This is the author-prepared version of the manuscript, accepted for publication and to appear in the ACM Proceedings of the 26th Annual Conference on Intelligent User Interfaces (IUI'21), 2021.

TextFlow: Screenless Access to Non-Visual Smart Messaging

Pegah Karimi

Indiana University-Purdue University Indianapolis, pekarimi@iu.edu

Emanuele Plebani

Indiana University-Purdue University Indianapolis, eplebani@iu.edu

Davide Bolchini

Indiana University-Purdue University Indianapolis, dbolchin@iupui.edu

Texting relies on screen-centric prompts designed for sighted users, still posing significant barriers to people who are blind and visually impaired (BVI). Can we re-imagine texting untethered from a visual display? In an interview study, 20 BVI adults shared situations surrounding their texting practices, recurrent topics of conversations, and challenges. Informed by these insights, we introduce *TextFlow*: a mixed-initiative context-aware system that generates entirely auditory message options relevant to the users' location, activity, and time of the day. Users can browse and select suggested aural messages using finger-taps supported by an off-the-shelf finger-worn device, without having to hold or attend to a mobile screen. In an evaluative study, 10 BVI participants successfully interacted with *TextFlow* to browse and send messages in screen-free mode. The experiential response of the users shed light on the importance of bypassing the phone and accessing rapidly controllable messages at their fingertips while preserving privacy and accuracy with respect to speech or screen-based input. We discuss how non-visual access to proactive, contextual messaging can support the blind in a variety of daily scenarios.

CCS CONCEPTS • Information Interfaces and Presentations • Assistive Technologies

Additional Keywords and Phrases: Text entry, Assistive technologies; Intelligent wearable and mobile interfaces; Aural navigation, and Ubiquitous smart environments

ACM Reference Format:

Pegah Karimi, Emanuele Plebani, and Davide Bolchini. 2021. *TextFlow*: Screenless Access to Non-Visual Smart Messaging: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages.

1 INTRODUCTION

Mobile texting mainly relies on on-screen keyboards that display characters visually, posing barriers to people who are blind and visually impaired (BVI) [2, 40]. Even when keyboards have accessible overlays (such as VoiceOver), they remain screen-centric: manipulating text is generally bound to holding a device out at all times and interacting with visual keypads primarily designed for sighted users, or with accessible keyboards that read aloud keys upon

This is the author's manuscript of the article published in final edited form as:

Karimi, P., Plebani, E., & Bolchini, D. (2021). Textflow: Screenless Access to Non-Visual Smart Messaging. IUI '21: 26th International Conference on Intelligent User Interfaces. https://doi.org/10.1145/3397481.3450697

touching [15, 23, 26, 41, 43]. Screen-bound interaction is especially problematic in nomadic contexts, when blind users are on-the-go, have to keep a cane in one hand and the other hand free to touch objects nearby. Users can resort to voice input [6] but are still often required to be bound to a mobile device for dictation and go through several repetitions to overcome recognition problems and ambient noise [5]. Studies have shown that users who are blind tend to prefer not to use speech-based interfaces in public places, due to concerns for privacy and social conspicuousness [1]. Another major issue of current mobile messaging is that, whereas daily routines and recurrent situations often lead users to periodically re-send similar messages, current approaches always require users to initiate an open-ended text composition. As a consequence, there is an undue burden to blind users, especially in light of the severe mechanical and interaction constraints they have to overcome to complete a text. Can we re-imagine entirely non-visual texting in a way that is quiet, proactive and untethered from a screen? To explore this question, the paper makes three main contributions:

- An AI model for generating short text suggestions from a large-scale dataset and driven by key contextual and situational factors such as user current location, activity, and time of the day. Message topics are informed by a study we conducted with 20 BVI users on their needs and practices of texting.
- TextFlow: a mixed-initiative context-aware system that generates entirely auditory message options potentially relevant to the user situation. TextFlow enables users to listen to, browse and send suggested aural messages without holding or attending to a mobile screen, using nimble finger-taps supported by an off-the-shelf finger-worn device. The system operates as a sequential auditory stream of fast-spoken topics and messages that can be browsed and selected for composing a text.
- An evaluative study with 10 blind participants who interacted with TextFlow through 15 tasks focused on daily text messaging. With an average task success rate of 88.6%, all participants were able to interact with TextFlow and send messages in entirely screenless mode. Participants consistently appreciated the positive experience of bypassing screen-centric methods for typing or voice input, especially when they are mobile and need to send short notifications without taking the phone out of their pocket.

2 RELATED WORK

2.1 Eyes-Free Mobile Text Input

Navigating mobile keyboards using touchscreen devices has been widely adopted by blind users. Technologies such as Apple's VoiceOver [28] allow users to touch the screen and get feedback via voice. While voiceover interaction makes mobile keyboards accessible, it requires users to continuously move the finger on the screen to find and select the intended letter. Researchers introduced better typing methods by rearranging the alphabets of mobile keypads. BrailleTap [22] and NavTap [23], for example, eliminate the need to memorize the letter position by grouping 3 or 4 letters as keys. Other technologies combine multi-touch input with audio feedback to enable fast typing. No Look Notes [10] is a multi-touch text entry that arranges letters in pie menus. Users can tap the finger on the screen to switch to the desired pie menu. These approaches enhance the accessibility of typing, but are still screen-centric, thus requiring the blind to hold out a phone at all times to interact with a visual display.

Speech input is now broadly available and enables voice dictation for texting [6]. When used outside, however, besides causing issues when background noise is present [5], speech input also introduces barriers to effective use. For example, it is challenging for the blind to review and edit the dictated text on the screen [6]. Social and privacy boundaries can also be easily broken by voice input, and this is problematic especially when the blind prefers not to draw unwanted attention in public spaces [16,42].

Wearable devices including armbands, hand- and finger-worn devices open additional input channels that can potentially bypass the constant reliance on a reference screen. NailO [27], for example, is a nail-mounted wearable surface that detects various swipe gestures to input emoticon or punctuation without the need to use the screen. However, users still have to rely on the mobile screen due to the difficulty of typing using only the nail surface. Recent work has shown the use of the Myo [37], a hand-worn band that recognizes and uses as input the arm's muscle movement, to control entirely auditory keyboards for the blind without the need for a reference screen [34]. The approach still requires performing quite ample hand gestures that are not as discreet as users would like, and may unnecessarily fatigue the arm.

