impaired users. It is unlikely that a Graffiti-based system, or a system involving multiple precise and sequential gestures, will be usable for many visually impaired users.

Yfantidis and Evreinov created a system that used extremely simple unidirectional single-finger directional gestures that mapped onto pie-menus of characters [7]. A gesture in any of the 8 compass directions (North, Northeast, East...) corresponded to a unique character. 24 characters were mapped onto 3 separate “layers,” or pie-menus, which are traversed based on the delay of the finger between touching the screen and commencing a gesture. This system also strictly used TTS for output.

The gesture-based approach of this system does not require the user to hit any targets. The time-delay to switch between screens could be improved upon by introducing multi-touch interaction techniques, but this was a hardware limitation and not a system implementation issue.

At Google, T. V. Raman has created a system that incorporates elements of both gestures and soft buttons [8]. As in Yfantidis and Evreinov’s system, characters are arranged into ‘rings’ which are selected by a gesture towards the desired character. Two representative letters from each ring are arranged on an additional default ring. Users enter letters by first selecting one of these representative characters, which transitions the screen to a secondary ring. On this secondary ring, users select a character, and begin the process anew. This design speeds text input by removing the need to wait to move between menu screens.

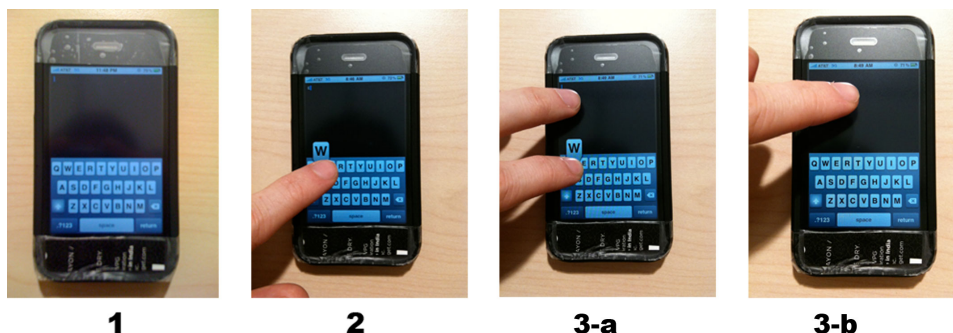


Fig. 1. Entering a character using VoiceOver: **1** Basic ‘soft’ QWERTY keyboard. **2** Rest finger on desired character. **3-a** Tap screen anywhere with second finger (‘split-tap’) or *alternatively*. **3-b** Lift up first finger and then double-tap the screen anywhere.

Both Yfantidis and Raman attempt to avoid the issue of accurately tapping by using gestures and have explicitly chosen simple, straight-line gestures. Despite this, accurately gesturing without visual feedback is challenging. Users must also first ensure their gesturing finger is oriented properly on the phone. No-Look Notes addresses the accuracy issues of targets by adding multi-touch instead of switching to gestures, thereby avoiding these potential pitfalls.

The largest problem with these two systems however is their lack of an effective exploration method. Users must continually swipe in different directions until the desired letter is found, deleting the character every time it is incorrect. After entering a letter, users must distinguish that series of swipes and delays from all the previous erroneous but highly similar input gestures. Introducing a completely novel and non-trivial organization of letters, as in Raman’s system, could further the confusion. Because it uses a multi-touch technique for text entry, No-Look Notes is able to reserve single-fingered interaction for exploration. Users may drag their finger around the screen seeking a letter without entering characters along the way.

2.2 Multi-Touch Interaction

During No-Look Note’s development, Apple announced “VoiceOver,” a system for eyes-free use of the iPhone that relies on touch-input and TTS output [9]. VoiceOver’s general scheme uses a ‘focus’ box which is moved between UI elements either by touching an element or flicking a finger in the direction the user wants focus to move. Double-tapping the screen activates the item currently inside the focus box. “Split-tapping,” in which the user first rests a finger on the desired item and then touches the screen with a second finger, may also be used to select an item. An early example of this general strategy of providing an ‘audio overlay’ for an unaltered GUI by moving focus between UI elements was demonstrated by Mynatt and Edwards [10].

Text entry in VoiceOver is subsumed by this basic focus-box system. Users interact with a soft keyboard (and any other application) by using this system:

touching near the key they wish to use, ‘swiping’ to move the focus box until it is correctly targeted if necessary, and split-tapping or double tapping to enter the key (see Figure 1). VoiceOver has additional functionality for examining the entered text. Users may swipe a finger up or down on the screen to move a cursor/insertion point through a line of text, reading each character the cursor moves over. In the rest of this paper, when referring to “VoiceOver,” we refer only to this text entry component.

Because it relies on a QWERTY keyboard layout, VoiceOver presents the user with a large number of targets. This can make it difficult to locate a letter, even though the double-tapping/split-tapping system helps avoid target-tapping accuracy issues. Additionally, the screen will be a non-sensical jumble of letters to users unfamiliar with computer keyboards.

