Towards Inviscid Text-Entry for Blind People through Non-Visual Word Prediction Interfaces

Kyle Montague
Open Lab,
Newcastle University,
kyle.montague@ncl.ac.uk

João Guerreiro, Hugo Nicolau
INESC-ID; Instituto Superior Técnico, Universidade de Lisboa
joao.p.guerreiro@ist.utl.pt
hman@inesc-id.pt

Tiago Guerreiro, André Rodrigues
LaSIGE, Faculdade de Ciências, Universidade de Lisboa
tjvg@di.fc.ul.pt
afrodrigues@fc.ul.pt,

Daniel Gonçalves
INESC-ID; Instituto Superior Técnico, Universidade de Lisboa
daniel.goncalves@inesc-id.pt

Abstract

Word prediction can significantly improve text-entry rates on mobile touchscreen devices. However, these interactions are inherently visual and require constant scanning for new word predictions to actually take advantage of the suggestions. In this paper, we discuss the design space for non-visual word prediction interfaces and finally present Shout-out Suggestions, a novel interface to provide non-visual access to word predictions on existing mobile devices.

Author Keywords

Text-Entry; Blind People; Word Prediction; Concurrent Speech; Touchscreens.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

Introduction

In recent years, mobile text-entry has experienced numerous innovations leading to greater input speeds and lower error rates. One such leap was word prediction [1], which will be familiar to anyone using a modern smartphone. During the typing process the device will suggest the most probable word, allowing users to correct misspelled words or complete them without having to type the remaining characters. Word predictions and auto-correct can be beneficial for most users, however the greatest gains are seen by users with slower text-entry rates, whom struggle to make key selections [5].

Due to the nature of the *Explore-by-touch* paradigm, text-entry on touchscreen devices is very inefficient for blind people, on both smartphones and tablet devices [8,9]. In [8], we explored the QWERTY typing behaviors of novice blind smartphone users and found that between 13%-32% of time was spend correcting errors; contributing to the slow entry rate of 4 wpm**.**

Alternative mobile keyboard solutions have previously been proposed for blind and visually impaired users

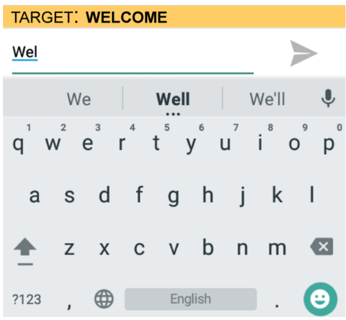

Figure 1. default word prediction on Android 5.0. The user has typed 'Wel', the system has identified 'Well' as the auto-correct suggestion but the target word was 'Welcome'.

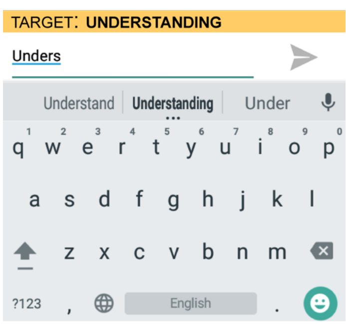

Figure 2. default word prediction on Android 5.0. The user has typed 'Unders', the most probable predicted word is 'Understanding' with two less probable alternative suggestions of 'Understand' and 'Under'.

[7,9,10]**.** Nevertheless, these solutions are not without additional challenges or shortcomings. First, they may require learning unfamiliar methods or keyboard layouts. Second, even methods such as chord-based braille input, which offer superior text-entry rates, do it at a significant cost to error rates. Existing chord-based braille input systems have reported average error rates of 45% [10], however Nicolau et al. 2014 [7] have since proposed a novel chord-based spellchecker to address the common chording errors. They demonstrated that their spellchecker could identify the correct word 69% of the time within the top 3 suggestions - yet, there is still a lack of support for non-visual suggestion selection and exploration on these devices, limiting the success of word predictions made, independently of their quality.

For these reasons, it is particularly appealing to provide suggestion selection to individuals with visual impairments with the aim of achieving input rates more inline with their inviscid text-entry rate [6]. A common bottleneck across all mobile non-visual text-entry methods is the feedback delivery; these methods rely on sequential audio output to both convey feedback about users' inputs and notify them of any content updates. Where a sighted person is able to quickly scan the entire display for input feedback, content changes, error notifications or auto-complete suggestions; a blind user receives only feedback for a single element at a given time. In this paper, we propose a novel non-visual auto-correction interface that enables blind users to leverage the existing text-entry suggestions through concurrent speech feedback [3].

