Background Perception and Comprehension of Symbols Conveyed through Vibrotactile Wearable Displays

Granit Luzhnica
Know Center
Graz, Austria
gluzhnica@know-center.at

Eduardo Veas
Interactive Systems and Data Science,
Graz University of Technology
& Know Center
Graz, Austria
eveas@know-center.at

ABSTRACT

Previous research has demonstrated the feasibility of conveying vibrotactile encoded information efficiently using wearable devices. Users can understand vibrotactile encoded symbols and complex messages combining such symbols. Such wearable devices can find applicability in many multitasking use cases. Nevertheless, for multitasking, it would be necessary for the perception and comprehension of vibrotactile information to be less attention demanding and not interfere with other parallel tasks. We present a user study which investigates whether high speed vibrotactile encoded messages can be perceived in the background while performing other concurrent attention-demanding primary tasks. The vibrotactile messages used in the study were limited to symbols representing letters of English Alphabet. We observed that users could very accurately comprehend vibrotactile such encoded messages in the background and other parallel tasks did not affect users performance. Additionally, the comprehension of such messages did also not affect the performance of the concurrent primary task as well. Our results promote the use of vibrotactile information transmission to facilitate multitasking.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Haptic devices**; **Empirical studies in HCI**.

KEYWORDS

haptic feedback; tactile feedback; skin reading; stimulation; haptic display; wearables; user study; multi-tasking; multimodal; HCI

ACM Reference Format:
Granit Luzhnica and Eduardo Veas. 2019. Background Perception and Comprehension of Symbols Conveyed through Vibrotactile Wearable Displays. In *24th International Conference on Intelligent User Interfaces (IUI '19), March 17–20, 2019, Marina del Rey, CA, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3301275.3302282

IUI '19, March 17–20, 2019, Marina del Rey, CA, USA

ACM ISBN 978-1-4503-6272-6/19/03…$15.00
https://doi.org/10.1145/3301275.3302282

1 INTRODUCTION

This paper presents evidence supporting the hypothesis that arbitrary symbol based messages transmitted using vibrotactile wearable displays can be comprehended while concentrating on other demanding primary tasks. Our results open up new possibilities to communicate vibrotactile patterns in the background without detrimental effects on a primary task.

Nowadays, wearable and mobile devices assist and enrich daily activities, for example, with information collected by the sensors within them. The primary feedback modalities of mobiles and wearables are visual and auditory. They compete for user attention and distract the user from important tasks. On the other hand, due to the proliferation of wearable devices, vibrotactile capabilities are accessible in wearable devices to a substantial number of end-users, but yet, are mostly utilised to complement the visual interaction. Nevertheless, tactile feedback can be used to transmit rich information without the need to perceive it through auditory or visual channels [15, 16]. The information can be delivered in the form of **tactons**, defined as vibrotactile patterns representing abstract meanings [2]. Nevertheless, for tactile communication to be a viable alternative in multitasking scenarios, it is necessary to assess whether it can be interpreted in the background and what are its effects on the primary task.

Perceiving information through the skin using vibrotactile wearable devices can be beneficiary in a broad spectrum of applications for visual and/or hearing impaired users by utilising sensory substitution where a sensory modality (e.g. vision or auditory) is captured, processed and then transmitted to a user via vibrotactile stimuli [14]. With that goal in mind, researchers tried transforming the entire speech signal of a spoken language to vibrotactile patterns. Such a method has been shown to work for a limited number of words [4]. As a result, numerous hearing aid solutions have been proposed [3, 19, 22–24, 34], which complement lip reading for users with hearing disabilities. Yet, with such methods it was not possible for users to comprehend messages consisting only of vibrotactile stimuli. They needed lip reading as well to understand the messages. Other researchers encode information by first encoding symbols [5, 10, 15, 16, 35] (e.g. letters or phonemes) using vibrotactile tactons and then combine such symbols into words [16, 35]. Luzhnica et al. [16] refer to this method as **skin reading**. Recent studies demonstrate that with only a few hours of training participants are able to decode symbols as well as interpret words and short sentences [13, 16] transmitted at a high speed and perceived with a high accuracy entirely using a glove-based wearable vibrotactile display.

Given its efficiency (perception accuracy, speed of transmission), skin reading can be useful beyond impaired sensory substitution, in general purpose applications to facilitate multitasking or reduce demands on the predominant visual displays. For instance, users would be able to perceive their phone notifications, SMS, emails etc... while performing other task e.g. driving, biking, working, etc... For most of such use case, multitasking is a crucial aspect. Extensive research has been carried out on optimizing discriminability (perception) and throughput (transmission speed) of vibrotactile patterns [12, 16]; and on perceiving information [10, 13, 15, 16, 35]. However, studies where such information is perceived in the background while performing other tasks are rare or rather missing.

In this paper, we present a user study which investigates whether vibrotactile tactons which represent letters of English Alphabet can be concurrently perceived in the background (as secondary task) while performing another attention demanding task. We recruit only participants who are trained in skin-reading and are proficient on understanding such tactons. We hypothesise that the perception of vibrotactile encoded symbols should be handled through automatic cognitive processing and should not affect the performance of the other primary task which is designed to be challenging and requires controlled processing. Thus, this work targets the following research questions:

RQ1. *Can trained users percieve and decode high-speed vibrotactile encoded symbols in the background while performing another attention demanding primary task?*

RQ2. *Does the background perception of such vibrotactile encoded symbols affect the performance of the primary task?*

RQ3. *Does the primary task affect the performance of the background perception of vibrotactile encoded symbols?*

2 RELATED WORK

Starting with Braille's invention of the Braille coding in 1824, haptic displays have long been widely used by people with visual impairments. Research on tactile displays equipped with actuators has been ongoing since at least 1924 [4], where Gault [4] used a piezoelectric unit to convert entire recorded speech to touch. Similarly, Kirman [7] used a 15×15 vibrator matrix on the palm to teach six participants to differentiate between the patterns of 15 different words. Other researchers attempted to utilise a visually oriented approach, where a low-resolution image of the object is projected to an array of stimulators. For instance, White [32] transformed images captured from a video feed to a 20×20 vibrotactile display placed on the back. After training, participants were able to distinguish simple shapes like circle, square and triangle. Bliss [1] developed the first commercial device capable of capturing text from the video feed and then imprinting each letter on the finger with a 6×24 matrix of vibrators.