No-Look Notes also uses split-tapping to allow the use of targets while avoiding their potential for accuracy issues. We further ease exploration by minimizing the number of targets on the screen at one time, using a simple alphabetical character-grouping scheme based on phone keypads.

As Apple’s official accessibility system included by default on the latest iPhone versions, VoiceOver will be distributed as the *de facto* accessible text entry system for a very large number of users. This makes VoiceOver both the first publicly released and first widely distributed multi-finger text entry system for touch screens. No-Look Notes was extensively and directly compared against VoiceOver in our evaluation, in which we evaluate both systems.

VoiceOver’s “split tap” is identical to the “second-finger tap” developed in Slide Rule by Kane *et al.* [11]. Slide Rule is a set of multi-finger interaction techniques for list-based applications, such as a phonebook and a music player. A complementary pair of techniques developed were the *one-finger scan* and *second-finger tap*, where sliding one finger on the screen is used to browse lists and a second finger tap, while still touching the screen with the first finger, selected an item in the list.

This approach both promotes risk-free exploration of items and allows blind users to select items without hitting a target on the screen. It is clear from this system that relatively simple multi-finger interaction techniques can greatly improve the accessibility of touchscreens. Our system also employs the split-tap from Slide Rule and applies it to text entry.

2.3 Haptics and Touchscreens

Haptic feedback for touchscreens tends to appear as an augmentation of button-based interfaces, as in work by Brewster [12], Leung [13] and Kaaresoja [14]. Brewster in particular showed simple haptic augmentation was beneficial for entering text on a small virtual keyboard. Recent work by Yatani and Truong explores other uses for haptics, using vibration to assist a gesture-based system [15].

Our system eschews haptic feedback because we wanted to develop a system that could be used on any touchscreen without modification. Touchscreens not integrated in mobile devices (*e.g.*, mobile phones) are unlikely to have vibration

motors included. Systems like Yatani’s require extensive modifications even to a phone that offers vibration.

3 DESIGN

3.1 Pilot Study

The design of No-Look Notes was informed by both the aforementioned work and by a pilot study of a prototype system. This pilot system was our first attempt at using multi-touch interaction to improve a gesture-based system. In our prototype, letters and letter groups (grouped by frequency of predicted use) were arranged in pie menus around the screen and as in the systems by Yfantidis [7] and Raman [8].

Rather than cycling through the pie-menus by selecting a representative character or adding a time delay, we tested a multi-touch interaction technique. By resting a finger on the screen, the user switched to a second pie menu. Users thus entered a letter or letter group by swiping in a letter’s direction, or resting a finger on the screen and then swiping.

We tested our prototype with five visually impaired participants, each of whom entered text for about 1.5 hours. This testing exposed three interrelated qualities that are absolutely key for a successful eyes-free text entry system: (1) Robust Entry Technique, (2) Familiar Layout, (3) Painless Exploration.

Robust Entry Technique Eyes-free systems must provide text entry techniques that are both exceptionally simple and exceptionally error-tolerant. Without visual confirmation, it is difficult for users to make precise gestures or hit precise targets. Mundane interaction techniques rapidly increase in complexity when used in an eyes-free context. Small distinctions, such as the difference between gesturing ‘Left’ and ‘Up-and-Left’ or the difference between a full circle and a three-quarters circle, are especially difficult.

Familiar Layout A layout that is easy to conceptualize and related to familiar interfaces or groupings is critical for an eyes-free system. Because touch-screens and multi-touch gestures are already foreign to most visually impaired users, an interface needs to include a recognizable layout it can to reduce the cognitive load on its users. Alphabetization, for example, will be far more successful than a layout based on theoretical character frequencies.

Painless Exploration Users of an eyes-free system must be able to painlessly explore the system’s layout, not just correct their mistakes. Exploring by repeatedly entering and undoing actions is not acceptable, there must be some first-class exploration technique. “Entering and undoing” adds to the user’s cognitive load by requiring them to remember which entry was correct among many similar entries. Painless exploration is aided by a familiar layout and robust entry technique, but does not necessarily follow from these.

3.2 No-Look Notes

No-Look Notes arranges characters around the screen in an 8-segment pie menu reminiscent of the systems proposed by Yfantidis and Raman. However, each section of the menu contains multiple characters, such that all 26 letters of the English alphabet appear (see Figure 2). The 8 character groups (*e.g.*, ‘ABC’, ‘PQRS’) correspond to the international standard mapping a phone keypad to letters (Familiar Layout) [16].

The segments are soft-buttons which must be touched. When the user touches a segment, either by dragging their finger to a new segment or touching the screen, the characters in that segment are announced audibly. The user may drag and tap their finger around the screen (for example, tracing the screen’s edges) without accidentally entering characters (Painless Exploration).

Resting one finger on a segment and tapping a second finger (*i.e.*, split-tapping or second-finger tapping) selects that segment, bringing the user to a secondary screen with that segment’s characters from that selection arrayed alphabetically from the top to the bottom of the screen. Users select the desired character the same way they selected a character group. The user drags a finger until they hear the desired character announced, then drop a second finger to the screen to select (Robust Entry Technique, see Figure 3).

