Keep in Touch: Combining Touch Interaction with Thumb-to-Finger μGestures for People with Visual Impairment

Gauthier Faisandaz, Alix Goguey, Christophe Jouffrais, Laurence Nigay

Keep in Touch: Combining Touch Interaction with Thumb-to-Finger µGestures for People with Visual Impairment

Gauthier Faisandaz[1], Alix Goguey[1], Christophe Jouffrais[2,3], Laurence Nigay[1]
[1]Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, Grenoble, France, [2]CNRS, IPAL, Singapore, Singapore, [3]CNRS, University Toulouse 3, Toulouse, France
name.surname@{ [1]univ-grenoble-alpes.fr, [2]cnrs.fr }

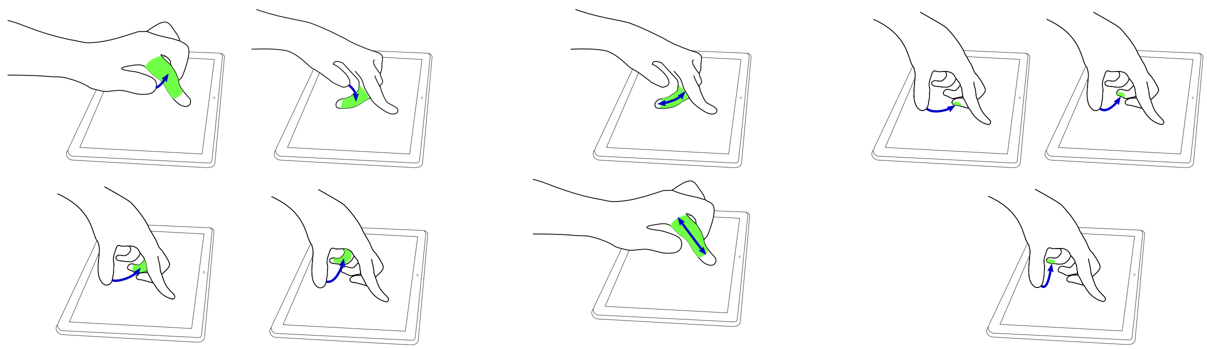


Figure 1: Set of TTF µGestures that can be performed while contacting a touchscreen. Left: three TAP performed by the thumb on the index, middle and either the ring or little finger. Center: two SWIPE along the index or the middle finger, in both directions. Right: three TAP on the nail of the middle, ring and little fingers.

Abstract

We present a set of 8 thumb-to-finger microgestures (TTF µGestures) that can be used as an additional modality to enrich touch interaction in eyes-free situations. TTF µGestures possess characteristics especially suited for people with visual impairment (PVI). They have never been studied specifically for PVI to improve accessibility of touchscreen devices. We studied a set of 33 common TTF µGestures to determine which are feasible and usable without seeing while the index is touching a surface. We found that the constrained position of the hand and the absence of vision prevent participants from being able to efficiently target a specific phalanx.

Thus, we propose a set of 8 TTF µGestures (6 taps, 2 swipes) balancing resiliency (i.e., low error-rate) and expressivity (i.e., number of possible inputs): as a dimension combined with the touch modality, it would realistically multiply the touch command space by eight. Within our set of 8 TTF µGestures, we chose a subset of 4 µGestures (2 taps and 2 swipes) and implemented an exploration scenario of an audio-tactile map with a raised-line overlay on a touchscreen and tested it with 7 PVI. Their feedback was positive on the potential benefits of TTF µGestures in enhancing the touch modality and supporting PVI interaction with touchscreen devices.

CCS Concepts

• **Human-centered computing → Human computer interaction (HCI)**.

Keywords

microgestures; accessibility; visual impairment; touch

ACM Reference Format:
Gauthier Faisandaz, Alix Goguey, Christophe Jouffrais and Laurence Nigay. 2022. Keep in Touch: Combining Touch Interaction with Thumb-to-Finger µGestures for People with Visual Impairment. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '22), November 7–11, 2022, Bengaluru, India.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3536221.3556589

1 Introduction

Microgestures (µGestures) enable interaction at your fingertips: they are small movements performed intentionally as inputs. Single hand µGestures - and specifically, thumb-to-finger (TTF) µGestures, which are made with the thumb on other fingers of the same hand - display great expressivity: they offer a variety of possible gestures (e.g., tap, swipe) that can be performed accurately on distinct parts of the fingers by taking advantage of the hand dexterity and the proprioception. Current research on µGestures shows that they are cognitively undemanding ([47, 57, 70]), eyes-free [69] and always available [57]: three characteristics especially suited for interaction for people with visual impairment (PVI). Yet, while µGestures have been studied in a variety of situations ([10, 15, 64, 69] to cite a few), to the best of our knowledge they have never been studied with PVI, despite the opportunities they could offer to this audience.

TTF µGestures could provide a way to improve the accessibility of various devices, and most notably touchscreen devices. Indeed, PVI rely more and more on touchscreen devices for everyday tasks

ICMI '22, November 7–11, 2022, Bengaluru, India

ACM ISBN 978-1-4503-9390-4/22/11…$15.00
https://doi.org/10.1145/3536221.3556589

[1, 29, 40] such as reading, social interaction, outdoor navigation or object recognition. Nowadays, PVI mainly use two input modalities for interaction: voice and touch. Voice is used a lot by PVI for textual inputs or punctual discrete commands (e.g., opening an application). However, it would rapidly become laborious if vocal commands had to be repeated several times (e.g., changing parameters in a word processing software), and voice interaction raises several concerns regarding privacy and discretion [51]. Being the main interaction modality of touchscreen devices, touch is massively used by PVI. However, interacting with current touchscreen devices still poses accessibility problems despite their built-in accessibility tools (e.g., Android's TalkBack[1], iOs VoiceOver[2]). These tools are commonly named "screen readers": as their name implies, they translate the 2D screen's interface into a 1D list of items to be explored sequentially via an audio feedback. While this allows PVI to access digital content to some extent, it considerably lengthens the interaction, because each item has to be accessed in order. Command inputs are limited in the same manner: for instance, a sighted person could directly select the required item in a menu, while a PVI would have to navigate with taps and swipes through each item until finding the right one. Not only navigating menus is especially long and tedious without visual feedback [49], but even simple commands such as copying and pasting can require a dozen interactions[3]. That is why accessibility tools propose gesture shortcuts (e.g., 2-finger Z-shaped pattern to close an alert) [4]. However shortcuts such as drawing and time-based on-screen gestures are difficult to perform by PVI [38] or are poorly recognized by devices [56]. Moreover, PVI prefer simple gestures (i.e., with one stroke, one finger, one direction, no angle) [14]. All in all, this shows that shortcuts are limited in their capability to improve touchscreen devices' accessibility for PVI.

To enhance touch modality, in this paper we explore a multimodal solution by using TTF μGestures. This is a form of "touch overloading": it consists of adding one or several modalities as an additional dimension to a touch input to trigger different commands [32]. For instance a touch with one finger on an icon application could open it, and the same touch on an icon while the thumb is touching the index could open a contextual menu. There are several approaches using other modalities for touch overloading, such as the pressure level [27, 30, 50] or the hand's posture [48] but most of them rely on visual perception to be used optimally [30, 38, 59] and lack accessibility for PVI. Grounded in this context, the current paper focuses on TTF μGestures as a modality to enrich the touch input space for PVI. To explore the combination of the touch modality with the TTF μGestures modality, we present the results of an experiment to identify which TTF μGestures are usable while touching the touchscreen of a device. Our goal is to identify a set of TTF μGestures to overload the touch modality. Concretely, we studied a set of 33 TTF μGestures based on the literature to determine which are feasible and usable while the index is touching a surface, thus extending previous literature on TTF μGestures studies [15, 34, 57, 70]. We made the hypothesis that some μGestures would be much harder to perform in this situation compared to studies where the hand position is not constrained. The set comprises taps, double taps and swipes performed by the thumb on all fingers of the same hand while differentiating the areas on which they are done (i.e., three phalanxes per finger plus nail). We designed a glove prototype and an experiment in order to systematically evaluate each TTF microgesture of the set in terms of error rate and comfort. We found that the constrained position of the hand and the absence of vision hinder participants when targeting a specific phalanx. In our experiment, 34.2% of the time, participants performed the gesture in the wrong area. By increasing the target size (i.e., the TTF μGesture can be performed on any phalanx of the finger), we drastically improved participants' performances. Thus, our results allow us to put forward a set of 8 TTF μGestures (6 taps, 2 swipes) that are usable while the index is touching a surface. This proposed set is a fair tradeoff between resiliency (i.e., low error-rate) and expressivity (i.e., number of possible inputs): as a dimension combined with the touch modality, it would realistically multiply the touch command space by eight. This gesture set can serve as a conservative baseline for researchers and designers that need a touch overloading modality for PVI - or even sighted participants.

Within our set of 8 TTF μGestures (Figure 1), we chose a subset of 4 TTF μGestures (2 taps and 2 swipes) and implemented an exploration scenario of an audio-tactile map with a raised-line overlay on a touchscreen and tested it with 7 PVI. Their feedback was positive on the potential benefits of TTF μGestures in enhancing the touch modality and supporting PVI interaction with touchscreen devices.