2.2 Context-Aware Mobile Communication

Research in ubicomp has studied the potential of leveraging contextual dimensions, such as location and time, to facilitate human communication through text messages. *Location* is a common type of contextual cue deployed in ubicomp applications [3]. LAMMS [9], for example, demonstrates the notion of location-based text messaging, enabling people to communicate locally with each other and receive messages related to that area. *Time* is another dimension often used when sending text messages. LATTE [39], for example, is an email-based system that incorporates both time and location to dynamically expand emails with the corresponding temporal and spatial information. Other forms of context-based mobile messaging systems allow users to define the context while sending a text. Jones and Neil [25] introduced a contextual constraints model, in which the user can place constraints on whom to send the message to (i.e. personal dimension), where the message has to be sent (i.e. spatial dimension), and when the message has to be sent (i.e. temporal dimension). Contextual information has also been used to provide feedback about the environment while an individual is on the go and holding the phone for typing [36]. Many people experience walking and typing at the same time they often fail to perform both tasks perfectly [35]. One method for providing feedback about the user's ambient environment is through audio [14, 32]. These systems provide real-time feedback to users allowing them to focus on typing while reducing errors by increasing awareness about the environment. However, the method can create difficulties for BVI users due to receiving auditory information from both typing and their ambient environment.

2.3 Use of Context in Recommender Systems

Context-aware recommender systems use the concept of 'context' to recommend items that are both relevant to the user's preferences and their specific context such as time and location [4, 7]. Magitti [8], for example, is a context-aware mobile recommender system that automatically detects a user's activity based on the user's context and behavior to recommend personal and timely leisure activities associated with a local environment. SoCo [30] is a social network aided recommender system that combines location and time with social information learned from friends with similar tastes to provide highly accurate recommendations. Bouidghaghen et al. [11] introduce a situation-aware recommender system to alleviate the problem of information overload. The system retrieves contextual information related to the user's location and time and provides a mobile user with personalized search results. However, context-aware recommender systems that rely on past experiences (i.e. exploitation) cannot model user's interest evolution because the learned rules will only reflect past user behavior; for this reason, a fraction of recommendations are selected at random or using a heuristic to obtain information about the user and discover better recommendations (i.e. exploration). In [29] authors use a bandit algorithm to exploit well-established advertisements for short-term capitalization as well as exploring less-known cases to find out potential advertisements that can be recommended in the future. Bouneffouf et al. [13] introduce a hybrid-ε-greedy algorithm that takes into account the context to deliver recommendations that are bound to the user's current situation and interests. The algorithm combines the user's situation based on time, location, and social ontologies with an ε-greedy algorithm and content-based filtering techniques.

3 ELICITING MOBILE TEXTING NEEDS AND PRACTICES

We conducted a formative, interview-based (IRB-approved) study to better understand the practice of text messaging adopted by blind and visually impaired individuals. The goal of our user study is twofold: 1) elicit recurrent messaging topics associated to BVI specific needs for mobile text communication; 2) explore the types of daily situations where BVI engage in text-based messaging.

3.1 Participants and Procedure

The study engaged 20 BVI participants (12 females and 8 males). Eight participants identified as totally blind, four legally blind, five had minimal light perception, and three had low vision. Participants' ages ranged from 35-79 (μ=51.7, σ=11.2) years old. 19 participants used iOS devices and one used Android. The average length of mobile usage was 11 years. We conducted the interviews through phone calls or Zoom meetings, depending on the participant's preference. Each interview lasted approximately 1 hour. Interview questions covered three foci: (1) places participants typically visit during daily routines and modes of transportation; (2) situations and circumstances in which they might send a text; (3) recurrent topics of their texts and most recent text messages they sent to others. Interviews were audio-recorded for transcription and analysis. Upon completing the study, each participant received a $30 Amazon gift card via email for their participation.

3.2 Analysis and Results

We performed a thematic analysis of the participants' responses. Overall, six high-level themes emerged that represent different types of text messages participants send on a regular basis: *notifying someone; offering assistance; scheduling/rescheduling plans; coordinating with someone; reminding someone; requesting assistance*.

3.2.1 ***Notifying someone*****.** Participants send different types of notifications according to their immediate situation. One such notification is to let others know that they might *arrive late* or specify their *arrival time*. For instance, P1 commented: "*If I'm running late to work, I may send a message to my boss saying, hey, I'm running late, I'll be there in 20 minutes*." In other instances, participants send a text to let others know they are *on their way*. For example, P7 would say "*The bus is ready to come, I am on my way*" when going to an appointment to meet a friend. Participants mentioned that when they are on their way to a specific location, it is common to inform their close ones that they arrived safely or to *signal their location*. For example, P3 described the type of message they would send when they are outside: "*I send a text to my husband and say I got here okay or I'll tell him that I got to the bus stop safely*." In some cases, users might *not be able to talk* due to privacy concerns and therefore might send a text stating that they will call back later. P2 gave an example of when they are waiting at the bus stop: "*Let me get back with you later, I am in the middle of commute*." Another typical notification that participants mentioned was to let others know that their initial *plan has changed*. For instance, P6 described their experience of sending a text while they were on their way to work, stating: "*while waiting at the bus stop, I was notified that our school trip that was supposed to go to Chicago got canceled, and so I had to text to colleagues that the plan for that day has changed*." On some occasions, participants specified that they had to send a text to emphasize that they are *waiting*. P11 gave an example of a text message they would send when they are waiting at the bus stop: "*Are you almost here? I am waiting for a long time*." Other, less frequent topics included letting others know they are in a *meeting* and on their way to go *shopping*.

3.2.2 ***Requesting assistance***. Participants may send a text asking for help due to access barriers, in particular when they need to *get to a location*. Most participants mentioned that when they reach a destination or they are inside a building, it is often difficult for them to find the entrance or exit. In these situations, they send a text to the person they planned to meet to look for them and guide them to the place. For instance, P2 explained an experience, stating: "*One time I was in a building, and I couldn't find my way. I was on my way, but I could not find my way back because I did not have my GPS on. So, I sent a text, and I said I got lost, and then I asked for help*." P8 described the type of text message they would send if they could not find the entrance: "*Come outside and get me*." Another type of request that participants mentioned is to ask someone to be *picked up*. P15 gave an example of the type of text message they typically send while walking to reach the bus stop: "*pick me up in five minutes*." In a few cases, participants mentioned they might ask someone to *bring food* when they are at work.

3.2.3 ***Scheduling/rescheduling plans***. When on the go or at work, it is common among BVI users to send texts to *make or change plans* for both formal and informal commitments. For instance, P5 provided an example of a text they would send to a friend when they are at work: "*where would you like to go for dinner?*" The same participant gave an example of a text sent on their way to meet a friend: "*I'll see you at 2 for our appointment.*" Users may also need to reschedule existing plans. P10 gave an example of a message they would send when at work: "*9'o clock meeting is canceled; meeting is at 10*." Participants also described sending a text to friends or family members about a change of plans. P16 explained a text sent to a friend when they realized they couldn't meet as planned: "*I told him I had the flu and I wouldn't be able to make it. And then I said could we reschedule?*" On other occasions, users might send a text to schedule a *phone call*. P3 gave an example of a text they would send on their way to work: "*when can we set up a call?*" Thus, BVIs need to be able to communicate their plans quickly while on the go or at work.