A more successful approach of transmitting information through haptics was provided by Geldard [5]. The device was named Vibratese and used five vibromotors on the chest to encode 45 symbols (letters, numbers and most frequent short words). The author reported that after 65 hours of training one participant was able to understand 38 wpm (words per minute). Recently, Luzhnica et al. followed a different encoding scheme using only the location of vibromotors to encode 26 letters of English alphabet [16]. The authors used six vibromotors on the back of the hand and were able to train users to perceive letters, words and phrases through skin within only five hours, albeit with repetition of stimuli. Later, Luzhnica and Veas [15] demonstrated the importance of encoding the alphabet as optimising it drastically improves the comprehension of encoded messages by users.

A crucial aspect of tactile displays is how they encode information and the choice of vibrotactile patterns. It requires providing patterns that are discriminative while delivering them as fast as possible. Typically a combination of variations in amplitude [28, 29, 33], frequency [28, 29, 33], duration [5, 6] and body locations [5, 17, 20, 27, 33] have been used. For instance, Geldard [5] used five locations, a variation of three durations and three intensities to encode the desired symbols. Recently, Novich [21] showed that spatiotemporal patterns, where vibromotors in a pattern are turned on and off sequentially one after the other, result in significantly better discrimination than the spatially encoded patterns where all vibromotors in a pattern onset simultaneously. Liao [10] utilised such a spatiotemporal patterns to encode symbols on the wrist. Although such encoding works well [10, 21] in terms of being identified by participants, it is many times slower than the spatial encoding. Luzhnica [16] used prioritised overlapping spatiotemporal patterns where vibromotors are activated in sequence after a gap, and they stay on until the pattern is finished (see Figure 3). This method resulted in better recognition accuracy compared to spatial patterns, and it is faster than spatiotemporal encoding, as vibromotors share most of the activated time. Recently, Luzhnica and Veas [15] demonstrated the importance of choosing the right alphabet encoding (mapping from patterns to letters) as optimising it drastically improveD the comprehension of encoded messages by users. Recent research also leverages spatial acuity (sensitivity of locations of actuators) for achieving a better perception of encoded information [11, 12].

Although, conveying alphanumerical symbols has been successfully demonstrated [5, 10, 13, 15–17], studies focusing on conveying trained messages in the background as a secondary task to the best of our knowledge to date are rather missing.

Cognitive processing theory suggests that when users are presented with multiple stimuli/tasks while multitasking, they prioritise or ignore some of them if attention bottlenecks occur [25]. There are two categories of cognitive processes: controlled and automatic which is defined by the amount of attention needed. Automatic processes occur without the need for attention and process initiation, are considered effortless and do not draw general processing resources. As such, they do not interfere with other parallel occurring thought processes [30]. On the other hand, controlled processes are considered very flexible but very costly at the expense of the attentional resources available [30]. Schneider et al. [26] showed that the very same tasks could be processed using one or other processing model depending on the user experience and proficiency on the given task. In their short memory experiment, in a condition where parameters (search target) of the task changed, users needed constant attention for solving it, and thus the task required controlled processing. When performing it in parallel with another task that required controlled processing, the performance on both tasks declined. On the contrary, in a condition where the same parameters were kept constant, after users gathered a lot of experience and became proficient with the task, the processing became automatic. When combined with another task that required controlled processing, the performances of both tasks were not affected by each other.

Lee and Starner [9] tested the perception of three vibrotactile patterns presented on the wrist as a secondary task. They used non-overlapping spatiotemporal patterns containing three vibromotors activated in a sequence for a total duration of 1.5 seconds. Participants were not trained to associate patterns with the meaning but asked to build their own mental model instead. As a primary task, a visual search task with three levels of difficulty was used. The authors reported that the primary task and secondary task were not significantly affected by each other in terms of accuracy, but they were affected in terms of reaction time. In our work, we use the same primary task as proposed and used by the authors as mentioned earlier [9]. As for the vibrotactile secondary task, first, we encode symbols representing letters of English alphabet and participants are trained to recognise the entire English alphabet prior to the study. Second, we use concise overlapping spatiotemporal patterns where each symbol is encoded with only 100-110 ms as we aim to maximise the throughput for information transmission. Both, the very short duration and the number of encoded symbols are expected to increase the difficulty of vibrotactile symbol identification which we hope to compensate by the pre-training where the recognition of such tactons (encoded symbols) is already well formed and crystallised.

3 USER STUDY

This user study aims at evaluating how well users can perceive tactons in background while performing another attention demanding primary task. The tactons used in this study represent letters of the English Alphabet.

Prior to this study, participants were trained in 5 sessions to recognise all 26 letters of English Alphabet. Three days after the last training session (session 5), participants were invited to take part in the study we are presenting. They were exposed to a recall session where they were tested for all 26 letters of English Alphabet. Additionally, they were tested how well they recognised letters of Alphabet in background while doing another primary task. For testing, while performing another task, we selected only 10 letters. Note that, participants were already trained to recognise all letters and the rationale to use only 10 letters for this test was simply to keep the study short and yet try different levels of difficulties in the primary task.