Word Prediction on Touchscreens

Unfortunately, today's word prediction, or auto-correction interfaces are not designed to support non-visual interactions. These solutions rely on the user's ability to quickly glance at a word completion prompt or the suggestion bar (typically above the top row of the keyboard). Although these suggestions are visually displayed on screen, blind users are not directly made aware of the available word completion suggestion.

Google's Android and Apple's iOS provide respectively, no feedback; or a simple *earcon* alert; to inform the user of a highly probable completion suggestion. Meaning the user has either, no idea that a suggestion is available; or no understanding of what that suggestion actually is until it has been accepted and replaces their typed text. Moreover, the default behavior for these platforms is to autocorrect the typed text if the suggestion meets a probabilistic threshold of being the intended word. Since recent studies have shown that blind users spend a large amount of time correcting errors [8], this behavior is particularly damaging when it provides incorrect or erratic suggestions (e.g. Figure 1). Moreover, users may ignore an early suggestion by continuing typing the intended word, thus not benefiting from the suggestions (Figure 2).

Design of Non-Visual Word Prediction Interfaces

Independently of the input method, blind users' unawareness of correction suggestions displayed on screen is impairing their ability to leverage them as they input text. We propose providing feedback about suggestions before users accept them, so they can avoid later error correction and leverage early suggestions. In this section, we discuss the design space of word prediction interfaces for blind people.

Additional Feedback
Current screen reader interfaces rely on a single, sequential auditory channel to convey digital information to blind users. Adding another sequential step would undermine users' ability to input text as fast as possible. Moreover, previous work about the perception of concurrent speech [3] demonstrated blind people's ability to receive and interpret concurrent speech sources, suggesting that current screen readers could be imposing limitations on the way auditory feedback is being provided. Therefore, adding a secondary, concurrent auditory channel could be used to provide information about suggestions without interrupting users' regular text-entry feedback.

In order to increase users' ability to cope with the simultaneous speech feedback, the design of auto-correction interfaces should consider attributes that enhance speech segregation, such as using different voices and spatial locations for each speech channel. Previous work has shown that speech perception reduces as the number of voices increases, suggesting the use of two voices with an upper limit of three [3]. Additional information about suggestions could be conveyed through other audio features such as pitch and volume (e.g. transmit confidence through the voice volume). One use case for these audio features could be to subtly communicate other non-autocorrect suggestions (which have a lower prediction confidence) as a whisper, such not to unnecessarily distract the user but make them aware of new word predictions.

Cognition and Suggestion Threshold
Current alternatives provide an earcon and auto-correct only highly probable suggestions. The ability to convey suggestions before the user inserts them provides the opportunity to reduce the confidence threshold and therefore communicate more suggestions. While such an outcome results in more feedback that can be leveraged by users, it could also increase cognitive load, thus impairing the overall text-entry task. Further research needs to be conducted to identify the best compromise between suggestions feedback provided and the cognitive load it imposes to blind users.

Inline with the *additional feedback* scenario, reducing the spoken suggestion threshold could enable solutions to present the blind user with multiple word predictions. To avoid cognitive overload, multiple suggestions could be read aloud in sequence through the secondary audio channel; using faster than normal speech rates [2]. This approach would now give blind users greater access to the de facto top three suggestions, commonly found in smart devices today.

Suggestion Selection
In current solutions, highly probable suggestions are automatically accepted after inputting a *space* or punctuation. While early feedback supports more informed decisions, it also enables interfaces that go beyond auto-correction. In turn allowing users to select a particular word among several suggestions. Therefore, researchers should investigate alternatives to both accept (instead of automatically) and navigate through suggestions. In touchscreens, such interaction mechanisms could rely on gestures.

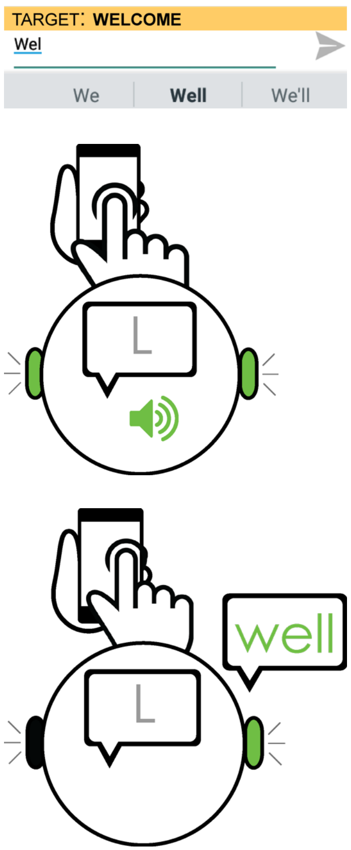

Figure 3. Feedback heard after inserting the character "L" when writing "*welcome*". In VoiceOver, the user listens the "L" and an *earcon* indicating an auto-complete suggestion. In Shout-Out-Suggestions, the user listens the "L" and the actual suggestion word ("*well*") on the right ear.