This simple arrangement makes it fast, easy, and risk-free to search for a character: simply drag a finger around the screen. Users may trace the edges of the screen and eventually reach any character, since every ‘pie slice’ reaches the edge of the screen. This also allows No-Look Notes to leverage some of the “pure” benefits of edges, as defined by Wobbrock *et al.* [17], such as higher accuracy.

In addition to character entry, No-Look Notes offers gestures for “space” and “backspace/undo.” Backspace is a quick swipe with one finger to the left, space is a quick swipe to the right. The backspace gesture is also used to cancel a selected character group without entering a character. Reading (TTS) and spelling currently entered text is triggered by a swipe down. These swipes require a minimum distance and speed in order to register, preventing a user exploring the screen by tapping or dragging their finger from accidentally activating them.

4 Evaluation

We performed a within-subjects evaluation of No-Look Notes and VoiceOver. Users spent 15 minutes learning a system in an interactive tutorial with an experimenter, then spent 1 hour using the system to enter words, and finally answered a brief questionnaire. The second system was then tested in the same way. Users were split into one group evaluating No-Look Notes first and another evaluating VoiceOver first to counterbalance order effects.

MacKenzie and Soukoreff deal extensively with evaluation in their treatise on mobile text entry [2]. The value of quantitative *and* qualitative results from users is emphasized. Measurement using words-per-minute, as well as examining the relationship between speed, accuracy and errors are also covered.



Fig. 2. No-Look Notes. The screen on the left is the main screen. The screen on the right is after the ‘ABC’ target is selected from the main screen. Visual representation was added to this figure for illustration.

4.1 Participants

We recruited 10 participants, 7 men and 3 women. Participants were recruited from the Atlanta Center for the Visually Impaired. The average age of participants was 40.8 (sd=10.85). All users were visually impaired, where visually impaired is defined as “requires assistive technology (screen readers, magnification) for computer use.” No participant was able to visually distinguish specific keys or letters on the phone screen for either system. Additionally, no participant could view the text they had entered.

4.2 Device

Participants entered text on an Apple iPhone, which has a 3.5 inch capacitive touchscreen. Non-touch sensitive points on the top and bottom of the phone (near the ear and mouthpiece) were taped over to give the screen tactile boundaries.

No-Look Notes was implemented as an iPhone OS 3.0 application. TTS for characters and actions (like backspace) was pre-synthesized using Mac OS X’s built-in TTS engine and loaded onto the phone as audio files. VoiceOver’s release introduced dynamic TTS on the iPhone, but developers do not have access to this functionality. No-Look Notes thus provided dynamic TTS using Flite [18].

VoiceOver’s iPhone OS 3.0 version was tested using a custom application to isolate the text entry portion of the system. No modifications to the actual text entry or TTS were made. However, target words were synthesized by Flite.

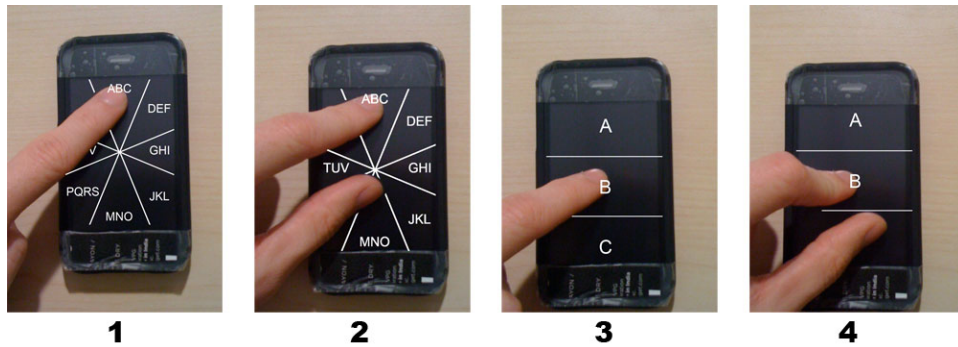


Fig. 3. Entering a character using No-Look Notes: **1** Rest finger on desired group **2** Tap screen anywhere with second finger ('split-tap') **3** Rest finger on desired character **4** Split-tap. Visual representation was added to this figure for illustration.

Participants tested VoiceOver in 'portrait' mode, with the keyboard taking up the bottom half of the screen.

4.3 Procedure

After their introduction to the iPhone's physical characteristics, participants ran the following procedure for each system. To avoid order effects, we counterbalanced participants such that half started with No-Look Notes and half started with VoiceOver.

First, participants spent 15 minutes learning the system. Experimenters taught the participants how to explore, enter and delete characters, as well as how to read characters on the screen. Participants were required to demonstrate their grasp of each new action, and encouraged to ask questions. Prior to beginning testing, participants demonstrated their competency by entering a test word.