3.1 Pre Study Training

Prior to this study, participants were exposed to 5 sequent days (≈5h in total) of training and testing in letters and words. During the training, participants learned the entire English Alphabet and were able to interpret words of 2-5 letters. We followed the study program of Luzhnica et al. [16] and thus for details of the training we encourage the reader to review [16]. For simplicity, we will refer to this training phase as pre-training throughout the paper. Please note that this pre-training was another user study on its own (see the User Study 2 in [15]), which aimed to evaluate the performance on vibrotactile skin reading of different layouts, identify systematic errors, optimise the layout and also optimise the alphabet encoding. However, this is out of the scope of this paper, and thus we will not further discuss it. For a complete overview and details of the pre-training we encourage the reader to refer to [15].

3.2 Patterns and Wearable Layout

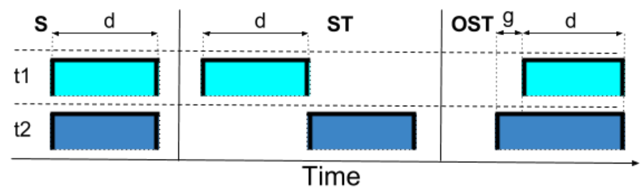


Figure 1: Patterns composed of two vibromotors/locations: spatiotemporal (ST), overlapping spatiotemporal (OST), spatial (S). Base duration (*d*) represents the activation time of a vibromotor (t1 and t2). The gap between the activation of vibromotors is denoted by *g*.

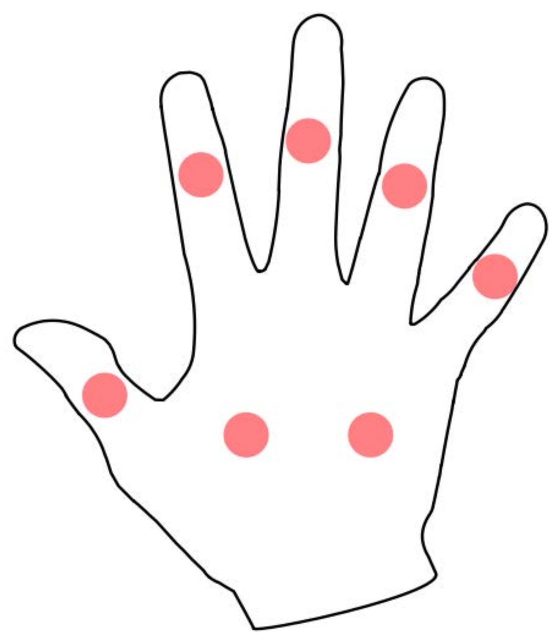


Figure 2: The hand based layout containing seven vibromotors (on the back of the hand) used for the study.

We use a hand based layout with seven vibromotors using locations suggested by Luzhnica and Veas [15] as illustrated in Figure 2. The vibromtors are placed on the back of the hand including the fingers. For the participants that take part in this study, the same layout and encoding of symbols was also used during the pre-training.

Each symbol is encoded with one or two vibromotors using an OST (overlapped spatiotemporal) stimulation as introduced by Luzhnica and Veas. [12, 16]. The encoding of each letter used in this

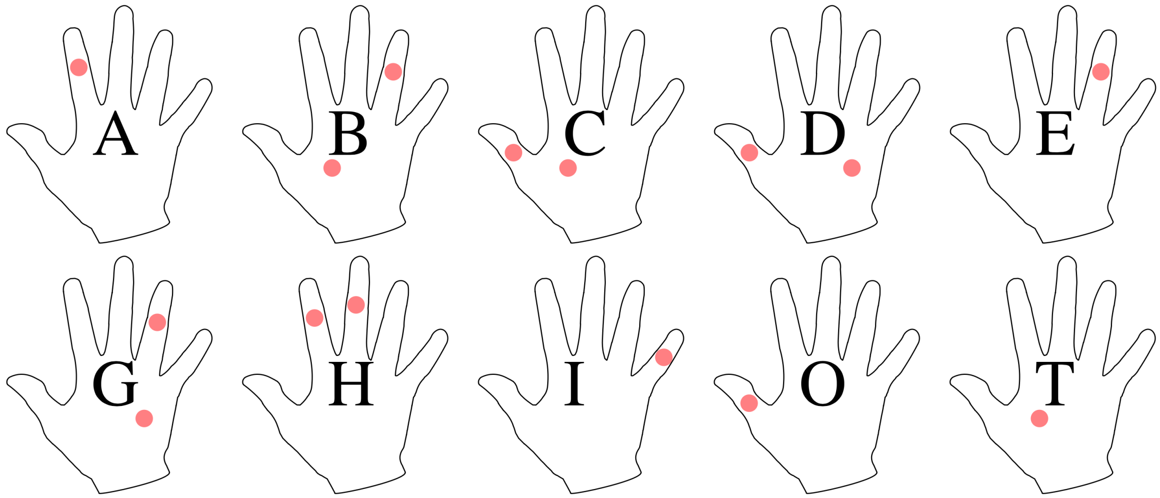


Figure 3: The encoding of symbols (letters) used for the study.

study is illustrated in Figure 3. Figure 1 illustrates the details of OST and the differences with other stimulation forms. Symbol encoding uses a base duration (d) of 100 ms and a 10 ms gap (g) between the activation of vibromotors. This means that the duration (ld) is 100 ms for one vibromotor symbols and 110 ms for two vibromotor encoded symbols.

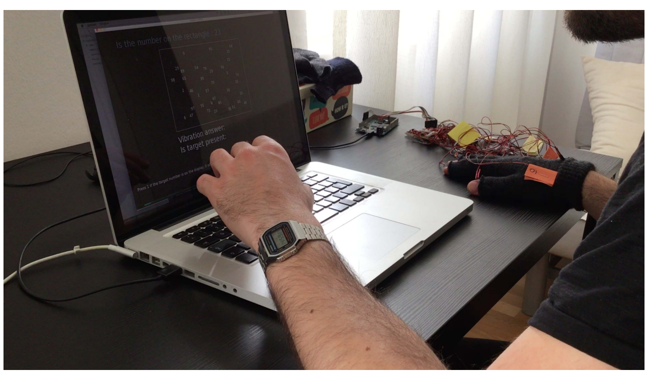


Figure 4: A participant performing the user study

3.3 Participants

Seven participants (six males and one female) aged between 21 and 34 years old participated in this experiment.