3.2.4 ***Coordinating with someone***. When on the go or in public places, users often need to text others to coordinate about location, time, and number of people they expect for a gathering. For instance, P1 gave an example of a text message they sent when meeting a friend at a restaurant: "*are you here at the restaurant yet?*" The participant is trying to *check-in location* to find out whether their friend has arrived at the restaurant or not. P9 gave an example of coordinating a group meeting: "*if we're supposed to be like a group of people, I can't remember everybody who's supposed to be there. So, I will send a text and say: who's going to be for dinner?*" The participant is *checking-in guests*, as they may not have instant access to a phone to verify the number of friends who confirmed coming for dinner. When asked about the context of the messages sent while returning from work, P15 stated: "*Since everybody is on a different schedule, we mostly text and say what time you're going to be home for dinner*."

3.2.5 ***Reminding someone***. Unlike sighted users, BVIs experience more difficulties accessing their calendar to check their appointments. In these situations, they often send each other reminders to confirm the upcoming *appointment* or *meeting*. P17 provided an example of a message sent to a friend while walking to the bus stop: "*See you tonight at dinner.*" When at work, participants might send reminders related to meetings or ask for confirmation. P10 provided an example of a text message they might send to a colleague: "*is the meeting today at 9:30?*" On other occasions, users might send a reminder related to an upcoming meeting to verify if others will be there. P13 commented: "*our meeting is today at 11 o'clock, please let me know if you can attend*." Therefore, users send reminders to confirm meetings and/or appointments both for themselves and others.

3.2.6 ***Offering assistance***. When on the way, users offer assistance to others regarding daily activities. For instance, P5 stated: "*If I'm in a vehicle, I have [to] kind of hover a lot to send a text to my son and say, do you want something specific from [the] drugstore?*" The participant emphasizes the amount of time it can take to send a text to someone and ask a question related to *shopping*. Users might also be in a specific store and send a text to ask if the recipient needs anything from that store. P15 gave an example of this type of text message: "*I'm here at the dollar store. Do you need anything?"* Both participants in these examples are offering help to avoid an extra trip to go to a store. Other types of offering assistance include *picking up* someone on the way or helping to *get to a location*. P15 explained their experience when they have an appointment with a friend, stating: "*I might text my friend and say, do I need to meet you at the bus stop to help guide you in?*" Therefore, BVIs frequently offer assistance to each other and to members of their household while outside the home.

In total, 17 specific topics emerged: *arrive late, arrival time, on my way, get to a location, check-in location, check-in guest, signal my location, appointment, meeting, shopping, making plan, changing plan, pick up, waiting, phone call, not able to talk,* and *bring food*. In addition, we identified specific situations in which some of these topics emerged frequently. Figure 1 shows the occurrence of topics that appeared three or more times for four different situations.

3.3 Generating Messages from Elicited Topics

For each topic, we generated a list of short candidate messages. For instance, if the topic is 'arrive late', a potential message would be 'I will arrive a few minutes late.' The dataset we used to generate messages is called Weibo [44] and it includes short text conversations from a Chinese microblog service. Weibo has more than 4.4 million samples, divided into posts and comments associated with each post. We selected this publicly available dataset because it contains a very large number of instances of the specific textual genre we are seeking to model, rather than broadcast, one-to-many communications typical of social media platforms. Samples are translated from Chinese to English using a tool called Youdao [45]. We performed data cleaning by removing redundant words such as 'come on' or 'oh' and discarding generic topics such as politics or advertisement. Given a topic, we used a pre-trained language model called RoBERTa [31] to retrieve related sentences from the dataset. The model has been trained on the Multi-Genre Natural Language Inference (MultiNLI) corpus [38] with pairs of sentences annotated as entailed, contradictory or neutral. For each topic, a reference text sample was selected from the user study. If the topic belongs to more than one category of text message, a sample was selected from each category. For example, if the topic 'appointment' belongs to both *reminding someone* and *coordinating with someone*, we selected a reference sample from both categories. We compared the reference sample with all the sentences in the dataset using pairwise sentence classification. We then retrieved the sentences classified as "entailment", i.e., the meaning of the query sentence is implied by the retrieved sentences, and "contradiction", i.e., the retrieved sentences imply the negation of the query sentence. The model returns a probability for each class as a function of the sentence pair:

$$(p_c, p_n, p_e) = f(s_1, s_2)$$

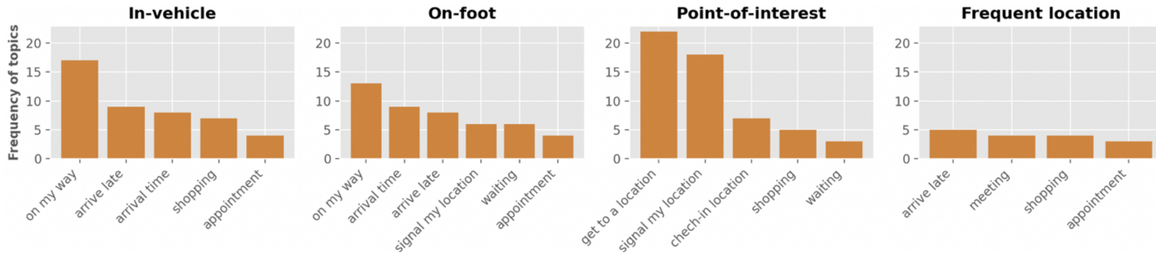

Figure 1: The frequency of topics identified by participants as relevant to four different situations: in-vehicle, traveling on-foot, nearby or inside points of interest, and when going to a frequently visited location (e.g. workplace).

where s_1 and s_2 are the two compared sentences and p_c, p_n, p_e are the probability of contradiction, neutral and entailment respectively. We ranked the results by p_e and p_c, i.e., the entailment and contradiction confidence scores. For entailment, we considered sentences with p_e higher than 0.2 (20%), whereas for contradiction we considered only sentences with p_c greater than 0.1 (10%). The choice was motivated by the fact that the entailment class provided more diverse and related sentences for the given query, whereas the retrieved sentences for the contradiction class were limited to a few samples. Low-scoring sentences, i.e., below 20% for entailment and 10% for contradiction, were not considered as they were often unrelated to the query.

Starting from the first retrieved recommendation, we selected those that included the topic and a single instance of contextual information based on factors such as time, location, audience (e.g., with someone), purpose (e.g., for dinner), or method (e.g., with Uber). For instance, the sentence "I've been waiting for a long time" consists of our topic (waiting) and information that concerns time. Recommendations that were not relevant to our topic or that had the same contextual information as the selected sentences are skipped. The third column in Table 1 shows the selected recommendations for the topic 'waiting'. Selected recommendations were then refined to represent a short message (fourth column of Table 1). The steps for refining sentences are the following: 1) we added or modified the subject to be in the first person; 2) recommendations that included more than one sentence were refined by removing sentences that did not contain the keyword related to the topic; 3) selected sentences that referred to specific locations or times were replaced by either the user's current situation or a plan elicited from the calendar.