We contribute the following results and insights: 1) a set of 8 TTF μGestures usable conjointly with the touch modality; 2) an application which use a subset of 4 TTF μGestures to interact with an audio-tactile map on tablet; 3) feedback of 7 PVI about the utility and usability of TTF μGestures to enhance touch interaction.

2 Related Work

Current touchscreen devices have built-in accessibility features (e.g., Voice Over for iOS and Talkback for Android) which include verbal descriptions and accessible input interaction features. Several studies evaluated these accessibility tools in a variety of contexts and many problems persist [2, 20]. Typically, accessible interactive features include eyes-free menus and shortcuts. Menus usually rely on a triggering gesture, and then multiple gestures to select the required element. These sequences are long and tedious (e.g., copying/pasting can require an input trajectory with a dozen of gestures). On-screen gesture shortcuts are quicker than menus. They can be simple (e.g., a double tap to zoom in) or more complex (e.g., 2-finger Z-shaped pattern to close an alert). They can however be difficult to perform by PVI [38] or be poorly recognized [56].

Accessibility features based on the touch modality have inherent problems that need to be tackled. In this section we first review the approaches for touch overloading before we focus on TTF μGesture as a new modality for PVI.

2.1 Touch Overloading: Enriching touch input

Touch overloading could be a valuable solution to improve the input bandwidth of touch interaction. On current touchscreen devices, touch overloading usually relies on the use of multiple contacts (e.g., pinch), repetition (e.g., double tap) and/or time (e.g., dwell or hold) as supplementary dimensions to enhance input possibilities.

[1] Talkback for Android
[2] Voice Over for iOS
[3] How to copy-paste with VoiceOver
[4] Apple's support: How to use gestures with VoiceOver

However the research community has explored other approaches. Some leveraged the shear force produced by the finger on the surface [31], or discriminated between different parts of the finger (e.g., knuckle, nail, tip) to trigger different commands [32, 45]. Several studies used various levels of pressure on the surface [27, 30, 50], while other relied on multitouch through bimanual interactions, for instance using a second hand to switch between modes [8]. This can also be achieved by using another finger of the same hand [28]. Relying on rhythmic patterns was also proposed [26] as well as associating command triggers to a specific number of contacts on a surface [44]. Such approaches increase the bandwidth for touch interaction, but most of them do not scale up (e.g,. for multitouch, number of limbs are limited), while repetition (e.g., quadruple taps) and time-based (e.g., dwell) options are tedious and delay the interaction flow. Besides, most of these approaches rely on visual perception to be used optimally [30, 59] and lack accessibility for PVI [38]. For instance, visual feedback (e.g., gauges) is necessary to precisely control pressure-based input [30, 59], while location and orientation-based inputs require spatial accuracy [38].

Research studies on touch overloading for PVI are scarce. In order to improve touchscreen interaction for PVI, many solutions rely on tangible and tactile overlays [23]. Kane et al. created physical cut out overlays for touchscreens marked with a fiduciary tag [37]. These overlays not only add tactile cues but also change the action of a touch event passing through the cut outs. As it technically uses another dimension (i.e., the overlays), it can arguably be considered as touch overloading. However solutions relying on tangible overlay can be expensive and cumbersome [37]. Since voice is frequently used by PVI as an input modality, another approach is to combine touch with speech commands. Previous research has shown that combined gesture and voice input can improve interaction in various contexts (e.g., desktop, mobile device, immersive 3D graphics, augmented reality) [9]. The benefit of combining gesture and speech is that they can be used to express complementary inputs [17] but all the studied combinations rely on a visual context for gesture, not applicable for PVI. For enhancing touch without visual context, speech commands define a possible approach but PVI can be reluctant to talk to their device in public spaces [51], or feel they lack control[7, 18, 73]. In this work we focus on a new modality for PVI, μGestures, to enrich touch interaction.

2.2 Touch Overloading for PVI

TTF μGestures: an Eyes Free Input Modality: μGestures are small movements performed intentionally as a form of input. They are part of what Ashbrook defines as "micro-interactions": interactions that last less than four seconds [4]. Wolf et al. [69, 70] describe μGestures as slight movements performed with the hands and fingers but other parts of the body can be used too (e.g., feet [3, 19], head [21], eyes [35, 61], even teeth [16]).

Single-hand μGestures, as defined by Chan et al. [15], are finger gestures performed using one hand on itself. In their elicitation study of 1632 gestures, the thumb was used 88% of the time. Thumb-to-finger (TTF) μGestures (i.e., the thumb touching another finger) were further studied by Soliman et al. [58], who identified its gestural primitives and introduced a four-dimensional design space: A) the touch initiator (i.e., which finger perform the touch), B) what location on another finger is touched, C) what touch action is performed (i.e., a tap, a swipe), D) how fingers are flexed.

Kuo et al. investigated thumb and finger functional work space (i.e., the range of movements) and found that the Thumb-Index and Thumb-Middle finger pairs have the maximal functional work space [41]. Huang et al. found that tap and stroke μGestures performed by the thumb on the index or middle finger of the dominant hand are comfortable [34]. Each finger can be divided in two or three segments (i.e., phalanxes) clearly delimited by knuckles and wrinkles. Several studies used them as touch targets [36, 55, 62, 68], while others further subdivided the segments [34, 63].

Preference and perceived comfort of the fingers' areas for TTF μGestures was explored by a number of studies [34, 36, 39, 62, 64]. They show that the index and middle fingers rate high in comfort of use, followed by the ring then little fingers in decreasing order. Comfort rating follows a similar trend along the finger, with the tip of the finger rating high, followed by the middle section and then the bottom one. Preference between the volar (i.e., palm side) and radial side (i.e., side facing the thumb) of the finger was studied and showed a tendency toward the radial side for the index and middle fingers, but the volar side for the ring and little fingers [36, 62]. The nail area was also studied: several studies conclude that it is a viable, always available, and appreciated input location [39, 42].

As far as we know, work on TTF μGestures for PVI has barely been addressed to-date [24, 52]. Partly because they allow for TTF μGestures, Feng found that hand-mounted controllers (e.g., ring) were preferred to other types of controllers such as hip-mounted or head-mounted (e.g., belt, glasses) [24]. Oh et al. studied preferences between several on-body touch input locations and found that same-hand gestures (not necessarily TTF μGestures) rated last out of the five locations in terms of ease of use and comfort, mostly because participants were unfamiliar with the interaction [52].

Combining touch and TTF μGestures: A touchscreen device allowing only for simple and double taps, combined with only two TTF μGestures (e.g., the thumb touching the index or the middle finger), could provide 2+2*2=6 types of taps (i.e., simple tap and double tap without the thumb touching any finger, same with the thumb touching the index, and same with the thumb touching the middle finger). This motivated the studies that used TTF μGestures to increase touch input expressivity [8, 63, 72]. For instance, Tsai et al. overloaded touch input on a smartphone by using TTF μGestures to switch between modes when the thumb touched specific areas on the index [63]. Surale et al. compared six overloading techniques for switching touch mode: 1) long press, 2) non-dominant hand, 3) two-fingers, 4) hard press, 5) knuckle, and 6) thumb-to-finger [8]. They found that the thumb-to-finger technique (i.e., a TTF μGesture) is the best technique when touch accuracy is required. In addition, it was highly rated by participants, both overall and on the ease-of-use aspect. Altogether, these results show that TTF μGestures are usable and enhance expressivity. However, no study have been conducted to evaluate and compare the usability of the most common TTF μGestures found in literature to be used in conjunction with touch modality. That is why, in the following, the first study focuses on which TTF μGestures can be used when the index is in contact with a touchscreen. Having identified the set of possible TTF μGestures, a second study shows how to use a subset of these TTF μGestures in a multimodal application.

3 Study 1: Usability of constrained μGestures

3.1 Experiment Design

Our study investigates from an ergonomic standpoint which TTF μGestures can be used as an extra modality conjointly to the touch modality when the index is in contact with a touchscreen.

Participants: 9 sighted volunteers (3 females) aged from 20 to 34 (Mean 25.9, Std. dev 4.6) participated in this experiment. All participants (except one) are right-handed, 3 play an instrument. The left-handed participant uses computers with a mouse and plays an instrument (bass guitar) as a right-handed person.

As it focuses on μGestures motor control, we do not need PVI participants for this study. Indeed, performance and comfort of doing μGestures are related to the biomechanics of the hand rather than whether the subject is visually impaired or not. Our rationale is based on Palani et al.[53]: in three experiments, they found no differences between PVI and blindfolded sighted participants in the ability to perform kinesio-tactile perceptual tasks. In addition, in a systematic review on empirical evaluations of technology for PVI, Brulé et al. [13] found that 178 studies tested interaction techniques "at the level of actions" with sighted participants, which shows that it is generally adequate to do so. Finally, these authors mentioned that, when possible, relying on sighted volunteers reduces constraints related to PVI participating in experiments (such as mobility issues and availability of volunteers) [13].