3.4 Apparatus

Our device consisted of an Arduino Due board which controls $3.4mm$ vibrotactile motors of type ROB-08449 (Voltage range: $2.3V \sim 3.6V$; Amplitude vibration: $0.8G$).

3.5 Procedure

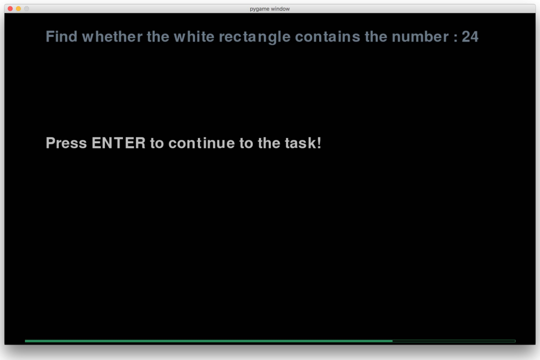
Participants were equipped with the device. Initially, they were exposed to a round of vibrotactile only (**VBO**) testing with all 26 letters to measure the recall accuracy. Here, participants were stimulated with a vibrotactile pattern corresponding to a letter and asked to provide the answer, and they were not performing any other task during this procedure. After that, they continued with the multitasking (primary and background task) study where first they were exposed to a trial phase and then finally continued with the tasks that were recorded to evaluate their performance.

During the study, participants were asked to solve a visual search as a primary task. During this task, participants initially were presented with an integer representing a **search target**. Then a new screen was presented to them with a set of integers within a box which we will refer to as **search set**. Their task was to determine whether the search target was within the search set. The screen with search set was visible only for 5 seconds during which participants could provide the answer.

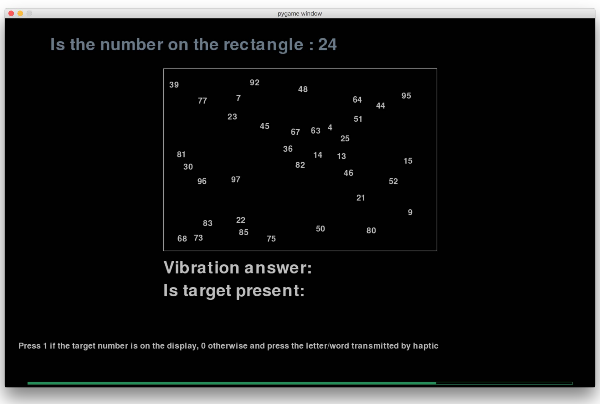
Additionally, in 50% of the cases, 0.5 seconds after the screen with the search set appeared a vibrotactile stimulus representing one of ten selected letters was presented to the participant and they needed to solve both visual search task and recognise the vibrotactile stimuli (**VSVB**) as a secondary/background task. In the other 50%, only visual search task was required (no vibrotactile stimuli) to solve (**VSO**). Participants were not informed priory whether the task contains a vibrotactile stimulus or not.

Participants had the chance to repeat the vibrotactile stimuli by pressing the SPACE bar as long as the search set screen was visible and they could also provide the answer during this time. After that, the visual set screen was replaced with a new one, where participants had the chance to provide or change the answer, but they were not able to neither see the visual search set nor repeat vibrotactile stimuli. To provide the answer for visual search task, participants used keyboard numbers 0 (no) and 1 (yes), whereas to provide the answer for vibrotactile stimuli, they typed the letter representing the stimuli.

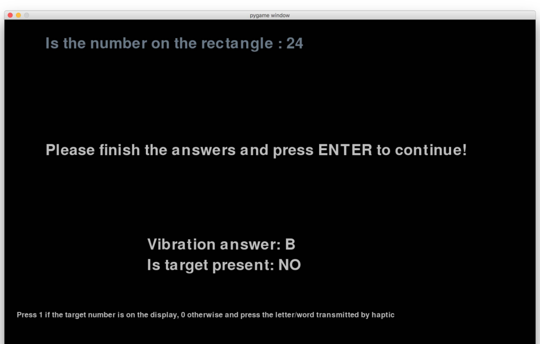
There were three difficulties of visual search task which was determined by the number of integers (size) within the search set. We used 9, 25 and 36 integers as such numbers have been used in the



(a) Presentation of the search target



(b) Visual search task screen



(c) Response providing screen

Figure 5: Visual search task user interface. First, the search target was presented on the screen (a). Then the visual search screen was presented for 5 seconds (b). While this screen was active, participants were stimulated with a vibrotactile cue in 50% of the tasks. In the response screen (c), participants could still provide/change their answer.

past [9] and suggested as three appropriate levels of difficulties. The position of the integers was randomly assigned. Additionally, the integers within the search set were unique and randomly selected from 1 to 99.

In the visual search and background vibrotactile tasks, each participant was exposed to 10 letters x 3 difficulties x 2 task type (VSO and VSVB) = 60 probes during the study. For seven participants, we collected 420 probes in total, 210 for each task type (70 for each difficulty). During the trial phase, they were exposed to 2 letters x 3 difficulties x 2 stimulation = 12 probes, but the responses were not recorded. In 50% of trials (balanced by difficulty and stimulation), the visual target was present in the visual search set. The entire probes appeared in random order. In the vibrotactile only test (VBO) we collected 26 probes (each letter once) for each participant. However, we will use the collected data only for the ten letters that are used in the visual search and background vibrotactile tasks. Thus for 7 participants, for this study analysis, we will use 70 probes (7 participants x 10 letters).