4 TEXTFLOW: MIXED-INITIATIVE AND SCREENLESS TEXTING

Based on the results of our study, we developed TextFlow (https://github.com/Banus/Textflow), a mixed-initiative text messaging system that enables BVI users to receive and interact with suggested messages relevant to their immediate situation, while at the same time bypassing the reliance on a visual screen. The system is made of three parts: a User Model, a Reasoning Model, and a Task Model (Figure 2).

4.1 The User Model

The User Model captures key aspects of the user's daily routine. The model adapts to the specific needs of BVI users for manipulating a text message and contains four types of data: the user's context, the user's situation, personal-related commitments, and a dynamic profile. To formally define the elements of the User Model, inspired by [12], the **user's context** C is defined as the ontology of location, time, and activity and we represent it as $C = (O_{location}, O_{time}, O_{activity})$. An instance of the user's context is a **user's situation**, defined as:

Table 1: Generating messages for the topic 'waiting'

Reference Sample	Class Type	Selected Recommendations	Messages
I am still **waiting**	Entailment	I am waiting for you to come	I am waiting for you to come
		I've been waiting for a long time! where have you been?	I've been waiting for a long time!
		I am waiting at the dentist	I am waiting at *
		I am still waiting for someone	I am still waiting for someone
		Wait a few minutes!	Wait a few minutes!
		Waiting outside the maternity ward	I am waiting outside the *
		Waiting for dinner!	I am waiting for **
		Still waiting for my ride	I am still waiting for my ride
		Wait a few seconds	Wait a few seconds
	Contradiction	There is no need to wait for this photo	There is no need to wait for me
		I am not waiting anymore	I am not waiting anymore

* current location based on the map. ** breakfast, lunch, or dinner based on the time.

$$S = (O_{location}x_i, O_{time}x_j, O_{activity}x_k)$$

where x_i is the location instance, x_j is the time instance, and x_k is the activity instance. We define two types of locations, retrievable from Google Maps: known locations and unknown locations. Known locations require either the user to manually label work, home, and school locations or the system to recognize a point of interest (e.g., train stations and restaurants). Locations that are not recognized as known are categorized as unknown locations. The location is represented as a longitude, latitude pair (log, lat).

To detect points of interest and labeled locations, we set a threshold to 0.1 miles on the distance from the user's current location. Time is divided into three periods of morning (5am-12pm), afternoon (12pm-5pm), and evening (5pm-9pm). The user activity is classified into three categories of *still*, *walking*, and *in-vehicle*. A confidence score is assigned to each activity and the one with the maximum score is detected as the user's current activity.

Personal-related commitments (PRCs) signal the user's recurrent and non-recurrent routines, and they are elicited from the personal calendar available on the mobile device and the data from the map. The calendar includes a list of commitments for each day with title, location, start and end time. The location must belong to one of the known locations and its coordinates are retrieved from the map. By combining the user's PRCs and the current situation we generate a **dynamic profile** that integrates real-time data from the user.

When the user initiates an interaction for sending a message, the system suggests topics and messages based on the User Model data. Topics are the entry points of TextFlow, and their order is determined by the user's situation, which is constituted by the type of activity, location and time (see Figure 1). When the activity is in-vehicle the system re-ranks topics for an in-vehicle situation regardless of the location. When the activity is classified as still or walking and the location is unknown, the system re-ranks topics for an on-foot situation. If the location is recognized as known, the system re-ranks topics either by frequent location or point of interest. Each topic corresponds to a list of messages, such that some of the messages depend on the location, time of the day, and PRC (see Table 1).

4.2 The Reasoning Model

The Reasoning Model provides rules to detect inconsistencies and infer new data [33]. These rules lead to 'high-level critical situations' [12], a class of situations in which the system initiates the interaction and selects topics bound to a specific rule. Based on the situations and messages learned from the initial study, we define four generic rules to determine when the system initiates the interaction and what topics are presented to the user.

Missing formal commitments: BVIs may miss important commitments due to limited access to time and traffic information. For each formal commitment, we estimate the arrival time based on the speed of the user. While users

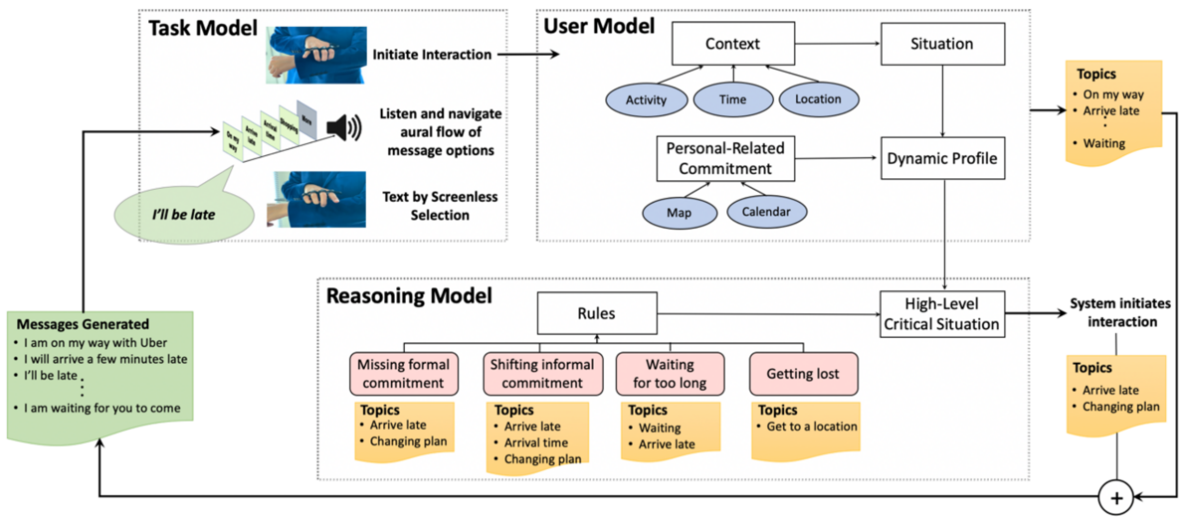

Figure 2: Three parts compose TextFlow: 1) the User Model leverages user's situational factors to generate a dynamic user profile; 2) the Reasoning Model operates on four rules based on the topics elicited from the initial study; 3) To support the user interaction, the Task Model enables the blind to listen to and select aural prompts in screenless mode via an existing finger-worn device (TapStrap).

are en route, starting 15 minutes before their scheduled commitment, the distance to the destination is recorded every 5 minutes and the average speed is computed as the distance covered in the last interval. The arrival time is estimated by extrapolating the average speed to the destination, and the rule is activated when the arrival time exceeds the time of commitment by 10 minutes. An example is when a person should be at work before 9 am, and the current situation is $S = (in - vehicle, 8{:}45am, \forall location)$. If the user's distance from work is 26 miles and 5 min later is 21 miles, the rule will be activated. The selected topics for this rule are 'arrive late' and 'changing plan'.