Following this practice session, participants spent 1 hour entering single words from a published phrase set for text entry, also by MacKenzie [19]. This was selected over requiring participants to memorize phrases to reduce the cognitive load on the user, as also discussed by MacKenzie [2]. Phrases were randomly chosen from the phrase bank to create a phrase set; this same phrase set was used by all participants for each system.

Timing begins after the user enters a character. Each input (including errors) is duly noted by the system. An incorrect character entry caused an error sound to play, and the user was required to backspace/delete the offending characters. Timing continues until the moment the user's entered text completely matches the target word. At this point, the timer pauses and the next target is queued, waiting for the user to touch the screen.

Participants were able to rest whenever they wished, but were encouraged to rest between words rather than mid-word. Participants were also reminded several times that this was a test of the system, not their speed or spelling skills.

After completing 60 minutes of text entry, participants responded to a brief questionnaire about their opinions of the system used in that condition. Participants then repeated the 15-minute practice/learning + 60 minutes of use + questionnaire cycle with the other system.

5 RESULTS

Overall, we collected 20 total hours of usage data (1 hour per system * 2 systems per participant * 10 participants) counting rest time between words, which yielded a total of 3921 correct characters and 1135 correct words entered.

5.1 Physical Comfort

Some participants mentioned that their hands or fingers were tired. One participant noted that his hands were “tingly.” Although some users required minor breaks, all users were able to complete the entire study.

5.2 Text Entry Speed

Due to technical limitations of the iPhone SDK, we were unable to access gesture information for the VoiceOver condition. Therefore, we started timing of target words at the first input we could measure in the condition: after a character was entered. The same timing was used for the No-Look Notes condition. Since we do not time until after the first character is entered, we considered target length to be $n - 1$ when calculating the text entry speed measure. We used the WPM (words-per-minute) measure for text entry speed, calculated as (characters per second) * (60 seconds per min) / (5 characters per word). This timing technique and text entry speed measure are identical to those used by Wigdor and Balakrishnan [20]. While we chose to consider a “word” to be 5 characters in this analysis, in actuality, the average length of the 1135 target words entered was 3.45 characters.

Although we counterbalanced the order conditions across participants, it is still possible to encounter asymmetric transfer effects. To test this, we performed one t-test on each system’s text entry speed for each of the two groups formed by counterbalancing condition order (No-Look Notes first, VoiceOver first) for two t-tests total. The results of these tests were non-significant (No-Look Notes speed: $p = 0.50$, VoiceOver speed: $p = 0.88$) which suggests that there was no order effect.

The overall text entry speeds were 0.66 WPM for VoiceOver and 1.32 WPM for No-Look Notes, a 100% increase in favor of No-Look Notes. This difference was determined to be significant using a paired t-test ($p < 0.001$). Figure 4 shows text entry speed performance for each user for each system; all but one participant achieved a higher speed with No-Look Notes.

To examine how participants’ speed varied over the course of the session, we split up each hour into 6 10-minute blocks and calculated speed within each

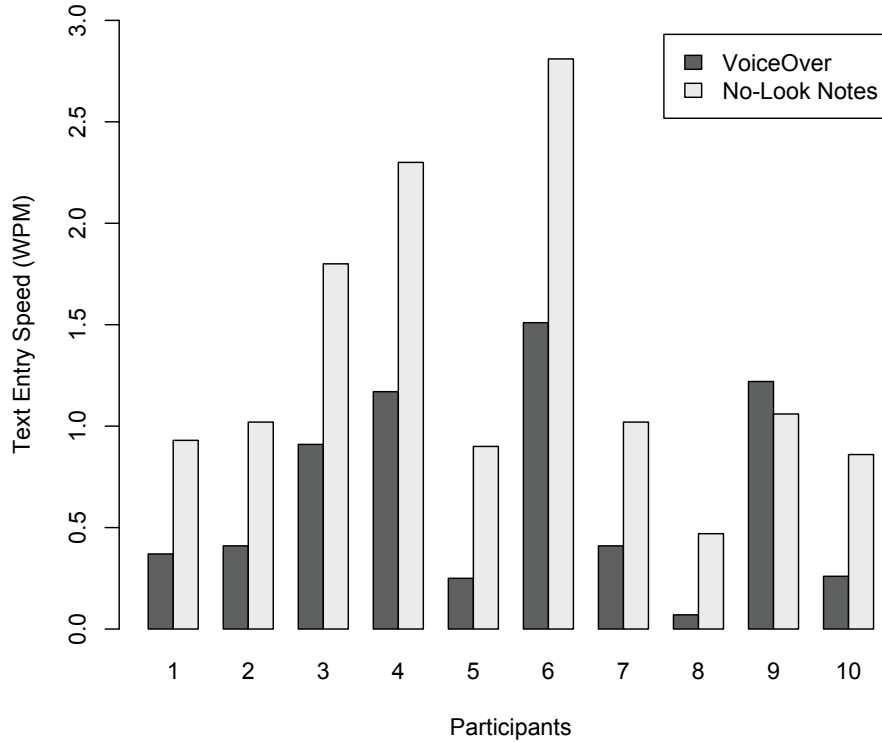


Fig. 4. Text entry speeds (WPM) for each participant for each system.

block. The averages across participants for block-level speed are shown in Figure 5. Text entry speeds for the first block for VoiceOver and No-Look Notes were 0.85 WPM and 0.61 WPM, respectively. The maximum text entry speed for each system occurred in the fourth block: 0.76 WPM for VoiceOver, 1.67 WPM for No-Look Notes. While these text entry speed curves may fit a general learning curve with many sessions over time, the participants' speed declined towards the end of our sessions, possibly due to fatigue.