The entire session took around 20 minutes (3-5 minutes vibrotactile only test, 3-5 minutes trial mode and 10-14 minutes the main study), although it varied as it depended on how fast participants responded. At the end, users filled a NASA TLX questionnaire where they were asked to self-assess the tasks in three different categories: (i) vibrotactile only (VBO) - where there was no visual search task, (ii) visual only (VSO) - where there were no vibrotactile stimuli during the visual search task and (iii) both (VSVB).

Note that, this study uses only 10 letters to keep the session short and avoid fatigue. Given that, when we encode the entire Alphabet (26 letters) [15], we have only use letters encoded by one or two vibromotors, we took 5 random letters with one vibromotor and 5 randomly chosen with two. Moreover, participants were not informed that only 10 letters would be used (or which ones) to avoid reducing the workload. Given that they were previously trained on 26 letters (in pre-training), they were prepared to respond for 26 letters of the English Alphabet during the VSVB tasks as well.

It is also worth noting that, both **VSO** and **VSVB** tests are unique to this study. Participants were not exposed to such tests during the pre-training.

3.6 Results

3.6.1 Performance. The recall test had dual purposes. But for the scope of this paper, we will include in the results only the ten letters that are used in primary/secondary task. Thus, we will ignore the responses of the rest of 16 letters. Furthermore, we define the following independent variables for the analysis of the data:

- Task type which takes values: **VBO** (vibrotactile only), **VBO** (visual only) and **VSVB** (visual and vibrotactile).
- Difficulty which is determined by the size of search set and takes values: 9, 25, 36 (for VSO and VSVB). We occasionally use the value 0 (e.g. in Figure 6) which indicates that there was no visual search task and thus representing VBO.

Additionally, we define the depended variables to be:

- Vibrotactile accuracy which represents the recognition accuracy letters represented by a vibrotactile stimulus.
- Visual accuracy which represents the accuracy on finding whether the search target was present in the search set during the visual search tasks.

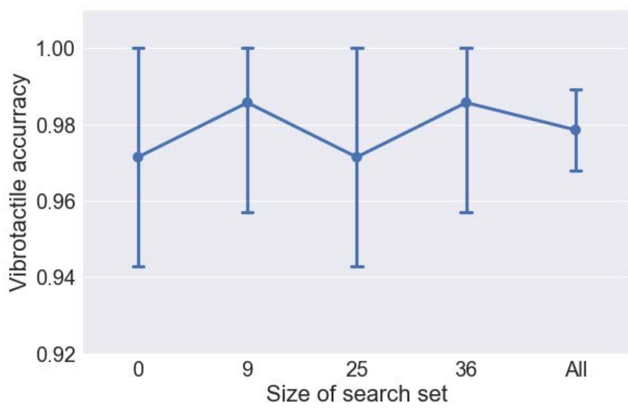


Figure 6: Vibrotactile letter recognition accuracy. The data is categorised depending on search set size and for all tasks combined. The set size of 0 indicates the VBO condition where no visual search task was present. Additionally, 'All' represents all set sizes combined (9, 25 and 36).

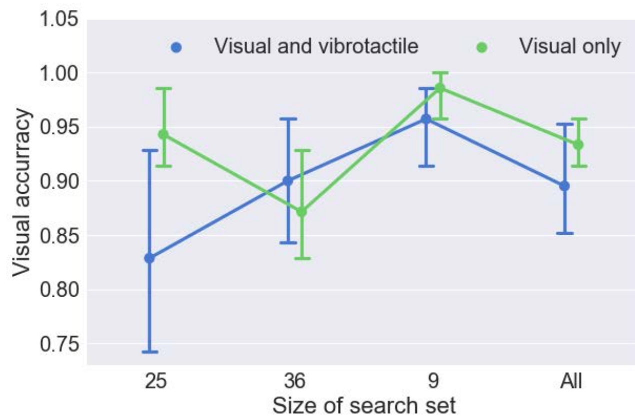


Figure 7: Visual search task accuracy depending on whether there was a vibrotactile task. The data is categorised depending on search set size and for all sizes combined ('All').

In our examination, first, we compare how the primary task affects the recognition of vibrotactile encoded symbols. For this, we compare the performance of the users on recognising letters in vibrotactile only (VBO) rounds with the recognition of letters as a secondary task while they had to perform the visual search task (VSVB) as a primary task. The results are visualised in Figure 6 which show that recognition accuracy is quite the same regardless of whether users were performing another primary task and its difficulty. The recognition accuracy is a binary variable set to be 1 if the participant recognised the letter or 0 otherwise. Given that we are also dealing with repeated measurements, we use McNemar's test in order to test for statistical significance between groups.

Overall, participants were able to recognise letters with a comparable and very high accuracy in both VBO ($\mu = 0.97, \sigma = 0.17$)[1] and VSVB ($\mu = 0.98, \sigma = 0.14$) conditions. When comparing particular difficulty levels of visual search task in VSVB with VBO, according to McNemar's tests, the differences in letter recognition are not significant in neither of the levels:

(1) Search size set of 9: **VBO** ($\mu = 0.97, \sigma = 0.17$) vs **VSVB** ($\mu = 0.99, \sigma = 0.12$); $\chi^2(1, N = 140) = 0.0, p = 1.0$
(2) Search size set of 25: **VBO** ($\mu = 0.97, \sigma = 0.17$) vs **VSVB** ($\mu = 0.97, \sigma = 0.17$); $\chi^2(1, N = 140) = 0.25, p = 0.62$
(3) Search size set of 36: **VBO** ($\mu = 0.97, \sigma = 0.17$) vs **VSVB** ($\mu = 0.97, \sigma = 0.12$); $\chi^2(1, N = 140) = 0.0, p = 1.0$.

We also examine the repetition of vibrotactile stimuli as this could be an indication of the difficulty of perceiving the vibrotactile message. Depending on whether there was a primary visual search task and the level of difficulty (search set size), between 4% and 6% of letters were repeated. McNemar's tests reveal that the differences are insignificant between any of the groups. The groups were compared similarly as in the case of the letter recognition accuracy (see above).