Shifting informal commitments: When BVIs are busy with one activity, they may forget to inform others about informal commitments, such as meeting a friend at Starbucks or going home. Shifting informal commitments are elicited from PRC based on the ending time of their formal commitments. In this study, the rule is activated when a person is not in the expected location or en route 15 minutes after the end of the commitment. An example is when a person leaves their place of work at $5{:}30\ pm$ and the situation is $S = (still, 5{:}45pm, work)$. Chosen topics for this rule are 'arrive late', 'arrival time', and 'changing plan'.

Waiting for too long: When BVIs wait in a specific place they may not realize the length of the time they spent there. Waiting for too long is defined as when the user stays still for at least 30 minutes and the location is labeled as 'station' or 'unknown'. An example is when a person is in the bus station and the situation is $S = (still, 2{:}15pm, bus - station)$ and after 30 minutes it is $S = (still, 2{:}45pm, bus - station)$. The selected topics for this rule are 'waiting' and 'arrive late'.

Getting lost: Finding the exact location of known places is not always feasible. While GPS is a possible solution, it may not always be accurate enough to support BVI users. In our study, getting lost occurs when a person is walking for at least 15 minutes in the same general area outdoors. An example is when a person has an appointment with their friend in a restaurant and after reaching the place the situation is: $S = (walking, 2pm, restaurant)$ and after 15 minutes is: $S = (walking, 2{:}15pm, restaurant)$. The selected topic for this rule is 'get to a location'.

4.3 The Task Model

Task models incorporate the logical steps associated with activities that an individual must perform to achieve their goal [21]. With TextFlow the user is either notified proactively of relevant topics based on the situation or the user can initiate the texting action, and this input triggers the User Model. A list of relevant topics and messages are offered to the user for interaction and topics are ordered based on the frequencies obtained from the initial user study. Topics with high frequency (shown in Figure 1) are read individually, whereas topics with a frequency of 2 or less are collected in a 'More' category, which plays on-demand after the last topic (Figure 3-c). These outputs are then fed into the Task Model.

The interaction is supported by a wearable device called TapStrap [24] (Figure 3-b) which uses finger movements to perform various operations such as selection or deletion to send a message via a messaging app (Figure 3-a). The system supports two different interaction modalities: self-disclosing flow and topic-by-topic browsing. Topic-by-topic browsing allows the user to navigate the system's suggestions with explicit forward and back commands, whereas self-disclosing flow reads the system's suggestions sequentially. We describe the steps that the user has to perform when interacting with each of the modalities:

Topic-by-topic browsing: In this interaction modality, users tap their index finger to listen to topics individually. They can navigate backward through suggestions by using the middle finger, tap their thumb to select a topic and then browse the message list in the same way. Users have the option to tap their ring finger to cancel the selection. When the system initiates the interaction, users can either tap their pinky finger to ignore the system's suggestions or continue browsing topics one by one.

Self-disclosing flow: To mitigate the need for using multiple fingers, we introduce an alternative TextFlow layout in which the system's suggestions are self-play, i.e., read out sequentially. We define three different dwell times, that is, the pause between the system suggestions. The first is node-to-node dwell time (150ms), which is the pause between topics or the pause between messages. The second is the dwell time between topics and messages (1s). The third is the flow-to-flow dwell time (1s), that is, upon hearing the intended topic, the user can tap their thumb to make a selection and if no selection is made the system will repeat the topics from the start after the dwell time. The user can tap their index finger to cancel the selection. Upon selecting a topic, the system will transition to messages. Similarly, the user can tap their thumb or index finger to make or cancel a message selection respectively.

We implemented and deployed TextFlow on a Nexus 6x phone with Android 10.0 and Android SDK v.29.0; the TapStrap device was interfaced using Android Tap SDK 0.3.3, while the aural rendering of the message options was realized through the default Android text-to-speech (TTS) engine [18]. To feed contextual data, we used Google Location and Context services [20] for the points of interest, the Google Calendar API [19] for personal-related commitments (PRC), and the Android Activity Recognition API [17] for data on the current user activity.

5 STUDY WITH BLIND PARTICIPANTS

We conducted an evaluative, in-person user study (IRB-approved) of TextFlow to scrutinize the usability and the user experience with the system for first-time users.

5.1 Participants, Setting and Procedure

We recruited 10 BVI individuals (6 females and 4 males) in the Indianapolis urban area. Six of them already participated in the initial study (Section 3). The participants' ages ranged from 33 to 68 years old (μ=47.7, σ=10.37). Eight used iOS and two used Android. The length of mobile usage ranged from 6 to 18 years (μ=11.8, σ=4.66). Five

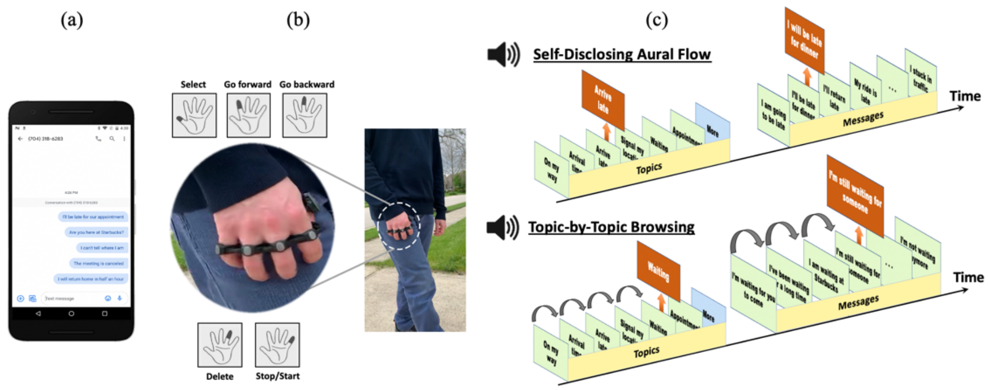

Figure 3: Detailed view of the Task Model. Users browse topics either as self-disclosing flow or by scanning them one-by-one (c); Messages are selected through an off-the-shelf device that supports five tapping gestures (b) and sent via a messaging app (a).

participants were identified as totally blind, three legally blind, and two had retinopathy of prematurity (ROP). The study was conducted in a conference room, with a table and chairs. Each participant remained standing while interacting with our system, and could rest, sit or take a break at any time. Based on our first study, we defined three scenarios to capture specific situations where a blind person might need to send a message: (1) On my way to meet a friend at Starbucks at 2 pm; (2) On my way to work riding with a friend; (3) Leaving the house at 1 pm to meet a friend for lunch. For each scenario, we asked participants to complete five tasks, such as "Letting a friend know I'll be late", or "Reminding a colleague about an upcoming meeting". Our study utilized a within-subject design where each participant completed 15 different tasks distributed evenly over three scenarios. In the first and third scenarios, four tasks required the user to initiate the interaction with TextFlow, whereas one task was associated with a rule in which the system initiated text suggestions. In the second scenario, two tasks were associated with rules and three tasks required the user to start messaging. When the user started messaging, the default interaction modality was topic-by-topic browsing. However, when the system initiated the interaction, the default modality was self-disclosing flow. The order of the three scenarios was counterbalanced to account for any ordering effects. Each scenario was executed indoors and described verbally to each participant to help them understand both the situation and the type of message they could send for each task. Participants were first introduced to the TapStrap, the hand position required while tapping with the device, and the role of each finger in sending a text message. In the first phase, we explained the purpose of each topic with one or two examples of text messages associated with the topic and introduced the two interaction modalities—topic-by-topic browsing and self-disclosing flow. In the second phase, participants were asked to practice and send a text message using the two modalities. Participants were provided with feedback and assistance if needed during the second training phase.