Apparatus: Participants are seated in front of a 10.1" tablet (Huawei MediaPad T5 Lite, 24*15,5 cm), laid on a table in front of them. A 17" screen is placed at head level in front of them to display instructions and feedback. To capture the μGestures performed during the experiment, some techniques are based on cameras, however, they suffer greatly from occlusion (i.e joints and finger segments are hidden from camera' sight) and require calibration [55, 58]. Other approaches based on remote sensors have been designed, including Tomography [75], radar [65], IMU [62], magnetometer [15], and a combination of a microphone and a gyroscope [74]. These methods are promising for non-intrusive recognition of μGestures but the set of recognized gestures is still limited. Other approaches rely on contact sensors [39, 66, 67]. More easily reproducible approaches are based on gloves with touch sensors [36, 43, 62, 71]. Our glove prototype uses conductive yarn and an Arduino board (Figure 2). A custom Python software runs the experiment, receives and processes data from the glove and logs the experimental data.

TTF μGesture set: As PVI mainly use their left hand when exploring a tactile or digital graphic [6], participants are asked to perform TTF μGestures with their left hand, without looking at it and with the index touching the tablet's screen. Based on previous studies on hands-free TTF μGestures [34, 36, 62], we selected tap

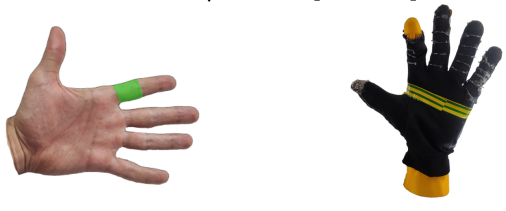


Figure 2: Left: Example of a stimulus shown to the participants: a TAP μGesture at the base of the index. Right: Glove prototype used in this experiment.

and swipe TTF μGestures as they are preferred by users [15]. The finger parts on which to perform TTF μGestures are the phalanxes [34, 36] and the nails [39]. In our experimental setting, the index finger is in contact with the surface, so the top phalanx and nail of the index were not reachable with the thumb and were thus discarded from the set. Our set includes 33 TTF μGestures in total: **11 TAP** (top, middle, bottom phalanxes except the index top phalanx × index, middle, ring, little); **11 2-TAP** (the same areas); **3 nail TAP** (middle, ring, little); **8 SWIPE** (top to bottom, bottom to top × index, middle, ring, little).

Task and Procedure: The task consists of doing with the left hand each of the 33 TTF μGestures shown on the screen (Figure 2) as fast and precisely as possible while keeping the index finger on the tablet's screen. Participants are asked not to look at their hand during the experiment, and the experimenter ensures they comply with the latter. To become familiar with the task, they go through a preliminary tutorial phase. They perform each of the 33 μGestures in a fixed order: TAP, 2-TAP then SWIPE. They must correctly perform each μGesture before going to the next step. Then, during the test phase, participants repeat eight blocks in which they have to perform the 33 μGestures presented in a pseudo random order. For each trial, participants start with the index finger on the tablet. They are then prompted with a μGesture picture on the computer screen (Figure 2). A trial ends when the participant performs the correct μGesture. All attempts were logged. Participants can take a break whenever they want in between trials. After the experiment, participants rate the perceived physical comfort for each μGesture on a 5-point Likert scale from 1 ("very uncomfortable") to 5 ("very comfortable"). A typical session lasts approximately 50 minutes.

3.2 Data and Analysis

We use a within-subjects design, with repeated measures, and recorded a total of 9 participants * 33 TTF μGestures * 8 blocks = 2376 gestures. Our data set is shared on osf.io.

Error Rate (ER): sum of Error Occurrences (i.e., fail on the first attempt) divided by the total number of trials. The ER is used as a "score" for studying and improving the gesture set.

Comfort: subjective rating between 1 ("very uncomfortable") and 5 ("very comfortable").

We filter out outliers (trials in which participants make more than 5 erroneous attempts) following the 95% distribution rule. In total, 140 trials out of 2376 were removed (5.9% of the data), resulting in 2236 trials. We use an estimation approach [22] to communicate the results instead of the null hypothesis statistical testing.

3.3 Results of the experiment

Error Rate: Figure 3 shows the mean ER and CI for the Tutorial and the Test sessions (all blocks, blocks 1 to 3 and blocks 4 to 8). We expected that participants would overall become better after a few blocks, hence we considered the first three blocks as a training period before participants reach a performance plateau. Following Dragicevic's advice, the highlight and dotted boxes show that CI's ranges of the first three blocks and the last five blocks do not overlap, strongly suggesting a learning effect. Figure 4a shows the ER for each type μGestures (TAP, 2-TAP, SWIPE), globally then per finger. CI's ranges of the SWIPE performed on the index and middle fingers do not overlap with the ring and little fingers. This strongly suggests

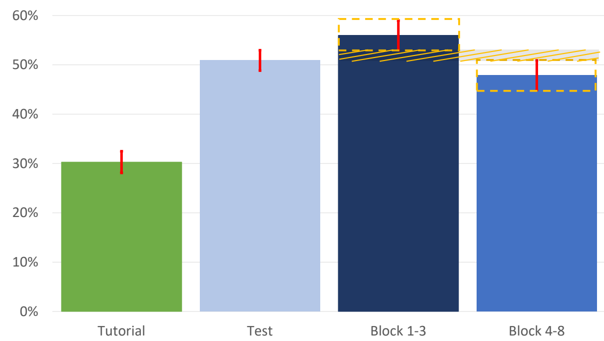


Figure 3: Error Rate for the Tutorial (left column) and Test phases (all other columns). The highlight and dotted boxes (in yellow) show that CI's ranges of the first three blocks and the last five blocks do not overlap.

that SWIPE performed on the index or middle fingers are less error-prone than SWIPE performed on the ring and little fingers. CI's ranges of the TAP performed on the index, middle and ring fingers scarcely overlap with CI's ranges of the 2-TAP performed on the same fingers. This suggests that TAP performed on the phalanxes of those three fingers are significantly less error-prone than 2-TAP performed on the same areas. CI's range of TAP performed on the nails does not overlap with anything. This strongly suggests that TAPS on nails are significantly less error-prone than any of the other µGesture. Figure 4a also shows that 2-TAP is the most error-prone µGesture: participants failed to perform it on the first attempt 67.9% of the time. On the contrary, TAP µGestures are less error-prone (42.4% including phalanx and nail TAP). In this figure, we show separately TAP made on nails separately (i.e., "Nails" labeled bar) from TAP made on finger phalanxes (i.e., finger labeled bars), as they are noticeably less error-prone."

Comfort: Figure 4b shows the average comfort rating, between 1 (lowest) and 5 (highest), given by participants during the post-experiment interview, by µGesture, target finger and target phalanx.

3.4 Analysis

Our results show that each of the 33 TTF µGesture is physically feasible even while the index is touching a surface. In this context, users' preference go to µGestures on fingers that are near the thumb (i.e., index and middle fingers) and on the top phalanxes (i.e., farthest phalanxes from the palm). TAP on nails are especially popular with participants. This is consistent with our quantitative results: TAP are the least error-prone µGestures (42.4%) followed by SWIPE (44.3%), and they are the most usable when performed on the index and the middle fingers. Nails are especially usable, with the lowest error-rate of all µGestures (9.2%). On the contrary, performing µGestures on the ring and little fingers is harder than on other fingers and is less preferred by participants. 2-TAP causes significantly more errors than other µGestures (67.9%).

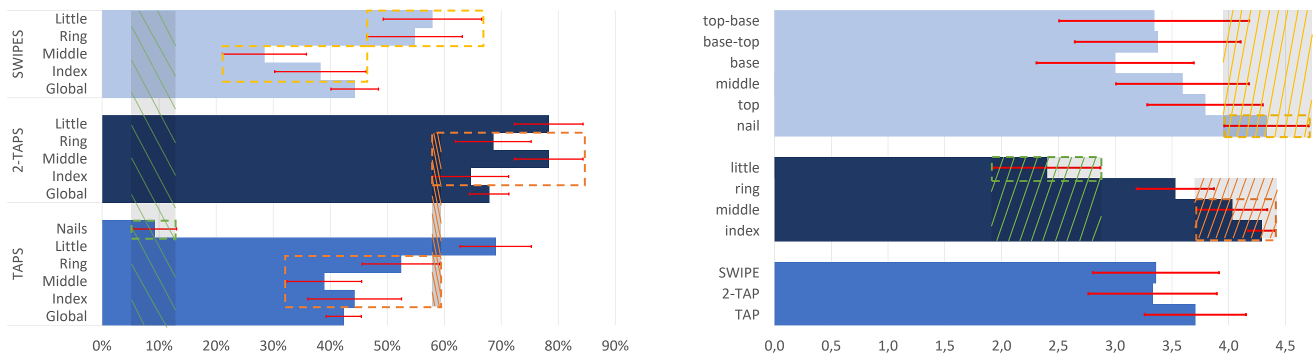
We aim at providing design guidelines and a µGesture set suitable for usage while the index is touching the screen. Our goal is to identify the set of µGestures with the lowest ER and the best Comfort: to do so, we are refining the initial set of 33 µGestures into a less error-prone set by reducing the error factors (e.g., precise targeting) and removing the most problematic µGestures. This refinement process is characterized by a tradeoff between resiliency (i.e., low ER) and expressivity (i.e., large set of µGestures). We base our analysis on a typology of errors that a participant may make during a trial: **Gesture-error**, the wrong µGesture is performed; **Finger-error**, the µGesture is performed on the wrong finger; **Area-error**, the µGesture is performed on the right finger but on the wrong part of the finger. If at least one of these errors happens before completing the trial, an Error Occurrence is counted once. These errors are not mutually exclusive. A participant can make several errors when performing a µGesture. For instance, if a participant is asked to perform a TAP on the base area of the middle finger but performs a 2-TAP on the top area of the middle finger, there is a Gesture-error and an Area-error. Out of 2249 trials, Gesture-errors represent an ER of 30.3% (677 occurrences, CI 4.3%). Finger-errors represent an ER of 12.5% (297 occurrences, CI 2.0%). Area-errors represent an ER of 34.2% (765 occurrences, CI 3.7%).