5.3 Text Entry Errors

We calculate the error rate as (incorrect characters entered) / (correct characters in target word), as is also done by Widgor and Balakrishnan [20].

As with text entry speed, we tested for the occurrence of order effects for error rate. Neither order was significant for either condition (No-Look Notes error rate: $p = 0.71$, VoiceOver error rate: $p = 0.62$), so the performed counterbalancing was also acceptable here.

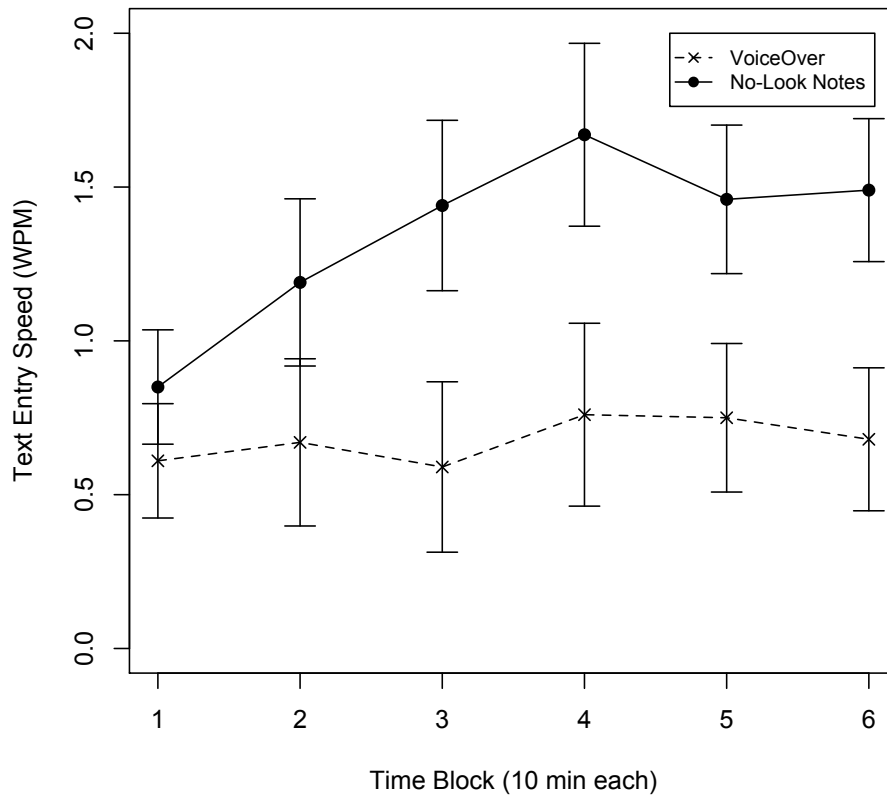


Fig. 5. Text entry speed (WPM) across participants for each 10-minute block (time interval).

The overall mean error rates across participants were 0.60 for VoiceOver and 0.11 for No-Look Notes, a 445% increase in errors for VoiceOver over No-Look Notes. Using a paired t-test, this difference was found to be significant ($p < .05$). Error rates for each participant for each system are shown in Figure 6.

5.4 Questionnaire Results

After each condition, participants responded to a brief questionnaire about their opinions of the system used in that condition. The questionnaire comprised 14 statements in which the participant would state their agreement on a 5-point Likert scale (1 = disagree strongly, 5 = agree strongly). All of the mean responses were higher for No-Look Notes than for VoiceOver. Using a paired Wilcoxon test,