After training, participants were required to start the tasks associated with the first scenario, and each scenario was read aloud by the researcher. During the interaction, users could tap on any part of the body that was convenient for them. Once the user sent a message, they were asked to determine whether the selected message was relevant to the context or not. Each scenario was video recorded, followed by a five-minute break. After the last session, we asked participants interview questions. The questions focused on situations in which the system could be useful, advantages and disadvantages of the system versus accessible keyboards and dictation, and comparison between the two interaction modalities. Answers to interview questions were audio recorded. For approximately two hours of participation, each participant received a $60 Amazon gift card.

5.2 Results

5.2.1 ***User Performance analysis***. To assess efficiency, we tracked the time participants took to send a message and computed the average time (μ) and standard deviation (σ) for task completion across all participants. To assess effectiveness, we defined and analyzed key measures of the task success for each task and participant.

Time on task. For the 11 tasks associated with topic-by-topic browsing, the time spent depended on where the target topic and message were in the auditory flow. Out of 11 tasks, six had both the target topic and the message among the top in the list and yielded the highest performance ($\mu = 17.42$; $\sigma = 9.43$), between 10.5 to 17.5 seconds. For three out of 11 tasks, either the target topic or the message was close to the end of the list: it took longer to identify and select the message ($\mu = 33.97$; $\sigma = 11.3$), between 29 to 35 seconds. Finally, for the two tasks where both the target topic and the message were at the end of the list, users selected the target message between 42 to 44.5 seconds ($\mu = 43.7$; $\sigma = 6.48$). The narrow range of standard deviations across all tasks (between 5.58 and 13.48 seconds) indicates that the performance was quite uniform across participants, with few exceptions. The exceptions are due to two main factors: the number of times that the user had to go backward (18 times in 11 tasks) and the number of times they had to cancel the selected topic (15 times in 11 tasks). The former occurred when the user browsed system suggestions very rapidly, inadvertently skipping the target topic or message. The latter happened when the user could not find the intended message within the first topic and had to get back to the list to select another topic. For the 4 tasks associated with the self-disclosing flow, the average time ($\mu = 35.6$) ranged between 28.7 to 46.3 seconds, while the standard deviation ($\sigma = 23.52$) sat between 18.30 to 27.92 seconds. Two factors affected the time spent on these tasks. The first factor was the number of times a selected topic or message was canceled (30 times in 4 tasks): when participants failed to select the target topic or message within one second, the system selected the next suggestion. The second factor was that 3 out of 10 participants preferred to listen to the system's suggestions for one or two loops to memorize choices and avoid canceling a selection.

Task success. For each task, we identified one or multiple target messages, defined as the messages that were designed and generated as most pertinent to a given situation. To determine task success, we defined the following criteria: *Success* (1): the participant selected the target message and confirmed it was relevant to the situation; *Partial Success* (0.5) captures two cases: when the message was off target but perceived as relevant; or when the message was on target, but the participant found it not relevant; *Failure* (0): the message was off target and deemed not relevant by the participant. To compute the task success, we normalized the value assigned to each task over the number of participants. The average task success rate across the three scenarios is 88.66% (see Figure 4). On average, 4 out of 15 tasks had a 100% success rate across all 10 participants. In the remaining 11 tasks, an average of 1.6 participants per task had partial success and an average of 0.72 failed. We observed that when users had to continuously search by either browsing over more topics or canceling their initial choices, they forgot the details of the task and ended up selecting a message that was either outside the target message or perceived as not relevant.

5.2.2 ***Analysis of Experiential Themes.*** The open coding thematic analysis of each participant's responses revealed that participants had a very positive experience with TextFlow and offered important insights, presented below:

Bypassing the phone: A theme that emerged from the participants' reflection on their experience is that TextFlow significantly reduces interaction barriers to access everyday texting in entirely auditory mode, especially when traveling. For example, P1 articulates how the phone-free approach of TextFlow can help him by saying: "*The system helps you send text messages without using your hands for the phone. And so, you could use your cane and fully*

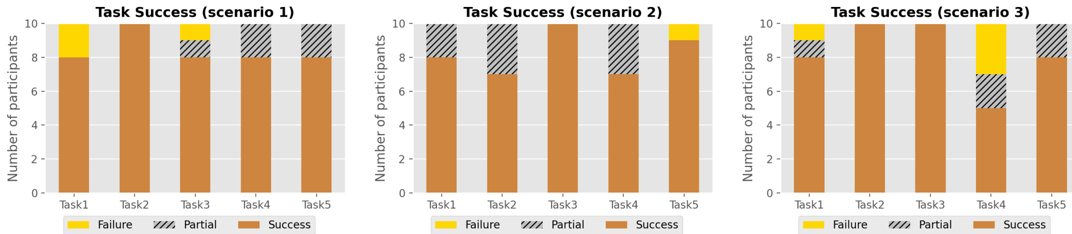

Figure 3: All participants successfully sent screenless messages. In five tasks, two participants on average selected a message off-target; in nine tasks, two participants on average did not perceive the message as relevant to the situation.

concentrate on traveling as opposed to texting and traveling at the same time." Another participant, P6, remarked: "*It can be very useful when traveling and you have a whole lot of items with you so that you do not have to remove your cell phone and run the risk of losing it or having it stolen*." P2 expressed a similar sentiment of comfort when able to rapidly access recurrent and ready-to-use messages, saying: "*I feel secure knowing that I got some quick access that I can contact my friend[s], to be able to say I am here, I am lost, or send my location*."

When reflecting on daily situations, participants appreciated how by initiating and suggesting message options relevant to the context, TextFlow partially relieves blind users from having to constantly monitor their surroundings. P7 described an exemplary scenario, stating: "*If you're on a bus, you're trying to listen to when it's moving, or when it stops and what's the next stop. So, if the system could pull one or two topics up automatically based on the situation, then you don't have to be worried about the time or what is going on*."