One such approach could be to use directional swipe gestures [4], to provide immediate navigation and selection from the top three suggestions.

Shout-Out-Suggestions

In order to explore the design space of text-entry auto-correction, we built Shout-Out-Suggestions, an interface that conveys text-entry suggestions using concurrent speech feedback. This system is based on Google's Keyboard and uses the Text-to-Speeches framework [3] to convey the concurrent feedback.

We replicated current mainstream solutions, that communicate the most probable suggestions with either a short 'beep' (earcon) or visual emphasis only, by adding additional speech feedback about highly probable suggestions before they are inserted. This means that Shout-Out-Suggestions reads aloud very confident suggestions (the ones that are automatically corrected after inserting *space* or punctuation), instead of playing the short 'beep' or providing no audio feedback. Currently, the system reads the regular *Explore-by-Touch* feedback in the main speech channel (female voice), while strong suggestions are read aloud in a secondary channel with a male voice, located near the user's right ear, assuming the use of headphones to enhance speech segregation (Figure 3).

We envision a comparison between Shout-Out-Suggestions and mainstream accessibility solutions as the means to determine how being aware of auto-correct/complete suggestions influences users' text-entry performance. However, the design space of auto-correction interfaces outlines different pathways that can be followed in order to explore the benefits of concurrent suggestions. For example, one may measure how reducing the confidence threshold affects user performance and cognitive overload. Moreover, users may be able to assess more suggestions beforehand, instead of exclusively the ones that would be automatically corrected. With such function, one could perform a swipe-right to accept the most probable suggestion; a swipe-up to ask for additional suggestions (read on different spatial locations – e.g. one suggestion at each ear) and afterwards swipe-left or –right to select the intended suggestion.

Conclusions

Currently there is a lack of support for non-visual access to word predictions, these interfaces require continuous monitoring to take advantage of the new suggestions. In this paper we propose the application of concurrent speech to provide access to word predictions for blind people.

Acknowledgements

This project was funded by EPSRC award number DERC EP/M023001/1 (Digital Economy Research Centre); the Portuguese Foundation for Science and Technology, under grants INCENTIVO/EEI/LA0021/2014, SFRH/BD/103935/2014, UID/CEC/50021/2013 and UID/CEC/00408/2013.

References

1. Mark D. Dunlop, and Andrew Crossan. 2000. "Predictive text entry methods for mobile phones." Personal Technologies 4.2-3: 134-143.
2. João Guerreiro and Daniel Gonçalves. 2015. Faster Text-to-Speeches: Enhancing Blind People's Information Scanning with Faster Concurrent Speech. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (ASSETS '15). ACM, NY, USA, 3-11.

3. João Guerreiro and Daniel Gonçalves. 2014. Text-to-speeches: evaluating the perception of concurrent speech by blind people. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility* (ASSETS '14). ACM, NY, USA, 169-176.
4. Shaun K. Kane, Jeff P. Bigham, and Jacob O. Wobbrock. 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 & accessibility* (ASSETS ‘08), ACM, NY, USA, 73–80.
5. Shaun K. Kane et al. 2008. TrueKeys: identifying and correcting typing errors for people with motor impairments. In *Proceedings of the 13th international conference on Intelligent user interfaces* (IUI '08). ACM, NY, USA, 349-352.
6. Per Ola Kristensson and Keith Vertanen. 2014. The inviscid text entry rate and its application as a grand goal for mobile text entry. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (MobileHCI '14). ACM, NY, USA, 335-338.
7. Hugo Nicolau et al. 2014. B#: chord-based correction for multitouch braille input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, NY, USA, 1705-1708.
8. Hugo Nicolau et al. 2015. Typing Performance of Blind Users: An Analysis of Touch Behaviors, Learning Effect, and In-Situ Usage. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (ASSETS '15). ACM, NY, USA, 273-280.
9. João Oliveira et al. 2011. Blind people and mobile touch-based text-entry: acknowledging the need for different flavors. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility* (ASSETS '11). ACM, NY, USA, 179-186.
10. Caleb Southern et al. 2012. An evaluation of BrailleTouch: mobile touchscreen text entry for the visually impaired. In*Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services* (MobileHCI '12). ACM NY, USA, 317-326.