Moreover, we analyse how the presentation of vibrotactile stimuli as a background/secondary task did affect the primary visual search task. Thus, we compare the visual search task accuracy when participants were not stimulated with the vibrotactile message (VSO) with the tasks where they were required to solve both visuals search task and recognise the message/letter (VSVB). Overall, participants were able to achieve slightly better performance on solving visual search task when they did not have any other secondary task (VSO) ($\mu = 0.93, \sigma = 0.25$) compared to when they did have a vibrotactile secondary task (VSVB) ($\mu = 0.9, \sigma = 0.31$).

To determine the statistical significance, the differences in among each level of difficulty will be analysed. When looking at particular levels of difficulty, for the set size of 9 and 36 there are only slight differences (in both directions) whereas for the search size of 25 participants were able to solve visual search task better when they did not have a secondary vibrotactile task. Nevertheless, performing McNemar's tests between conditions (VSO vs VSVB) reveal that regardless of the level of difficulty, there are no statistically significant differences in the performance of solving visual search task between VSO and VSVB:

(1) Search size set of 9: **VSO** ($\mu = 0.99, \sigma = 0.12$) vs **VSVB** ($\mu = 0.96, \sigma = 0.2$); $\chi^2(1, N = 140) = 0.25, p = 0.62$
(2) Search size set of 25: **VSO** ($\mu = 0.94, \sigma = 0.23$) vs **VSVB** ($\mu = 0.83, \sigma = 0.38$); $\chi^2(1, N = 140) = 3.5, p = 0.061$
(3) Search size set of 36: **VSO** ($\mu = 0.87, \sigma = 0.34$) vs **VSVB** ($\mu = 0.9, \sigma = 0.3$); $\chi^2(1, N = 140) = 0.06, p = 0.80$

3.7 Questionnaire

The results of NASA TLX are presented in the Figure 8. In addition to the six metrics contained in NASA TLX, we calculated the workload using the simplified R-TLX method (averaging all metrics where the performance is inverted). For analysis, we compare the workload between three different task type conditions, namely VBO vs VSVB and VSO vs VSVB.

Considering that the workload of each task normally distributed (Shapiro-Wilk: p > 0.05) and the variances of each compared groups are homogenous (Levene: p > 0.05) for each compared pairs, we

[1] We denote mean of a group by μ and standard deviation by σ.

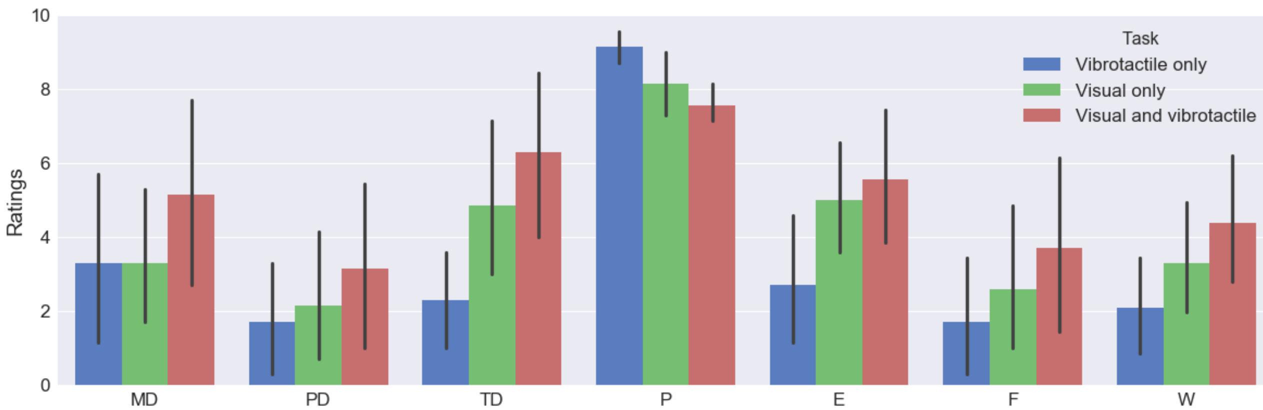


Figure 8: NASA TLX self evaluation metrics for the given tasks. Metrics notation : MD - mental demand, PD - physical demand, TD - temporal demand, P - performance, E - effort, F - frustration and W - workload.

use the paired t-test for determining whether there is a significant difference in workload. A paired t-test analysis reveals that the workload for VBO task ($\mu = 2.1, \sigma = 1.94$) was significantly lower compared to the workload of VSVB task ($\mu = 4.38, \sigma = 2.65$); $t(14) = -5.69, p = 0.001$. Additionally, the workload for VSO task ($\mu = 3.29, \sigma = 2.26$) was lower compared to the workload of VSVB task ($\mu = 4.38, \sigma = 2.65$); but not enough to be considered significant; $t(14) = -1.65, p = 0.15$.

4 DISCUSSION AND CONCLUSION

Our study was designed to investigate the perception and comprehension of tactons (representing letters of English Alphabet) in the background (as a secondary task) while performing another attention demanding primary task. For encoding we use overlapping spatiotemporal patterns [12, 16] and use a very high speed of transmission where vibration patterns for a symbol take 100-110 ms. We recruit only previously trained users who are proficient in recognising such tactons. This way, we hypothesised that the recognition of such tactons would be processed using automatic cognitive processing and thus performed effortlessly and without interfering with the primary task. We use attention demanding visual search tasks with three difficulties as primary tasks during our user study.