Privacy, efficiency, and accuracy: When reflecting on how TextFlow compares to dictation, users pointed to privacy, efficiency, and accuracy as key advantages with TextFlow. P4 stated: "*With dictation, I have to speak…everybody else can hear what I'm saying. But, if I was using this system and I had earphones on, nobody else can hear. So, there's privacy, which is a big deal for blind people*." "*You might end up saying something two or three times, which makes Siri less efficient. Whereas with this, you can get to your message in a much shorter time*." [P9] When comparing their use of TextFlow to accessible keyboards (e.g., VoiceOver), P6 shared: "*The thing about the keyboard is that a lot of times when you select the wrong letter, you have to delete and go back, whereas with this it's not like you're doing individual letters. You may send the wrong message, but you can go back and fix that*."

Control preferred over self-disclosing flow: Participants expressed that browsing the messages one by one gives them more time to think before moving to the next item, whereas the self-disclosing flow requires heightened attention and more practice. P6 commented: "*when you browse, you can control the way it repeats them for you. The one that reads topics it takes more time before it becomes visible for most people and you have to memorize them first*." Similarly, P3 stated: "*the browsing work[s] better for me because you can pick the speed at which you're going.*" Other participants indicated that the self-disclosing flow has the potential to better support multi-tasking. P2 stated: "*If I'm in the middle of doing something […], I think the one that reads suggestions is better. I think that would be quicker. The browsing is better when you're calm, and not rushed*." "*If you're walking using a cane, it'll be more convenient to use the one that the system reads the suggestions to you and you have to use only your thumb*." [P1]

Envisioned use in additional scenarios: When reflecting on additional circumstances in which TextFlow could be useful, participants mentioned: 'finding a place', 'standing on public transportation', 'being in crowded places', and 'being in a meeting'. P7 commented: "*Even though I can get to the building, but I have a hard time finding a way to get in or to actually even get back out of the place. Because GPS works to get me to the spot, but it won't tell me where the door itself is*." "*if I got lost on the wrong street, I could text somebody to help me out faster, get me in the* *right direction.*" [P8] Another setting mentioned is public transportation, when it is difficult to type. P3 stated: "*It is useful like if you're standing in a train station or on the bus where your hands are full, but you could use the one hand to send a quick message whereas you wouldn't be able to type*." P10 expressed that the system would be useful when standing in crowded places, stating: "*when I'm in a crowd of people, noise and like say, I'm standing at the door of the market.*" Participants also commented on the benefit of the system when they are in a meeting: "*When I am in a meeting, I can inform my family or friends without leaving the meeting or interrupting others.*" [P2]

6 DISCUSSION AND FUTURE WORK

Users who are blind or visually impaired face an additional overhead of mechanical constraints and cognitive effort to access basic mobile services such as texting. The problem has two sources. On the one hand, users have to overcome a screen-centric visual paradigm to interact with entirely aural prompts; on the other hand, current text entry methods require users to initiate text composition from scratch even for frequently sent messages. To address this problem, TextFlow uniquely combines in one solution three interactive approaches: (1) AI-driven, mixed-initiative interfaces that generate suggested, situationally relevant content for candidate messages; the specific intelligent components of our system include the recognition of the user's activity (e.g. in-vehicle, still, walking), which in turn influences the order of the message topics presented to the user in the auditory flow. (2) Entirely auditory, accessible interfaces that leverage the BVI's auditory bandwidth to attend to and rapidly process aural prompts; (3) Screenless input techniques that leverage finger-worn devices to discreetly control, browse and manipulate text, thus bypassing the reliance on a visual display. Overall, our approach provides the basic foundations to generate accessible auditory messages based on a variety of contextual cues. Although our system instantiates specific rules that are empirically informed from the results of a user study, the three conceptual components of the system (the user model, the reasoning model, and the task model) are generic enough to cater to a variety of message types and contextual dimensions. The results of our evaluative study show that the overall experiential response of participants toward TextFlow is positive. Participants were able to interact fluidly with the TextFlow system and send messages quickly and discreetly. Users emphasized that the system bypasses the limitations of existing technologies in terms of reliability, time efficiency, and privacy, in particular when users are mobile and need to send a quick text message while focusing on their surroundings. The feedback from the participants has also pointed to relevant limitations and directions for future work: (1) The paradigm of the messages is still quite limited and does not cover the variety of messages that the user might like to send. For instance, participants mentioned the desire to add important details to a message, such as: 'I am lost, I have a red dress'. It is in fact critical for a blind person to communicate visual signals to be found by friends in public spaces. (2) A second limitation is the extent of personalization. Participants expressed the desire to access the most frequently used topics at beginning of the aural stream, and to add custom messages for rapid re-use. We are working on extending the paradigm of messages supported and on the order of the auditory flow of messages based on the user's history. A major challenge is to strike a balance between message options suggested by the system and the complexity of text editing when manipulating strings in screenless, aural environments.

ACKNOWLEDGMENTS

This research is based on work supported by the National Science Foundation under Grant IIS #1909845. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the NSF. We thank all study participants and the staff at BOSMA Enterprises.

REFERENCES

[1] Ali Abdolrahmani, Ravi Kuber, and Stacy M Branham. 2018. " Siri Talks at You" An Empirical Investigation of Voice-Activated Personal Assistant (VAPA) Usage by Individuals Who Are Blind. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. 249–258.

[2] Ali Abdolrahmani, Ravi Kuber, and Amy Hurst. 2016. An empirical investigation of the situationally-induced impair- ments experienced by blind mobile device users. In Proceedings of the 13th Web for All Conference. 1–8.

[3] Gregory D Abowd, Agathe Battestini, and Thomas O'Connell. 2002. The Location Service: A framework for handling multiple location sensing technologies. College of Computing and GVU Center, Georgia Institute for Technology, Atlanta,. Georgia, USA (2002).

[4] Gediminas Adomavicius and Alexander Tuzhilin. 2011. Context-aware recommender systems. In Recommender systems handbook. Springer, 217–253.

[5] Tousif Ahmed, Roberto Hoyle, Kay Connelly, David Crandall, and Apu Kapadia. 2015. Privacy concerns and behaviors of people with visual impairments. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 3523–3532.

[6] Shiri Azenkot and Nicole B Lee. 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. 1–8.

[7] Linas Baltrunas, Bernd Ludwig, Stefan Peer, and Francesco Ricci. 2012. Context relevance assessment and exploitation in mobile recommender systems. Personal and Ubiquitous Computing 16, 5 (2012), 507–526.

[8] Victoria Bellotti, Bo Begole, Ed H Chi, Nicolas Ducheneaut, Ji Fang, Ellen Isaacs, Tracy King, Mark W Newman, Kurt Partridge, Bob Price, et al. 2008. Activity-based serendipitous recommendations with the Magitti mobile leisure guide. In Proceedings of the sigchi conference on human factors in computing systems. 1157–1166.