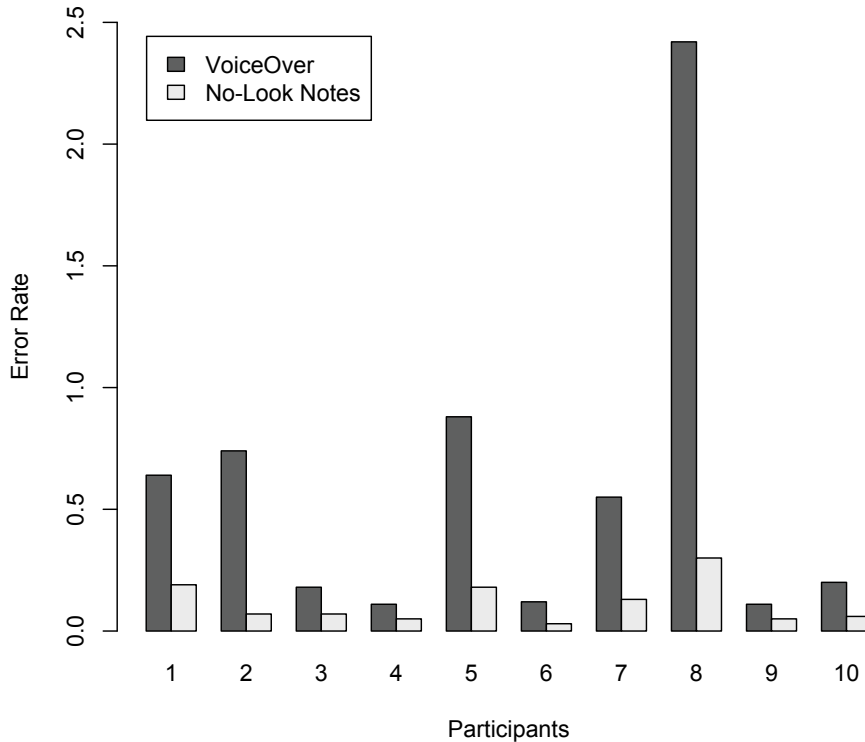


Fig. 6. Error rates for each participant for each systems.

we determined that 8 of 14 differences were significant ($p < .05$), all in favor of No-Look Notes, while 6 of the 14 were not significant. The list of statements, mean responses, and significant differences are shown in Table 1.

No-Look Notes was rated significantly higher than VoiceOver for the majority of statements, including key statements such as “easy to learn,” “fast to use” and “felt in control.” Our questionnaire results suggest that most users felt No-Look Notes was both easier to learn and faster overall. This supports our speed and error monitoring results, which shows that this was actually the case.

5.5 Qualitative Feedback

Seven participants responded negatively to the size and number of targets in VoiceOver, finding it difficult to locate keys. Participant 3 said “*If I wasn’t familiar with the QWERTY [keyboard layout], it would have been hell*”. Indeed, participant 8, our only participant with no QWERTY experience, achieved a rate of only 0.07 wpm with VoiceOver, saying mid-use: “*I want to cry right now.*”

Statement	VoiceOver	No-Look Notes
Easy to use*	2.6 (1.43)	4.2 (0.63)
Fun to use	3.4 (1.71)	4.7 (0.67)
Fast to use*	3.0 (1.83)	4.4 (0.70)
Felt in control*	3.0 (1.63)	4.2 (0.79)
Easy to learn*	3.2 (1.48)	4.6 (0.70)
Intuitive	3.4 (1.43)	4.4 (0.97)
Familiar	2.7 (1.57)	3.4 (1.26)
Features clear to me*	3.7 (0.95)	4.5 (0.71)
Improve with practice*	3.7 (1.25)	4.8 (0.42)
Would use this system	3.9 (1.37)	4.7 (0.67)
Made entering text accessible*	3.7 (1.25)	4.8 (0.42)
Aware of text I was entering*	2.8 (1.40)	4.5 (0.85)
Audio feedback clear	3.2 (1.55)	3.3 (1.70)
Easy to undo mistakes	3.1 (1.45)	4.4 (0.84)

Table 1. Questionnaire results (mean, sd) for responses given on a Likert scale (1 = disagree strongly, 5 = agree strongly). The mean value for each question favored No-Look Notes; asterisks indicate where this difference was significant.

Participants responded positively to the familiar aspects of both systems. Participants with knowledge of QWERTY keyboards (nine out of ten) said that thinking of the keyboard layout aided them in locating keys in VoiceOver. The fact that groups of characters in No-Look Notes are the same groups used on a phone keypad was also mentioned favorably by several participants. It was also useful for some participants to envision No-Look Notes’ layout as a clock (*e.g.*, ‘ABC’ at 12 o’clock). Participant 1 mentioned this, saying “*a lot of the learning I’ve done is clock learning anyway.*”

Feedback was mixed on whether the system should repeat a character when it was entered (VoiceOver) or merely give an audible “click” (No-Look Notes). The suggestion of reading the character back but in a different voice (*e.g.*, female) met with a favorable response.

Despite voicing their frustrations with both systems, at the end of the experiment participants were uniformly enthused about learning them. Participants maintained that the experiment had been fun, despite the frustration that was sometimes apparent during testing.

6 DISCUSSION

Our participants’ perseverance, despite clear frustration, was impressive. This high tolerance could be evidence that visually impaired users are inured to struggling through poorly designed systems. Tolerance could also be due to the novelty of using a touchscreen or desire that touchscreen phones gain accessibility (participant 1: “*If they were accessible I would buy one*”). Regardless of this tol-

erance’s source, visually impaired users are eager for a touchscreen accessibility solution and willing to put in the time to learn a new system.