3.5 Refining the µGesture set

TAP & 2-TAP: We are refining TAP and 2-TAP together, as they are the same kind of inputs. Seeing as TAP µGestures on nails are the least error-prone by far (Figure 4a) and the most comfortable µGestures (Figure 4b), we decided to keep them as is in our set of µGestures, without the need to refine them.

When taking only TAP and 2-TAP on phalanxes (i.e., no TAP on nails) into account, Table 5 shows an ER of 59.5%. It also shows that Area-errors are the main source of error (47.4%). The context of the experiment could be one explanation: because of the constrained position of the hand and the eyes-free situation, participants probably have difficulties targeting a phalanx precisely. One solution is to increase the target size by considering only one area on each finger. In this situation, the ER for TAP and 2-TAP together is reduced from 59.5% to 41.1%, as shown in Figure 6. To obtain this score we consider that performing the task on the required finger – independently of the required area, aside from nails – was sufficient to validate the trial. TAP µGestures performed on nails are comparatively less error-prone (9.2%), as Figure 4a indicates, which justifies treating them separately . The lower right highlight and box (in orange) in Figure 6 show that CI's range for this refinement step does not overlap with the previous CI's range, which strongly suggests that this refinement step has a significant impact on the ER. The TAP & 2-TAP µGesture set is therefore reduced from 14 to 11 in total (4 TAP and 4 2-TAP, one per finger + 3 nail TAP).

Gesture-errors are the second source of errors (31.8%). As TAP and 2-TAP are the same kind of input, they can be merged to prevent confusion between the two µGestures. Figure 4a shows that CI's ranges of the TAP performed on the index, middle and ring fingers scarcely overlap with CI's ranges of the 2-TAP performed on the same fingers. This suggests that TAP performed on the phalanxes of those three fingers are significantly less error-prone than 2-TAP performed on the same areas, and that correctly performing 2-TAP is more difficult. Thus we merge 2-TAP with TAP as a second refinement step: we consider all 2-TAP performed as TAP instead. In this configuration, Figure 6 shows that the global ER is reduced from 41.1% down to 25.9%. The lower right highlight and box (in orange) in Figure 6 show that CI's range for this refinement step does not overlap with the next CI's range, which strongly suggests



(a) Error Rate for each µGesture, globally and for each finger. Dotted boxes (in yellow) show no CI overlap between index and middle fingers SWIPE and ring and little fingers SWIPE. Highlight and dotted boxes (in orange) show almost no CI overlap between index, middle and ring fingers TAP and 2-TAP. Highlight and dotted box on the left (in green) shows no CI overlap between nail TAP and the rest.

(b) Comfort rating and CI sorted by µGesture (bottom), finger targeted (middle), and area targeted (top). Highlight and dotted box (in yellow) show close to no CI overlap between nail TAP and the rest. Highlight and dotted box (in green) show no CI overlap between little finger µGesture and the rest. Highlight and dotted box (in orange) show almost no CI overlap between index and middle fingers µGesture and ring finger ones.

Figure 4: Error Rate for each Gesture, globally and for each finger (left). Comfort rating by µGesture, finger and area (right).

that this refinement step has a significant impact on the ER. The TAP gesture set is reduced from 11 (4 TAP and 4 2-TAP, one per finger + 3 nail TAP) to 7 (4 TAP, one per finger + 3 nail TAP).

As shown in Table 5 and Figure 4a, µGestures performed on the little finger are unanimously the most error-prone. Participants also rate them with the lowest comfort, as shown in Figure 4b. This is

			# Trials	Error Occurrences	Error Rate	CI
Test sessions: TAP & 2-TAP only	Global		1473	876	59.5%	±2.5%
	Error type	Gesture	1473	468	31.8%	±5.2%
		Finger		197	13.4%	±2.6%
		Area		699	47.4%	±4.9%
	Target finger	Index	278	142	51.1%	±5.9%
		Middle	411	211	45.1%	±4.8%
		Ring	394	237	52.1%	±4.8%
		Little	390	286	63.3%	±4.4%
			# Trials	Error Occurrences	Error Rate	CI
Test sessions: SWIPE only	Global		547	243	44.4%	±4.2%
	Error type	Gesture	547	208	38%	±4.1%
		Finger		64	11.7%	±2.7%
		Area		58	10.6%	±2.6%
	Target finger	Index	141	54	38.3%	±5%
		Middle	144	41	28.5%	±4.3%
		Ring	136	75	55.1%	±5.8%
		Little	126	73	57.9%	±6%

Figure 5: Error Rate by type of µGesture, globally and by fingers, plus error type, for (up) TAP and 2-TAP only and (bottom) SWIPE only.

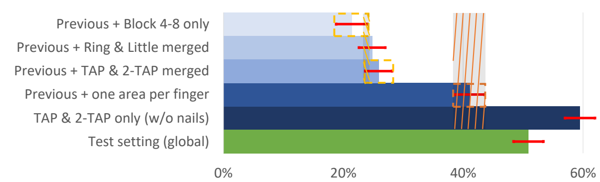


Figure 6: Error rate with CI for TAP and 2-TAP µGestures of the Test sessions, for each step of the refinement process. Highlight and dotted boxes (in yellow) show that CIs barely overlap between the *one area per finger, TAP and 2-TAP merged* configuration and the last refinement step. Highlight and dotted box (in orange) show no CI overlap between the *one area per finger* configuration and the rest.

consistent with literature findings [34, 36, 55]. Instead of completely discarding µGestures on the little finger, we merged the little finger with the ring finger: we considered these two fingers as one finger (and one area), except for Gestures on nails. If a TAP on the ring or little finger is required, it can be performed on either of these two. Doing so, it reduces the ER from 25.9% to 24.8%, as shown in Figure 6. The TAP gesture set is reduced from 7 to 6 (3 TAP + 3 nail TAP).

In Figure 3, blocks 1 to 3 have a higher mean ER than blocks 4 to 8. This suggests a learning effect, as participants gradually become better at performing the µGestures. We thus base our predictions on the remaining blocks. As shown in Figure 6, taking only blocks 4 to 8 into consideration decreases the ER from 24.8% to 21.4%. The upper left highlight and boxes (in yellow) in Figure 6 do not overlap much, which suggest that the last two refinement steps (i.e., Ring Little finger fusion, plus block 4-8 only) have a significant impact on the ER. After refining the TAP µGesture set, we have 6 TAP µGestures available: 3 on fingers (index, middle, ring / little) and 3 on the nails (middle, ring, little). As shown in Figure 6, this refinement process reduces the ER from 59.5% to 21.4%.

SWIPE: We refine SWIPE apart, as they are the only bi-directional inputs. When taking only SWIPE into account, Table 5 shows an ER of 44.4% and that 11.7% of the errors in the Test phase are Finger-errors. To prevent accidentally performing a SWIPE on another finger, we keep only one SWIPE active at a time, on a finger that triggers the fewest errors. This means we consider that any SWIPE performed on any other finger is not happening thus not triggering an error. Table 5 shows that the middle finger is the least error prone (28.5%) in the Test phase with SWIPE µGestures only, followed by the index finger (38.3%). Figure 4b shows that CI's ranges of the SWIPE performed on the index and middle fingers do not overlap with the ring and little fingers. This strongly suggests that SWIPE performed on the index or middle fingers are less error-prone than SWIPE performed on the ring and little fingers. In addition, Figure 4b shows that the index and middle fingers are the fingers preferred by users. Hence, to prevent confusion and false recognition, we recommend using only one finger on which to perform the SWIPE, either on the index or middle finger. In this configuration, the ER of SWIPE decreases from 44.4% to 38.3% with the index finger activated only, and to 28.5% for the middle finger. The SWIPE gesture set is thus reduced from 8 (1 SWIPE per finger along each direction) to 2 (1 SWIPE on the index or middle finger, along each direction).

We cannot further improve the ER: Table 5 shows that most of the errors in the Test phase are Gesture-errors (38%), which means SWIPES are detected as TAP or 2-TAP. The context of the experiment could be one explanation. Because of the constrained position of the hand and the eyes-free situation, participants probably performed the SWIPE µGestures in two steps: first they targeted the area and made contact, then they performed the swiping gesture. But the prototype was designed with a time threshold of 500ms initiated at first contact and after which, in the absence of other information, the contact was considered as a TAP.

As we did for the TAP µGestures, we are considering only blocks 4 to 8. Figure 7 shows that it decreases the ER from 38.3% to 33% for the index finger, and from 28.5% to 25.6% for the middle finger. The right highlight and box (in orange) in Figure 7 show that CI's range for SWIPE µGestures do overlap with CI's range for SWIPE µGestures performed on the index, but not on the middle finger. This suggests that using the index only to perform SWIPE µGestures is not significantly different than using SWIPE with all fingers activated. However, this is not the case after users have passed the learning period (i.e., block 1-3): CI's range for SWIPE µGestures only slightly overlap with CI's range for SWIPE µGestures performed on the index for block 4-8, suggesting that the difference is significant.