[9] Jan Blom, Anu Kankainen, Tomi Kankainen, and Sauli Tiitta. 2003. Location-aware multiuser messaging: exploring the evolution of mobile text-based communication services. Helsinki Institute for Information Technology Technical Report 2003 2 (2003).

[10] Matthew N Bonner, Jeremy T Brudvik, Gregory D Abowd, and W Keith Edwards. 2010. No-look notes: accessible eyes-free multi-touch text entry. In International Conference on Pervasive Computing. Springer, 409–426.

[11] Ourdia Bouidghaghen, Lynda Tamine-Lechani, and Mohand Boughanem. 2009. Dynamically personalizing search results for mobile users. In International Conference on Flexible Query Answering Systems. Springer, 99–110.

[12] Djallel Bouneffouf, Amel Bouzeghoub, and Alda Lopes Gançarski. 2012. A contextual-bandit algorithm for mobile context-aware recommender system. In International Conference on Neural Information Processing. Springer, 324–331.

[13] DjallelBouneffouf,AmelBouzeghoub,andAldaLopesGançarski.2012.Followingtheuser'sinterestsinmobilecontext- aware recommender systems: The hybrid-E-greedy algorithm. In 2012 26th International Conference on Advanced Information Networking and Applications Workshops. IEEE, 657–662.

[14] StephenBrewster,JoannaLumsden,MarekBell,MalcolmHall,andStuartTasker.2003.Multimodal'eyes-free'interaction techniques for wearable devices. In Proceedings of the SIGCHI conference on Human factors in computing systems. 473– 480.

[15] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. 2014. Designing a text entry multimodal keypad for blind users of touchscreen mobile phones. In Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility. 131–136.

[16] Benjamin R Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. " What can i help you with?" infrequent users' experiences of intelligent personal assistants. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services. 1–12.

[17] Android Documentation. 2020. Activity Recognition API. https://developers.google.com/location-context/activity- recognition

[18] Android Documentation. 2020. android.speech.tts Reference page. https://www.tapwithus.com

[19] Android Documentation. 2020. Google Calendar API. https://developers.google.com/calendar

[20] Android Documentation. 2020. Location and Context APIs. https://developers.google.com/location-context

[21] W Keith Edwards and Rebecca E Grinter. 2001. At home with ubiquitous computing: Seven challenges. In International conference on ubiquitous computing. Springer, 256–272.

[22] Tiago Guerreiro, Paulo Lagoá, Pedro Santana, Daniel Gonçalves, and Joaquim Jorge. 2008. NavTap and BrailleTap: Non-visual texting interfaces. In Proc. Rehabilitation Engineering and Assistive Technology Society of North America Conference (Resna).

[23] Tiago Guerreiro, Hugo Nicolau, Joaquim Jorge, and Daniel Gonçalves. 2009. NavTap: a long term study with excluded blind users. In Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility. 99–106.

[24] Tapstrap Inc. 2016. Tap Strap 2. https://www.tapwithus.com

[25] Simon Jones and Eamonn O'Neill. 2009. Context-aware messaging: how personal, spatial and temporal constraints affect text-based communication. In Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia. 1–10.

[26] Shaun K Kane, Jeffrey 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 and accessibility. 73–80.

[27] Hsin-Liu Kao, Artem Dementyev, Joseph A Paradiso, and Chris Schmandt. 2015. NailO: fingernails as an input surface. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 3015–3018.

[28] BarbaraLeporini,MariaClaudiaBuzzi,andMarinaBuzzi.2012.InteractingwithmobiledevicesviaVoiceOver:usability and accessibility issues. In Proceedings of the 24th Australian Computer-Human Interaction Conference. 339–348.

[29] Wei Li, Xuerui Wang, Ruofei Zhang, Ying Cui, Jianchang Mao, and Rong Jin. 2010. Exploitation and exploration in a performance based contextual advertising system. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 27-36).

[30] Xin Liu and Karl Aberer. 2013. SoCo: a social network aided context-aware recommender system. In Proceedings of the 22nd international conference on World Wide Web. 781–802.

[31] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692 (2019).

[32] Joanna Lumsden and Andrew Gammell. 2004. Mobile note taking: Investigating the efficacy of mobile text entry. In International Conference on Mobile Human-Computer Interaction. Springer, 156–167.

[33] Estefanía Martín, Pablo A Haya, and Rosa M Carro. 2013. Adaptation technologies to support daily living for all. In User modeling and adaptation for daily routines. Springer, 1–21.

[34] Reeti Mathur, Aishwarya Sheth, Parimal Vyas, and Davide Bolchini. 2019. Typing Slowly but Screen-Free: Exploring Navigation over Entirely Auditory Keyboards. In The 21st International ACM SIGACCESS Conference on Computers and Accessibility. 427–439.

[35] David E Meyer and David E Kieras. 1997. A computational theory of executive cognitive processes and multiple-task performance: Part I. Basic mechanisms. Psychological review 104, 1 (1997), 3.

[36] Sachi Mizobuchi, Mark Chignell, and David Newton. 2005. Mobile text entry: relationship between walking speed and text input task difficulty. In Proceedings of the 7th international conference on Human computer interaction with mobile devices & services. 122–128.

[37] Myo. 2018. Welcome to MyoTM Support. https://support.getmyo.com/hc/en-us

[38] Nikita Nangia, Adina Williams, Angeliki Lazaridou, and Samuel R Bowman. 2017. The repeval 2017 shared task: Multi-genre natural language inference with sentence representations. arXiv preprint arXiv:1707.08172 (2017).

[39] Andronikos Nedos, Alexander O'Connor, Graham Abell, Siobhán Clarke, and Vinny Cahill. 2004. LATTE: Location and Time Triggered Email.. In International Conference on Wireless Networks. Trinity College Dublin, Dublin, Ireland, 813–819.

[40] Hugo Nicolau, Tiago Guerreiro, Joaquim Jorge, and Daniel Gonçalves. 2010. Proficient blind users and mobile text-entry. In Proceedings of the 28th Annual European Conference on Cognitive Ergonomics. 19–22.

[41] André Rodrigues, Kyle Montague, Hugo Nicolau, and Tiago Guerreiro. 2015. Getting smartphones to talkback: Understanding the smartphone adoption process of blind users. In Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility. ACM, New York, NY, USA, 23–32.

[42] Kristen Shinohara and Jacob O Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 705–714.

[43] Radu-Daniel Vatavu. 2017. Visual impairments and mobile touchscreen interaction: state-of-the-art, causes of visual impairment, and design guidelines. International Journal of Human–Computer Interaction 33, 6 (2017), 486–509.

[44] Hao Wang, Zhengdong Lu, Hang Li, and Enhong Chen. 2013. A Dataset for Research on Short-Text Conversations Proceedings of the Conference on Empirical Methods in Natural Language Processing. ACL.

[45] Youdao. 2020. Translator. http://fanyi.youdao.com