While using the systems, users experimented with different ways of holding the device. Some used a single hand, others used thumbs, others set the device on the table and used either one or two hands. Users generally stuck with the same style of input for both No-Look Notes and VoiceOver. The majority of users preferred to leave the device on the table.

Related to device positioning was our decision to test only in a portrait orientation. VoiceOver also allows for a landscape orientation, which would give users more options and, more tangibly, allow VoiceOver to increase the size of its targets by about 25%. While this would likely improve VoiceOver’s performance, we feel it is unlikely to have a drastic effect. A slight change in target size would not provide the advantages of full screen use (such as tracing the screen’s edges) or the low number of targets present in No-Look Notes, nor would it address the audio feedback issue discussed below.

Some users also displayed a hesitance to enter a character if they weren’t sure of the results, hovering their finger above the screen as they meditated on which direction to gesture. There may be a significant psychological difference between a truly risk free ‘exploration gesture’ and actually entering a character, then deleting it.

6.1 Simple Entry Gestures

Both systems put emphasis on simple input gestures in an attempt to make accidental entry of characters difficult. The split-tapping technique, developed by Slide Rule as a tool for making selections off of lists [11], proved effective at selecting arbitrarily located targets, provided the targets were large enough. VoiceOver’s targets were small enough that the participants occasionally moved off of the target while preparing to tap the screen with a second finger.

VoiceOver’s double-tapping entry technique was generally more difficult for users, with only one participant preferring it to split-tapping. The key issue was tapping with sufficient speed. If a user was too slow, the system would interpret the tap as a touch and re-locate the system’s focus, losing the user’s place.

6.2 Text-Awareness

Participants expressed a desire for the ability to easily determine the text they had entered. While some were able to use VoiceOver’s cursor system (one even used the cursor to insert missing characters), others found the system difficult to grasp. No-Look Note’s simple read-and-spell gesture was usable for all participants, although waiting for the system to spell the word was galling to some.

It is key that users be able to quickly check both their recently entered text and a larger amount of their text. When users get lost or forget their place, they can be forced to erase their text and start anew. A combination of VoiceOver’s ability to precisely examine each character combined with a more rapid version of No-Look Note’s direct access to spelling may be best.

6.3 Audio Feedback

Audio was a key component of both systems, used for both locating characters and reading text. Users found both systems difficult to understand. The dynamic text to speech engine used in both systems, Flite [18], was particularly disliked.

Though Apple has created a dynamic TTS system for the iPhone, it is inaccessible to developers at the time of this writing. VoiceOver also effectively cuts developers off from interpreting gestures themselves, making it impossible to use many applications. These design decisions severely restrict developers, preventing them from harnessing the power of gestures and restricting them to button-based systems. This could lead to developers creating unnatural or unusable ports of gesture-based applications, or simply failing to make an accessible version at all.

Nearly all participants had difficulty distinguishing between like-sounding letters (*e.g.*, ‘M’ and ‘N’, ‘C’ and ‘Z’). This led to extreme frustration in VoiceOver’s QWERTY based layout, where like-sounding letters were often adjacent. Even the experimenters found it almost impossible to aurally distinguish certain letter groups. Using this TTS in a noisy environment or using the phone’s built-in speaker will exacerbate the problem, though headphones could help avoid this. A serious portion of VoiceOver’s high error-rate can be attributed to this issue.

No-Look Notes was able to escape this problem with its alphabetical groups, largely keeping like-sounding characters apart. Groups like ‘MNO’ were still usable, as users proved more easily able to locate the correct letter when faced with only a few large targets in alphabetical order.

6.4 Deployment

Our results show that No-Look Notes has a number of advantages over VoiceOver. VoiceOver’s focus-box input technique simply did not translate well to the large number of targets in a soft keyboard. We believe best way to use No-Look Note’s strengths is not to replace VoiceOver, however, but to integrate with it. VoiceOver is used not only for text-entry, but for navigating the entire iPhone interface. Rather than invoking a virtual keyboard when a text field is selected, No-Look Notes could be activated. A simple gesture to return to VoiceOver navigation would complete integration.

7 FUTURE WORK

Text-Messaging Evaluation Many users expressed a desire to send text messages using either system. An extended evaluation of both systems in which users actually send text messages would show how effective each system is at enabling this apparent ‘killer app.’

Visual Feedback No-Look Notes has no visual feedback to help with text entry or the system’s layout. Users with partial vision would benefit from clear graphics

on the touchscreen. Graphics need not even be text — colored or pattern areas could be used to designate pie segments, for example, or the screen could flash when a character was entered. This graphically augmented system could then be tested with both visually impaired and sighted users.

Refine Character-Entry Feedback Experimenters suggested using different voices for different types of feedback, such as using a different tone of voice to repeat characters as they were entered. Participants were enthusiastic about the idea. This could be an effective way of adding additional contextual information to eyes-free text entry systems. This would also help avoid confusion over whether hearing a character read meant that character had actually been entered or merely been touched.