After refining the SWIPE µGesture set, the resulting set includes 2 SWIPE µGestures available on either the index or middle finger. As shown in Figure 7, the refinement process decreases the ER from 44.4% to 33% for the index, and to 28.5% for the middle finger. The upper left highlight and boxes (in yellow) in Figure 7 show that CI's ranges for SWIPE µGestures performed on the index and middle finger for block 4-8 overlap a lot. While we cannot assert that they are strictly equivalent, we believe that they can be used interchangeably with minimal differences between the two.

3.6 Final µGesture set

Our refinement process applied to our initial set of 33 µGestures leads us to define a final set of 8 µGestures (Figure 1: 3 finger TAP (i.e., index, middle, ring/little), 3 nail TAP (i.e., middle, ring, little) and 2 SWIPE (i.e., index or middle finger, along two directions)

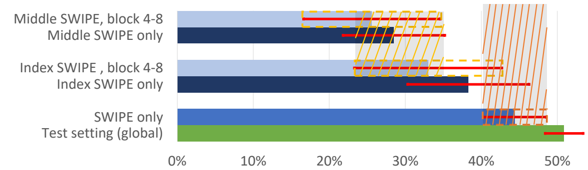


Figure 7: Error rate with CI for SWIPE µGestures of the Test sessions, for each step of the refinement process. Highlight and dotted boxes (in yellow) show that CIs overlap between index and middle finger SWIPE for block 4-8. Highlight and dotted box (in orange) show that CIs overlap between the overall SWIPE and index SWIPE.

which can be used in conjunction with the index touching a surface. Using all 8 µGestures results in an Error Rate of 27.9% (450 Error Occurrences out of 1616 Trials, CI 2.2%). We want to stress that the Error Rate is dependent on our experimental setting and that it serves only for internal consistency. It is not our intention to make it our contribution nor is it to assess whether our glove is the best device for our interaction technique. Based on our observations and results, we developed a second prototype for Study 2: this new glove contains fewer sensors than the first one as it is built to recognize the identified set of µGestures. This new prototype glove is simpler and less bulky, and should help reduce the overall ER.

4 Study 2: Use case study of µGestures for PVI

This follow-up study aims at illustrating the use of the identified set of µGestures as part of multimodal interaction for PVI and obtaining qualitative feedback from PVI. Our scenario is based on the use of an audio-tactile device [11, 12] and uses only a subset of the µGestures identified in Study 1: an index TAP, a little finger nail TAP, a bi-directional index SWIPE. We chose these µGestures as they are among the easiest and most comfortable ones (Figure 4b).

4.1 Experiment Design

Participants We recruited 7 participants with VI, all right-handed, aged from 23 to 63 y.o (mean 36.3 , Std. Dev 14), one female. We carefully selected them so that they have different backgrounds and their levels of visual impairment range between category 1 (i.e.,

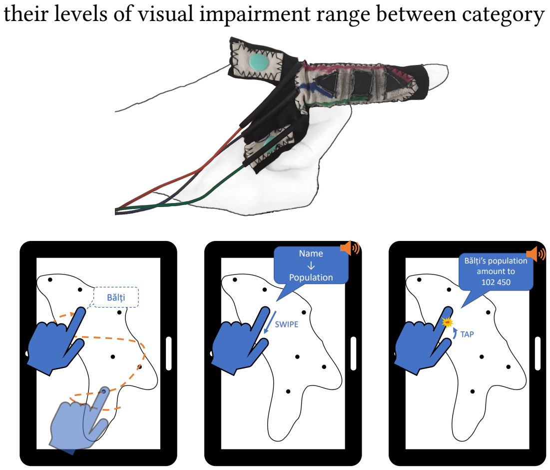


Figure 8: (Up) New glove. (Bottom) Usage illustration of the audio-tactile map: touching a POI selects it; an index SWIPE changes category; an index TAP starts an audio feedback.

mild visual impairment) and 5 (i.e., complete blindness) of the WHO classification [25]). P1 (24 y.o), has no central vision and can only use peripheral vision (WHO cat. 2, diagnosed at 7 y.o). P2 (46 y.o) and P3 (34 y.o) have low vision (both WHO cat. 1, both diagnosed at 1 y.o). P4 (30 y.o), has complete blindness (WHO cat. 5, since birth). P5 (23 y.o), has severe visual impairment (WHO cat. 3, since birth). P6 (34 y.o), has low vision (WHO cat. 1, since birth). P7 (63 y.o), has pigmentary retinopathy with near complete blindness (WHO cat. 4, diagnosed at 6 y.o). Only P2, P4, P5, P7 have used tactile documents.

Apparatus: For our second prototype, we used capacitive components, allowing us to use the bare thumb instead of covering it with conductive fabric. Sensors can be better felt, improving the affordance (Figure 8). The software was similar to Study 1.

Usage Scenario: PVI have a different paradigm of interaction with tactile documents than sighted users have with graphical ones. Because they explore the document with their hands, integration of the content must be done more sequentially compared to a sighted user, by aggregating a succession of incomplete details until enough has been gained to synthesize and understand the whole picture [46]. When exploring audio-tactile documents, PVI rely on bimanual strategies mainly based on movements of the two index finger tips touching the tactile document [5, 6, 54, 76]. But existing audio-tactile devices require back and forth movements between the document being explored and a command menu placed sideways, which interrupts exploration strategies and increases cognitive load [11]. TTF µGestures aim to address this issue as input gestures in addition to the touch modality for exploration. Our goal is to use TTF µGestures to trigger located feedback and commands without breaking the contact with the explored graphic. Drawing from this use case, we adapted the audio-tactile map application from [11, 12]. We represented Moldova's geography on a tablet with a tactile overlay. Using this application, users can retrieve information about Moldova's cities and regions (Figure 8. Users select a POI by touching it with his left index. They can then circle through a menu with category items ("Name", "Population" and "Surface area") using index finger SWIPE µGestures. Users can switch between abstraction levels ("City" and "Region") using a little finger nail TAP. At any time, a TAP on the index triggers an audio feedback reading the selected POI, category and level information (e.g., Surface area of the city of Chisinau).

Task and procedure: Each session started with a presentation of the TTF µGestures and apparatus. Participants then tried all four TTF µGestures: in the air first, then with the index on the map and finally with the glove on. After trying all interactions a few times, they had to explore the map to answer a series of 14 questions such as "How many inhabitants are there in city X?". These questions ensured a wide variety of multimodal interaction (i.e., changing "POI" and/or "category" and/or "abstraction level"). They were asked to think aloud. After the exploration task, participants answered a questionnaire about the perceived usability and utility. Usability questions covered ease to perform and learn the µGestures, and potential discomfort. Utility questions covered perceived utility and qualities of the multimodal interaction technique, and other potential usage scenarios. This study took 15min on average.

4.2 Results

Usability: All participants agreed that the index TAP was easy to perform. P2 and P7 had some issues with the nail TAP, and P3 said that while easy to do, it was the hardest µGesture of the set. All but P4 found the SWIPE gestures easy to perform. P5 and P7 said "Once you get it, it is easy" and "Everybody can do it". Three participants said that TTF µGestures were quick to perform. Five said they were practical. Two said they were readily available. All participants strongly agreed on the ease of learning the interaction technique. In situ observations further support their comments: they all became independent after a few minutes of usage, even though P2 needed some reminders throughout the session. P1, P3, P6 found the interaction technique "very intuitive". P3 said "he could picture the menu along his finger". P6 said the µGesture set can be "remembered easily". Apart from issues with the TAP on the little finger's nail (P2, P7), none of them mentioned anything about the interaction technique but pointed out the prototype's shortcomings (e.g., lag, size of the glove...) instead.

Utility: All participants stated that this interaction technique can be useful for eyes-free interaction while interacting with a touchscreen device. P1, P3 and P7 spontaneously mentioned that TTF µGestures allow you to interact and keep your hand in place instead of having to move the hand away (e.g., to perform an on-screen gesture or to interact with a side menu) and lose context. P2 and P7 also said that TTF µGesture could be useful for changing parameters on the fly (e.g., stopping a long or unwanted audio feedback, changing audio output speed). As for using TTF µGestures in other situations, six participants mentioned at first an educational context (e.g., learning geography, anatomy or geometry at school). P3 said that it could be used on any surface. He also mentioned it could be used as additional input with a computer (e.g., use of macro shortcuts) or modifier-keyboard-keys (e.g., CTRL, ALT) but for touchscreen. He also said that µGesture-based techniques could be particularly useful for people suffering from motor handicap (e.g., tetraplegia). P5 and P6 said it would be interesting for games.

Overall: In summary, all participants found TTF µGestures useful, usable, intuitive and comfortable when being used to explore an audio-tactile document. They were really interested in the interaction possibilities TTF µGestures open and were especially curious about concrete scenarios that we would develop in the future.

5 General Discussion and Limitations

We want to stress that our aim is not to assess whether our glove is the best device for our interaction technique but rather if the µGestures it senses can provide benefits in our context. Having shown the potential of µGestures as an input modality combined with touch, further research effort can be put into technologies to accurately capture the identified µGestures according to the context of use (e.g., a camera in a fixed setting [55], a ring in a mobile setting [10]). Notably, the second prototype we made for Study 2 has less sensors and is therefore less bulky and easier to use.