Our results show the recognition of tactons was very accurate (≥ 97%) regardless of whether this was done as a single task or as a secondary task along the visual search task (see Figure 6). Similarly, the visual search tasks did not have any effect on the repetition of vibrotactile stimuli. Second, the performance of the primary visual search task does not seem to be significantly deteriorated by the presence and recognition of tactons in the background. Although, on average the performance did slightly decrease. As shown in Figure 7, clearly such deterioration seem not to be accelerated with the increase of primary task difficulty. For instance, when the search set size was 36, which represents the highest difficulty, participants even performed better on the visual search task when having to simultaneously recognise a tacton in the background compared to the cases when they did not perform any task in the background. Additionally, the self-assessed workload (see Figure 8) did also not significantly increased when users performed the visual search task along with a parallel background vibrotactile task.

The results of our study evdience that transmitting information through vibrotactile wearable devices through overlapping spatiotemporal patterns, even at very high speed, is very efficient and can be used along with other user cognitive activities. As such, it provides many opportunities to support and facilitate multitasking especially considering that we are using wearable devices. It suggests that information conveyed using vibrotactile wearable devices can be effortlessly comprehended while performing other tasks as well, assuming that users are adequately trained to recognise and associate the vibrotactile patterns.

We use tactons which represent letters of the English Alphabet. However, such tactons could represent any other abstract meaning such as commands, warnings, states etc.. as long as users are trained to recognise them. Additionally, we use only 10 tactons in order to keep the study short and avoid fatigue. However, participants were already trained in a pre-training period (as part of another study [15]) to recognise the entire Alphabet (26 letters). During our user study, participants were not aware that they would be tested on only 10 symbols/letters and which of them will be used. Thus, we argue that they would have had the same cognitive load if they were tested on all 26 symbols/letters in multitasking experiment.

Although 10 or even 26 tactons might be limiting for all use cases, yet they can encode sufficient information in a lot of use cases (commands, warnings, states, etc..). On the other hand, when perceiving vibrotactile encoded information, it has already been demonstrated that individuals are able to understand words as a series of letters [13, 15, 16] or phonemes [35] as a primary task (without any other parallel task). It would be certainly very useful if users would be able to perceive such complex messages (words, sentences) composed of several tactons in the background as the application possibilities for multitasking would broaden drastically.

5 FUTURE WORK

The insights gathered in this work will serve as foundation to further investigate the comprehension of textual information in background transmitted by wearable vibrotactile displays. One limitation of this work is that we did not test participants on comprehending words in background. Exposing them to words for longer periods would be necessary in order for them to be able to read words as units and thus get enough proficiency to be able to comprehend the words using automatic processing cognitive model. In other forms of reading such as visual, it is a well-established theory, that fast reading is attributed to words being read as units instead of letter by letter [8, 18]. Such a word recognition as a unit is achieved through exposure to words (practice) [8, 31]. Analogously, for vibrotactile encoded information, we think that for testing in the background, users should be at a stage of proficiency where they would interpret words as units. Therefore, we plan to conduct such longitudinal user studies in our future work.

ACKNOWLEDGMENTS

This work is partially funded by MoreGrasp (H2020-ICT-2014-1 643955). The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

REFERENCES

[1] J. C. Bliss, M. H. Katcher, C. H. Rogers, and R. P. Shepard. 1970. Optical-to-Tactile Image Conversion for the Blind. *IEEE Transactions on Man-Machine Systems* 11, 1 (March 1970), 58–65. https://doi.org/10.1109/TMMS.1970.299963

[2] Stephen Brewster and Lorna M. Brown. 2004. Tactons: Structured Tactile Messages for Non-visual Information Display. In *Proceedings of the Fifth Conference on Australasian User Interface (AUIC)*. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 15–23. http://dl.acm.org/citation.cfm?id=976310.976313

[3] K L Galvin, G Mavrias, A Moore, R S Cowan, P J Blamey, and G M Clark. 1999. A comparison of Tactaid II+ and Tactaid 7 use by adults with a profound hearing impairment. *Ear and hearing* 20 6 (1999), 471–82.

[4] Robert H Gault. 1924. Progress in experiments on tactual interpretation of oral speech. *The Journal of Abnormal Psychology and Social Psychology* 19, 2 (1924), 155.

[5] Frank A. Geldard. 1957. Adventures in tactile literacy. *American Psychologist* 12, 3 (1957), 115–124.

[6] Eric Gunther. 2001. *Skinscape: A tool for composition in the tactile modality.* Ph.D. Dissertation. Massachusetts Institute of Technology.

[7] Jacob H. Kirman. 1974. Tactile perception of computer-derived formant patterns from voiced speech. *The Journal of the Acoustical Society of America* 55, 1 (1974).

[8] Kevin Larson. 2004. The Science of Word Recognition. *Advanced Reading Technology, Microsoft Corporation* (July 2004). https://www.microsoft.com/typography/ctfonts/WordRecognition.aspx

[9] Seungyon "Claire" Lee and Thad Starner. 2010. BuzzWear: Alert Perception in Wearable Tactile Displays on the Wrist. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 433–442. https://doi.org/10.1145/1753326.1753392

[10] Yi-Chi Liao, Yi-Ling Chen, Jo-Yu Lo, Rong-Hao Liang, Liwei Chan, and Bing-Yu Chen. 2016. EdgeVib: Effective Alphanumeric Character Output Using a Wrist-Worn Tactile Display. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 595–601. https://doi.org/10.1145/2984511.2984522

[11] Granit Luzhnica, Sebastian Stein, Eduardo Veas, Viktoria Pammer, John Williamson, and Roderick Murray Smith. 2017. Personalising Vibrotactile Displays Through Perceptual Sensitivity Adjustment. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. ACM, New York, NY, USA, 66–73. https://doi.org/10.1145/3123021.3123029

[12] Granit Luzhnica and Eduardo Veas. 2017. Vibrotactile Patterns using Sensitivity Prioritisation. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. ACM, New York, NY, USA, 148–155. https://doi.org/10.1145/3123021.3123032

[13] Granit Luzhnica and Eduardo Veas. 2018. Investigating Interactions for Text Recognition Using a Vibrotactile Wearable Display. In *23rd International Conference on Intelligent User Interfaces (IUI '18)*. ACM, New York, NY, USA, 453–465. https://doi.org/10.1145/3172944.3172951