Extend Character Set No-Look Notes could be extended to allow entry of numbers or symbols. This could involve adding more targets to the pie menu, or perhaps creating a modal input method of swapping between a ‘letters menu’ and a ‘numbers menu.’ VoiceOver features a numbers/symbols keyboard accessed by entering a ‘more’ button on the soft keyboard.

8 CONCLUSION

We introduced No-Look Notes, a system for eyes-free mobile text entry using multi-touch input. No-Look Notes was designed to take advantage of other work on accessibility and text entry. We also developed three design principles that are key for eyes-free text entry: Robust Entry Technique, Familiar Layout and Painless Exploration.

We have implemented No-Look Notes on an Apple iPhone. We conducted user trials with visually impaired participants to evaluate both No-Look Notes and Apple’s own VoiceOver system, offering a comparison of the two systems. Our study of VoiceOver is also a first look of what will become the first widely distributed system for eyes-free text entry on the iPhone platform.

Acknowledgements We are indebted to Anisio Correia and Leigh Cooper, our collaborators at the Center for the Visually Impaired of Atlanta. Rosa Arriaga lent essential guidance and expertise to our statistical analysis. Jeffrey Bigham and Bruce Walker gave invaluable advice, especially concerning our evaluation.

References

1. McGookin, D., Brewster, S., Jiang, W. Investigating touchscreen accessibility for people with visual impairments. In Proceedings of the 5th Nordic conference on Human-Computer Interaction: Building Bridges, 298-307. (2008)
2. MacKenzie, I. S., Soukoreff, R. W. Text Entry for Mobile Computing: Models and Methods, Theory and Practice. *Human-Computer Interaction*, 17(2), 147. (2002)

3. Sánchez, J., Aguayo, F. Mobile Messenger for the Blind. *Universal Access in Ambient Intelligence Environments*, 369-385. (2007)
4. Tinwala, H., MacKenzie, I. S. Eyes-Free Text Entry on a Touchscreen Phone. *Proceedings of the IEEE Toronto International Conference Science and Technology for Humanity TIC-STH 2009*, 83-89. (2009)
5. Perlin K. Quikwriting: continuous stylus-based text entry, *Proceedings of the 11th annual ACM symposium on User interface software and technology*, 215-216. (1998)
6. Plimmer, B., Crossan, A., Brewster, S. A., Blagojevic, R. Multimodal Collaborative Handwriting Training for Visually-Impaired People. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 393-402. (2008)
7. Yfantidis, G., Evreinov, G. Adaptive Blind Interaction Technique for Touchscreens. *Universal Access in the Information Society*, 4(4), 328-337. (2006)
8. Speech Enabled Eyes Free Android Applications <http://code.google.com/p/eyes-free/>
9. VoiceOver <http://www.apple.com/accessibility/iphone/vision.html>
10. Mynatt, E. D., Edwards, W. K. Mapping GUIs to Auditory Interfaces. In *Proceedings of the 5th annual ACM symposium on User Interface Software and Technology*, 61-70. (1992)
11. Kane, S. K., Bigham, J. P., Wobbrock, J. O. Slide Rule: Making Mobile Touch Screens Accessible to Blind people using Multi-Touch Interaction Techniques. *Proceedings of the 10th international ACM SIGACCESS Conference on Computers and Accessibility*, 73-80. (2008)
12. Brewster, S., Chohan, F., Brown, L. Tactile Feedback for Mobile Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing System*, 159-162. (2007)
13. Leung, R., MacLean, K., Bertelsen, M. B., Saubhasik, M. Evaluation of Haptically Augmented Touchscreen GUI Elements under Cognitive Load. In *Proceedings of the 9th International Conference on Multimodal Interfaces*, 374-381. (2007)
14. Kaaresoja, T. Snap-Crackle-Pop: Tactile Feedback for Mobile Touch Screens. *Proceedings of Eurohaptics*, 565-566. (2006)
15. Yatani, Koji, Truong, Khai N. SemFeel: A User Interface with Semantic Tactile Feedback for Mobile Touch-screen Devices. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, 111 – 120. (2009)
16. E.161: Arrangement of Digits, Letters and Symbols on Telephones and Other Devices that can be Used for Gaining Access to a Telephone Network <http://www.itu.int/rec/T-REC-E.161-200102-I/en>
17. Wobbrock, J. O., Myers, B. A., Kembel, J. A. EdgeWrite: A Stylus-Based Text Entry Method Designed for High Accuracy and Stability of Motion. In *Proceedings of the 16th annual ACM Symposium on User Interface Software and Technology*, 61-70. (2003)
18. Black, A., Lenzo, K.. Flite: A Small Fast Runtime Synthesis Engine. *4th ISCA Speech Synthesis Workshop*, 157-162. (2007)
19. MacKenzie, I. S., Soukoreff, R. W. Phrase Sets for Evaluating Text Entry Techniques. *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 754-755. (2003)
20. Wigdor, D., Balakrishnan, R. *TiltText*: Using Tilt for Text Input to Mobile Phones. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, 81-90. (2003)