Study 1 shows that users' preferences go to TTF µGestures on fingers that are near the thumb (i.e., index, middle finger) and on the top phalanxes (i.e., farthest phalanxes from the palm) when the index is in contact with a surface. This is in line with the literature and we show that these existing results hold when the index is in contact with a surface. Most of the errors (34,2%) in Study 1 are triggered because of the difficulty to precisely target a phalanx (i.e., Area-errors). This can be explained by the constrained position of the hand and the absence of visual feedback. Indeed, in a hand-free situation with visual feedback, performing a TTF μGesture is a trivial task for the users [41, 70]. So our recommendation is to consider the entire finger as a unique area. This reduces the accuracy needed and allows users to target a comfortable area on a given finger, but also reduce the number of possible μGestures per finger. To further reduce errors, we have opted for a solution that merges the sources of confusion (e.g., 2-TAP considered as TAP), which in turn results in a decrease in the number of inputs. Still, the set of 8 μGestures that we propose is a fair tradeoff between resiliency (i.e., low Error Rate) and expressivity (i.e., number of possible inputs). Indeed, as a modality combined with the touch modality, it would realistically multiply the touch command space by eight. While we conscientiously thought during the experimental design about PVI and sighted having developed different manual dexterity, we decided to go for a conservative condition. First, regardless of their handedness, PVI mostly use their left hand to explore tactile graphics [6]. Second, asking sighted users to use their left hand although they are right-handed should be considered as a worst-case scenario. Hence, we hypothesized that PVI should be able to use μGestures with more ease than sighted users with their left hand.

Study 2 shows that PVI appreciate TTF μGestures as an input modality in combination with the touch modality. Mostly, they find it "quick", "intuitive", "easy to perform" and "practical". All gestures were considered easy to perform while keeping their index on the tablet. Two PVI participants mentioned that the TAP on the little finger's nail was the most difficult of the 4 μGestures. We suppose that nails TAP having a good comfort rating in study 1 being perceived as hard in study 2 can be explained by several factors: nails TAP might be perceived as quite efficient when compared to the other 33 μGestures in general (study 1) but not so significantly when compared to the easiest 3 μGestures (e.g., TAP on index) (study 2); in study 1 the comfort rating was influenced by the error rate and nails TAP are particularly resilient while in study 2 an error rate was not present; in study 1 nails TAP comfort scores were grouped together, while in study 2, only the little finger's nail TAP was used, which might be the hardest of the lot. We are conscious that we cannot generalize findings of Study 2 with 7 participants. Also, study 2 may have suffered from the novelty effect. Nonetheless, they validated the usability of the multimodal technique and unanimously reported the same characteristics (i.e., quick, intuitive, practical, easy to perform) for justifying the appeal of the technique as compared to the technique they use on a daily basis. This makes us confident that the positive feedback is not solely due to the novelty effect and strongly suggests that TTF μGestures are promising and offer desirable qualities for PVI interaction on tactile devices.

The participants were adamant about this technique being useful. Nevertheless it was hard for them during the questionnaire to identify use-cases that would benefit from it beyond the scenario we proposed. One explanation we put forward is that they do not use much of the possibilities offered by tactile devices (e.g., copy/paste) because of their complexity with current accessibility tools [1]. Most of the PVI would not even know such features are possible, therefore they could not envision that it can be done. Six of them said that it could be useful in education and two mentioned games. They further suggested that the multimodal technique would be pertinent in situations where it could be used concomitantly while exploring a document (e.g., changing a parameter on the fly, while continuing the exploration) rather than sequentially (e.g., having to interrupt the primary task with the document). Without explicitly stating it this way, the participants suggested that the multimodal approach taken could be promising for concurrent tasks and not just for synergistic usage for a given task like in Study 2. The exploration of a document defines a promising scenario involving concurrent tasks. Indeed [76], in line with [33, 60], showed that as a part of a cognitive strategy to interpret a tactile document, fingers frequently stop and go during the exploration. TTF μGestures could be a complementary input modality of choice as μGestures could be triggered on the fly during those stops and with minimal impact on the cognitive process of understanding the document.

6 Conclusion

Being the main interaction technique of tablets and smartphones, touch as an input modality is increasingly ubiquitous but still poses usability challenges to PVI. To improve the touch modality for PVI, in this paper we investigated the TTF μGesture modality, a modality not yet studied for accessibility but possessing many desirable characteristics (i.e., eyes-free, unobtrusive, cognitively undemanding) for PVI. Our first study contributes to a first exploration of which TTF μGestures are physically feasible and comfortable while touching a surface in an eyes-free situation. Our results suggest that in this context, targeting a specific phalanx is difficult. Moreover, despite our constrained context, the results are coherent with literature on the following points: 1) performing TTF μGestures on the ring and little fingers is harder than on other fingers and is less preferred by participants [34, 36, 39, 62, 64]; 2) taps on middle, ring and little finger's nails are especially usable and popular with participants [39, 42].From this first study we identify a set of 8 TTF μGestures (i.e., 6 TAP and 2 SWIPE) that can be usable as a touch overloading modality in an eyes-free situation. This gesture set can serve as a conservative baseline for designers that need a touch overloading modality for PVI as well for sighted participants. We illustrated the use of this baseline set of TTF μGestures by designing a multimodal application combining the touch and μGesture modalities for PVI. To do so we selected 4 TTF μGestures from the baseline set. We experimentally tested the multimodal application with 7 participants with visual impairment. This second study shows that PVI can easily engage with this multimodal interaction technique. In particular, they appreciate it for being quick, instinctive, practical and easy to perform. The collected feedback from PVI also highlights that the multimodal technique could be useful in several contexts of use. Participants were really interested by the possibility of using TTF μGestures in eyes-free interaction. They notably expressed that the technique could be useful in educational contexts, in public spaces (e.g., audio-guides in museums with tactile representations of paintings), for games, and for software applications with parameters that can be adjusted on the fly (e.g., granularity of text selection in word processors).

As future work, we plan to further investigate the multimodal technique based on the identified set of 8 TTF μGestures with PVI. One use case would be an alternative to current text edition, which seems to pose many problems to PVI [40]: they could swipe towards the base of their index finger to copy the selected text (as a metaphor of grabbing it in your hand), and drop the copied text with the opposite SWIPE (as a metaphor of putting it down).

7 Acknowledgments

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) funded by the French program Investissement d'avenir. We also thank our partners: Laboratory Cherchons pour Voir, CNRS CESDV-IJA

References

[1] Carl Halladay Abraham, Bert Boadi-Kusi, Enyam Komla Amewuho Morny, and Prince Agyekum. 2021. Smartphone usage among people living with severe visual impairment and blindness. *Assistive Technology* 0, 0 (2021), 1–8. https://doi.org/10.1080/10400435.2021.1907485

[2] Nancy Alajarmeh. 2021. The extent of mobile accessibility coverage in WCAG 2.1: sufficiency of success criteria and appropriateness of relevant conformance levels pertaining to accessibility problems encountered by users who are visually impaired. *Universal Access in the Information Society* (2021). https://doi.org/10.1007/s10209-020-00785-w

[3] Jason Alexander, Teng Han, William Judd, Pourang Irani, and Sriram Subramanian. 2012. Putting your best foot forward: Investigating real-world mappings for foot-based gestures. *Conference on Human Factors in Computing Systems - Proceedings* (2012), 1229–1238. https://doi.org/10.1145/2207676.2208575

[4] Daniel Ashbrook. 2010. *Enabling Mobile Microinteractions.* Ph. D. Dissertation.

[5] Sandra Bardot, Marcos Serrano, and Christophe Jouffrais. 2016. From Tactile to Virtual: Using a Smartwatch to Improve Spatial Map Exploration for Visually Impaired Users. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 100–111. https://doi.org/10.1145/2935334.2935342

[6] Sandra Bardot, Marcos Serrano, Bernard Oriola, and Christophe Jouffrais. 2017. Identifying how visually impaired people explore raised-line diagrams to improve the design of touch interfaces. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2017-May. ACM Press, New York, New York, USA, 550–555. https://doi.org/10.1145/3025453.3025582

[7] Grace M. Begany, Ning Sa, and Xiaojun Jenny Yuan. 2016. Factors Affecting User Perception of a Spoken Language vs. Textual Search Interface: A Content Analysis. *Interact. Comput.* 28 (2016), 170–180.