[14] Granit Luzhnica and Eduardo Veas. 2018. Skin Reading Meets Speech Recognition and Object Recognition for Sensory Substitution. In *Proceedings of the 2018 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2018 ACM International Symposium on Wearable Computers (UbiComp '18)*. ACM, New York, NY, USA. https://doi.org/doi.org/10.1145/3267305.3267604

[15] Granit Luzhnica and Eduardo Veas. 2019. Optimising the Encoding for Vibrotactile Skin Reading. In *ACM CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA. https://doi.org/10.1145/3290605.3300465

[16] Granit Luzhnica, Eduardo Veas, and Viktoria Pammer. 2016. Skin Reading: Encoding Text in a 6-channel Haptic Display. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. ACM, New York, NY, USA, 148–155. https://doi.org/10.1145/2971763.2971769

[17] Granit Luzhnica, Eduardo Veas, and Caitlyn Seim. 2018. Passive Haptic Learning for Vibrotactile Skin Reading. In *Proceedings of the 2018 ACM International Symposium on Wearable Computers (ISWC '18)*. ACM. https://doi.org/10.1145/3267242.3267271

[18] Susanna Millar. 2004. *Reading by Touch in Blind Children and Adults.* Springer Netherlands, Dordrecht, 437–457. https://doi.org/10.1007/978-94-017-1731-1_23

[19] P. Milnes, J. C. Stevens, B. H. Brown, I. R. Summers, and P. G. Cooper. 1996. Use of micro-controller in a tactile aid for the hearing impaired. In *Proceedings of 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 1. 413–414 vol.1. https://doi.org/10.1109/IEMBS.1996.657019

[20] Hugo Nicolau, João Guerreiro, Tiago Guerreiro, and Luís Carriço. 2013. UbiBraille: Designing and Evaluating a Vibrotactile Braille-reading Device. In *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '13)*. ACM, New York, NY, USA, Article 23, 8 pages. https://doi.org/10.1145/2513383.2513437

[21] Scott D. Novich and David M. Eagleman. 2015. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. *Experimental Brain Research* 10 (2015). https://doi.org/10.1007/s00221-015-4346-1

[22] Parivash Ranjbar. 2008. Vibrotactile Identification of Signal-Processed Sounds from Environmental Events Presented by a Portable Vibrator: A Laboratory Study. *Iranian Rehabilitation Journal* 6 (2008). arXiv:http://irj.uswr.ac.ir/article-1-274-en.http://irj.uswr.ac.ir/article-1-274-en.html

[23] Charlotte M Reed and Lorraine A Delhorne. 2003. The reception of environmental sounds through wearable tactual aids. *Ear and hearing* 24, 6 (2003), 528–538.

[24] Jerker Rönnberg, Ulf Andersson, Björn Lyxell, and Karl-Erik Spens. 1998. Vibrotactile Speech Tracking Support: Cognitive Prerequisites. *The Journal of Deaf Studies and Deaf Education* 3, 2 (1998), 143–156. https://doi.org/10.1093/oxfordjournals.deafed.a014344

[25] Walter Schneider, Sue T Dumais, and Richard M Shiffrin. 1982. *Automatic/Control Processing and Attention.* Technical Report. ILLINOIS UNIV CHAMPAIGN HUMAN ATTENTION RESEARCH LAB.

[26] Walter Schneider and Richard M Shiffrin. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological review* 84, 1 (1977), 1.

[27] Caitlyn E. Seim, David Quigley, and Thad E. Starner. 2014. Passive Haptic Learning of Typing Skills Facilitated by Wearable Computers. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 2203–2208. https://doi.org/10.1145/2559206.2581329

[28] Ian R Summers, Jon J Whybrow, Denise A Gratton, Peter Milnes, Brian H Brown, and John C Stevens. 2005. Tactile information transfer: A comparison of two stimulation sites. *The Journal of the Acoustical Society of America* 118, 4 (2005), 2527–2534.

[29] David Ternes and K E MacLean. 2008. Designing Large Sets of Haptic Icons with Rhythm. In *EuroHaptics*, M. Ferre (Ed.). Springer, Madrid, Spain, 199–208. http://www.cs.ubc.ca/labs/spin/publications/spin/ternes2008eh.pdf

[30] James S Uleman and John A Bargh. 1989. *Unintended thought.* Guilford Press.

[31] D. H. Whalen. 1991. Infrequent words are longer in duration than frequent words. *The Journal of the Acoustical Society of America* 90, 4 (1991), 2311–2311. https://doi.org/10.1121/1.401072 arXiv:https://doi.org/10.1121/1.401072

[32] Benjamin W. White, Frank A. Saunders, Lawrence Scadden, Paul Bach-Y-Rita, and Carter C. Collins. 1970. Seeing with the skin. *Perception & Psychophysics* 7, 1 (1970), 23–27. https://doi.org/10.3758/BF03210126

[33] Cheng Xu, Ali Israr, Ivan Poupyrev, Olivier Bau, and Chris Harrison. 2011. Tactile Display for the Visually Impaired Using TeslaTouch. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. ACM, New York, NY, USA, 317–322. https://doi.org/10.1145/1979742.1979705

[34] Hanfeng Yuan, Charlotte M. Reed, and Nathaniel I. Durlach. 2005. Tactual display of consonant voicing as a supplement to lipreading. *The Journal of the Acoustical Society of America* 118, 2 (2005), 1003–1015. https://doi.org/10.1121/1.1945787 arXiv:http://dx.doi.org/10.1121/1.1945787

[35] Siyan Zhao, Ali Israr, Frances Lau, and Freddy Abnousi. 2018. Coding Tactile Symbols for Phonemic Communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA. https://doi.org/10.1145/3173574.3173966