[8] Hemant Bhaskar Surale, Fabrice Matulic, and Daniel Vogel. 2017. Experimental Analysis of Mode Switching Techniques in Touch-based User Interfaces multi-touch; touch input; mode switching. (2017). https://doi.org/10.1145/3025453.3025865

[9] Mark Billinghurst. 2013. Hands and Speech in Space: Multimodal Interaction with Augmented Reality Interfaces. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction* (Sydney, Australia) *(ICMI '13)*. Association for Computing Machinery, New York, NY, USA, 379–380. https://doi.org/10.1145/2522848.2532202

[10] Roger Boldu, Alexandru Dancu, Denys J.C. C Matthies, Pablo Gallego Cascón, Shanaka Ransir, and Suranga Nanayakkara. 2018. Thumb-In-Motion: Evaluating Thumb-to-Ring Microgestures for Athletic Activity. In *Proceedings of the Symposium on Spatial User Interaction (SUI '18)*. Association for Computing Machinery, New York, NY, USA, 150–157. https://doi.org/10.1145/3267782.3267796

[11] Anke M. Brock, Philippe Truillet, Bernard Oriola, Delphine Picard, and Christophe Jouffrais. 2015. Interactivity Improves Usability of Geographic Maps for Visually Impaired People. *Human-Computer Interaction* 30, 2 (2015), 156–194. https://doi.org/10.1080/07370024.2014.924412

[12] Emeline Brule, Gilles Bailly, Anke Brock, Frédéric Valentin, Grégoire Denis, and Christophe Jouffrais. 2016. MapSense: Multi-sensory interactive maps for children living with visual impairments. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, 445–457. https://doi.org/10.1145/2858036.2858375

[13] Emeline Brulé, Brianna J. Tomlinson, Oussama Metatla, Christophe Jouffrais, and Marcos Serrano. 2020. *Review of Quantitative Empirical Evaluations of Technology for People with Visual Impairments.* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi-org.ins2i.bib.cnrs.fr/10.1145/3313831.3376749

[14] Maria Claudia Buzzi, Marina Buzzi, Barbara Leporini, and Amaury Trujillo. 2017. Analyzing visually impaired people's touch gestures on smartphones. *Multimedia Tools and Applications* 76, 4 (2017), 5141–5169. https://doi.org/10.1007/s11042-016-3594-9

[15] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Dong Yang, and Frank Maurer. 2016. User Elicitation on Single-Hand Microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 3403–3414. https://doi.org/10.1145/2858036.2858589

[16] Victor Chen, Xuhai Xu, Richard Li, Yuanchun Shi, Shwetak Patel, and Yuntao Wang. 2021. Understanding the Design Space of Mouth Microgestures. (jun 2021). arXiv:2106.00931 https://arxiv.org/abs/2106.00931v1

[17] P. R. Cohen, M. Dalrymple, D. B. Moran, F. C. Pereira, and J. W. Sullivan. 1989. Synergistic Use of Direct Manipulation and Natural Language. *SIGCHI Bull.* 20, SI (mar 1989), 227–233. https://doi.org/10.1145/67450.67494

[18] Eric Corbett and Astrid Weber. 2016. What Can I Say? Addressing User Experience Challenges of a Mobile Voice User Interface for Accessibility. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 72–82. https://doi.org/10.1145/2935334.2935386

[19] Andrew Crossan, Stephen Brewster, and Alexander Ng. 2010. Foot Tapping for Mobile Interaction. (2010). https://doi.org/10.5555/2146303.2146366

[20] Rafael Jeferson Pezzuto Damaceno, Juliana Cristina Braga, and Jesús Pascual Mena-Chalco. 2018. Mobile device accessibility for the visually impaired: problems mapping and recommendations. *Universal Access in the Information Society* 17, 2 (2018), 421–435. https://doi.org/10.1007/s10209-017-0540-1

[21] James W. Davis and Serge Vaks. 2001. A perceptual user interface for recognizing head gesture acknowledgements. *ACM International Conference Proceeding Series* 15-16-Nove (nov 2001), 21. https://doi.org/10.1145/971478.971504

[22] Pierre Dragicevic. 2016. Fair Statistical Communication in HCI. Springer, Cham, 291–330. https://doi.org/10.1007/978-3-319-26633-6_13

[23] Julie Ducasse, Anke M. Brock, and Christophe Jouffrais. 2018. Accessible Interactive Maps for Visually Impaired Users. In *Mobility of Visually Impaired People*. Springer International Publishing, 537–584. https://doi.org/10.1007/978-3-319-54446-5_17

[24] Catherine Feng. 2016. Designing wearable mobile device controllers for blind people: A co-design approach. In *ASSETS 2016 - Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. Association for Computing Machinery, Inc, 341–342. https://doi.org/10.1145/2982142.2982144

[25] W H O Programme for the Prevention of Blindness and Deafness. 2003. Consultation on development of standards for characterization of vision loss and visual functioning : Genveva, 4-5 September 2003. , WHO/PBL/03.91 pages.

[26] Euan Freeman, Gareth Griffiths, and Stephen A. Brewster. 2017. Rhythmic micro-gestures: Discreet interaction on-The-go. *ICMI 2017 - Proceedings of the 19th ACM International Conference on Multimodal Interaction* 2017-Janua, September (nov 2017), 115–119. https://doi.org/10.1145/3136755.3136815

[27] Alix Goguey, Sylvain Malacria, and Carl Gutwin. 2018. Improving Discoverability and Expert Performance in Force-Sensitive Text Selection for Touch Devices with Mode Gauges. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018). https://doi.org/10.1145/3173574

[28] Alix Goguey, Daniel Vogel, Fanny Chevalier, Thomas Pietrzak, Nicolas Roussel, and Géry Casiez. 2017. Leveraging finger identification to integrate multi-touch command selection and parameter manipulation. *International Journal of Human-Computer Studies* 99 (mar 2017), 21–36. https://doi.org/10.1016/J.IJHCS.2016.11.002

[29] Nora Griffin-Shirley, Devender R Banda, Paul M Ajuwon, Jongpil Cheon, Jaehoon Lee, Hye Ran Park, and Sanpalei Nylla Lyngdoh. 2017. A Survey on the Use of Mobile Applications for People who Are Visually Impaired. *Journal of Visual Impairment & Blindness* 111 (2017), 307–323.

[30] LeeLik Hang, LamKit Yung, LiTong, BraudTristan, SuXiang, and HuiPan. 2019. Quadmetric Optimized Thumb-to-Finger Interaction for Force Assisted One-Handed Text Entry on Mobile Headsets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (sep 2019), 1–27. https://doi.org/10.1145/3351252

[31] Chris Harrison and Scott E Hudson. 2012. Using Shear as a Supplemental Two-Dimensional Input Channel for Rich Touchscreen Interaction. (2012).

[32] Chris Harrison, Julia Schwarz, and Scott E Hudson. 2011. TapSense: Enhancing Finger Interaction on Touch Surfaces. *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11* (2011). https://doi.org/10.1145/2047196

[33] Everett W. Hill, John J. Rieser, Mary-Maureen Hill, Mary-Maureen Hill, John A. Halpin, and R. Halpin. 1993. How Persons with Visual Impairments Explore Novel Spaces: Strategies of Good and Poor Performers. *Journal of Visual Impairment & Blindness* 87 (1993), 295 – 301.

[34] Da Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong Hao Liang, De Nian Yang, Yi Ping Hung, and Bing Yu Chen. 2016. Digitspace: Designing Thumb-to-fingers touch interfaces for one-handed and eyes-free interactions. *Conference on Human Factors in Computing Systems - Proceedings* January 2019 (2016), 1526–1537. https://doi.org/10.1145/2858036.2858483

[35] Robert J.K. Jacob. 1990. What you look at is what you get: Eye movement-based interaction techniques. *Conference on Human Factors in Computing Systems - Proceedings* (mar 1990), 11–18. https://doi.org/10.1145/97243.97246

[36] Haiyan Jiang, Dongdong Weng, Zhenliang Zhang, and Feng Chen. 2019. HiFinger: One-Handed Text Entry Technique for Virtual Environments Based on Touches between Fingers. *Sensors (Switzerland)* 19, 14 (jul 2019), 1–24. https://doi.org/10.3390/s19143063

[37] Shaun K Kane, Meredith Ringel Morris, and Jacob O Wobbrock. 2013. Touchplates: Low-Cost Tactile Overlays for Visually Impaired Touch Screen Users. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility* (2013). https://doi.org/10.1145/2513383

[38] Shaun K Kane and Jacob O Wobbrock. 2011. Usable Gestures for Blind People: Understanding Preference and Performance. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011). https://doi.org/10.1145/1978942

[39] Hsin Liu Kao, Artem Dementyev, Joseph A. Paradiso, and Chris Schmandt. 2015. NailO: Fingernails as an input surface. *Conference on Human Factors in Computing Systems - Proceedings* 2015-April, April (apr 2015), 3015–3018. https://doi.org/10.1145/2702123.2702572

[40] Akif Khan and Shah Khusro. 2021. An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. *Universal Access in the Information Society* 20, 2 (2021), 265–298. https://doi.org/10.1007/s10209-020-00733-8

[41] Li Chieh Kuo, Haw Yen Chiu, Cheung Wen Chang, Hsiu Yun Hsu, and Yun Nien Sun. 2009. Functional workspace for precision manipulation between thumb and fingers in normal hands. *Journal of Electromyography and Kinesiology* 19, 5 (oct 2009), 829–839. https://doi.org/10.1016/J.JELEKIN.2008.07.008

[42] Do Young Lee, Soo Hwan Lee, and Ian Oakley. 2020. Nailz: Sensing Hand Input with Touch Sensitive Nails. *Conference on Human Factors in Computing Systems - Proceedings* (apr 2020). https://doi.org/10.1145/3313831.3376778

[43] Lik Hang Lee, Kit Yung Lam, Tong Li, Tristan Braud, Xiang Su, and Pan Hui. 2019. Quadmetric Optimized Thumb-to-Finger Interaction for Force Assisted One-Handed Text Entry on Mobile Headsets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 94 (sep 2019), 27 pages. https://doi.org/10.1145/3351252

[44] G. Julian Lepinski, Tovi Grossman, and George Fitzmaurice. 2010. The design and evaluation of multitouch marking menus. *Conference on Human Factors in Computing Systems - Proceedings* 4 (2010), 2233–2242. https://doi.org/10.1145/1753326.1753663

[45] Guanhong Liu, Yizheng Gu, Yiwen Yin, Chun Yu, Yuntao Wang, Haipeng Mi, and Yuanchun Shi. 2020. Keep the Phone in Your Pocket: Enabling Smartphone Operation with an IMU Ring for Visually Impaired People. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (jun 2020). https://doi.org/10.1145/3397308

[46] Jack M. Loomis, Roberta L. Klatzky, and Susan J. Lederman. 1991. Similarity of Tactual and Visual Picture Recognition with Limited Field of View. *Perception* 20 (1991), 167 – 177.

[47] Denys J. C. Matthies, Bodo Urban, Katrin Wolf, and Albrecht Schmidt. 2019. Reflexive Interaction: Extending the Concept of Peripheral Interaction. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction* (Fremantle, WA, Australia) *(OZCHI'19)*. Association for Computing Machinery, New York, NY, USA, 266–278. https://doi.org/10.1145/3369457.3369478

[48] Fabrice Matulic, Daniel Vogel, and Raimund Dachselt. 2017. Hand contact shape recognition for posture-based tabletop widgets and interaction. *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, ISS 2017* (oct 2017), 3–11. https://doi.org/10.1145/3132272.3134126

[49] Oussama Metatla, Fiore Martin, Tony Stockman, and Nick Bryan-Kinns. 2014. "Non-Visual Menu Navigation: the Effect of an Audio-Tactile Display". 213–217. https://doi.org/10.14236/ewic/hci2014.25

[50] Takashi Miyaki and Jun Rekimoto. 2009. Graspzoom: Zooming and scrolling control model for single-handed mobile interaction. *MobileHCI09 - The 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (2009). https://doi.org/10.1145/1613858.1613872

[51] Aarthi Easwara Moorthy and Kim-Phuong L Vu. 2015. Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. *International Journal of Human–Computer Interaction* 31, 4 (2015), 307–335. https://doi.org/10.1080/10447318.2014.986642

[52] Uran Oh and Leah Findlater. 2014. Design of and subjective response to on-body input for people with visual impairments. In *ASSETS14 - Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM Press, New York, New York, USA, 115–122. https://doi.org/10.1145/2661334.2661376

[53] Hari prasath Palani, Paul Fink, and Nicholas Giudice. 2020. Design Guidelines for Schematizing and Rendering Haptically Perceivable Graphical Elements on Touchscreen Devices. *International Journal of Human-Computer Interaction* 36 (04 2020), 1–22. https://doi.org/10.1080/10447318.2020.1752464

[54] Chris Perkins and Ann Gardiner. 2003. Real world map reading strategies. *Cartographic Journal* 40, 3 (2003), 265–268. https://doi.org/10.1179/000870403225012970

[55] Manuel Prätorius, Dimitar Valkov, Ulrich Burgbacher, and Klaus Hinrichs. 2014. DigiTap: An Eyes-Free VR/AR Symbolic Input Device. *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology - VRST '14* (2014). https://doi.org/10.1145/2671015

[56] 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 and Accessibility (ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 23–32. https://doi.org/10.1145/2700648.2809842

[57] Adwait Sharma, Joan Sol Roo, and Jürgen Steimle. 2019. Grasping Microgestures: Eliciting Single-Hand Microgestures for Handheld Objects. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19, Chi)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300632

[58] Mohamed Soliman, Franziska Mueller, Lena Hegemann, Joan Sol Roo, Christian Theobalt, and Jürgen Steimle. 2018. FingerInput: Capturing Expressive Single-Hand Thumb-to-Finger Microgestures. *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces* 18, December (2018), 177–187. https://doi.org/10.1145/3279778.3279799

[59] Craig Stewart, Michael Rohs, Sven Kratz, and Georg Essl. 2010. Characteristics of Pressure-Based Input for Mobile Devices. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10* (2010). https://doi.org/10.1145/1753326

[60] J M Tellevik. 1992. Influence of Spatial Exploration Patterns on Cognitive Mapping by Blindfolded Sighted Persons. *Journal of Visual Impairment & Blindness* 86, 5 (1992), 221–224. https://doi.org/10.1177/0145482X9208600508

[61] Tribhuwan Kumar Tewari, Anshul Arya, and Sameer Rastogi. 2010. Message Reading Through Eye Blinking. *International Journal of Computer Applications* 2, 6 (jun 2010), 1–4. https://doi.org/10.5120/676-951

[62] Hsin Ruey Tsai, Lee Ting Huang, Cheng Yuan Wu, and Yi Ping Hung. 2016. ThumbRing: Private interactions using one-handed thumb motion input on finger segments. *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, MobileHCI 2016* (2016), 791–798. https://doi.org/10.1145/2957265.2961859

[63] Hsin Ruey Tsai, Te Yen Wu, Min Chieh Hsiu, Jui Chun Hsiao, Da Yuan Huang, Yi Ping Hung, Mike Y. Chen, and Bing Yu Chen. 2017. SegTouch: Enhancing touch input while providing touch gestures on screens using thumb-to-index-finger gestures. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. Part F1276. ACM, New York, NY, USA, 2164–2171. https://doi.org/10.1145/3027063.3053109

[64] Jérémy Wambecke, Alix Goguey, Laurence Nigay, Lauren Dargent, Daniel Hauret, Stéphanie Lafon, and Jean Samuel Louis De Visme. 2021. M[eye]cro: Eye-gaze+Microgestures for Multitasking and Interruptions. *Proceedings of the ACM on Human-Computer Interaction* 5, EICS (may 2021), 22. https://doi.org/10.1145/3461732

[65] Saiwen Wang, Jie Song, Jaime Lien, Ivan Poupyrev, and Otmar Hilliges. 2016. Interacting with Soli: Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 851–860. https://doi.org/10.1145/2984511.2984565

[66] Martin Weigel, Tong Lu, Gilles Bailly, Antti Oulasvirta, Carmel Majidi, and Jürgen Steimle. 2015. ISkin: Flexible, Stretchable and Visually Customizable On-Body Touch Sensors for Mobile Computing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2991–3000. https://doi.org/10.1145/2702123.2702391

[67] Martin Weigel, Aditya Shekhar Nittala, Alex Olwal, and Jürgen Steimle. 2017. SkinMarks: Enabling Interactions on Body Landmarks Using Conformal Skin Electronics. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3095–3105. https://doi.org/10.1145/3025453.3025704

[68] Eric Whitmire, Mohit Jain, Divye Jain, Greg Nelson, Ravi Karkar, Shwetak Patel, and Mayank Goel. 2017. DigiTouch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (sep 2017), 1–21. https://doi.org/10.1145/3130978

[69] Katrin Wolf. 2016. Microgestures—Enabling Gesture Input with Busy Hands. In *Peripheral Interaction: Challenges and Opportunities for HCI in the Periphery of Attention*, Saskia Bakker, Doris Hausen, and Ted Selker (Eds.). Springer International Publishing, Cham, 95–116. https://doi.org/10.1007/978-3-319-29523-7_5

[70] Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 2011. Taxonomy of Microinteractions: Defining Microgestures Based on Ergonomic and Scenario-Dependent Requirements. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part I (INTERACT'11, Vol. 6946 LNCS)*. Springer-Verlag, Berlin, Heidelberg, 559–575. https://doi.org/10.1007/978-3-642-23774-4_45

[71] Pui Chung Wong, Kening Zhu, and Hongbo Fu. 2018. FingerT9: Leveraging Thumb-to-Finger Interaction for Same-Side-Hand Text Entry on Smartwatches. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3173574.3173752

[72] Haijun Xia, Tovi Grossman, and George Fitzmaurice. 2015. Nanostylus: Enhancing input on ultra-small displays with a finger-mounted stylus. *UIST 2015 - Proceedings of the 28th Annual ACM Symposium on User Interface Software and*

Technology (nov 2015), 447–456. https://doi.org/10.1145/2807442.2807500
[73] Mary Zajicek, Richard Wales, and Andrew Lee. 2004. Speech interaction for older adults. , 122–130 pages. https://doi.org/10.1007/s10209-004-0091-0
[74] Cheng Zhang, Anandghan Waghmare, Pranav Kundra, Yiming Pu, Scott Gilliland, Thomas Ploetz, Thad E. Starner, Omer T. Inan, and Gregory D. Abowd. 2017. FingerSound: Recognizing Unistroke Thumb Gestures Using a Ring. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 120 (sep 2017), 19 pages. https://doi.org/10.1145/3130985
[75] Yang Zhang and Chris Harrison. 2015. Tomo: Wearable, Low-Cost Electrical Impedance Tomography for Hand Gesture Recognition. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software amp; Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 167–173. https://doi.org/10.1145/2807442.2807480
[76] Kaixing Zhao, Sandra Bardot, Marcos Serrano, Mathieu Simonnet, Bernard Oriola, and Christophe Jouffrais. 2021. Tactile Fixations: A Behavioral Marker on How People with Visual Impairments Explore Raised-Line Graphics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 27, 12 pages. https://doi.org/10.1145/3411764.3445578