Learning to Feel Words: A Comparison of Learning Approaches to Acquire Haptic Words

Jennifer Chen, Robert Turcott, Pablo Castillo, Wahyudinata Setiawan, Frances Lau, Ali Israr
Facebook, Inc.
1 Hacker Way, Menlo Park, CA 94025
USA
{jenniferchen, rturcott, pablocastillo, wahudinata, flau, aliisrar}@fb.com

ABSTRACT

Recent studies have shown that decomposing spoken or written language into phonemes and transcribing each phoneme into a unique vibrotactile pattern enables people to receive lexical messages on the arm. A potential barrier to adopting this new communication system is the time and effort required to learn the association between phonemes and vibrotactile patterns. Therefore, in this study, we compared the learnability and generalizability of different learning approaches, including guided learning, self-directed learning, and a mnemonic device. We found that after 65 minutes of learning spread across 3 days, 67% of participants, including both native and non-native English speakers, following the guided learning could identify 100 haptic words with over 90% accuracy, while only 20% of participants using the self-directed learning paradigm could do so.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)** → HCI design and evaluation methods

ADDITIONAL KEYWORDS AND PHRASES

Vibrotactile, haptic language, learning, transcutaneous language communication, speech-to-touch

ACM Reference format:
J. Chen, R. Turcott, P. Castillo, W. Setiawan, F. Lau, and A. Israr. 2018. Learning to feel words: A comparison of different learning approaches to acquire haptic words. In *Proceedings of ACM Symposium on Applied Perception, Vancouver, British Columbia Canada, August 2018 (SAP'18)*, 8 pages.

1 INTRODUCTION

Various efforts have been made to enable people with impaired vision or hearing to communicate through touch and vibration since the 1800s. For example, visually impaired people can touch raised dots (Braille) to read; and deaf-blind people can use the Tadoma method [1], in which the thumb is placed on a speaker's lips and the fingers are placed on the jaw and throat, to perceive speech based on vibration and movement.

On the other hand, research on transcutaneous communication (TLC) in people with normal vision and hearing started later. In 1924, Gault reported the first case of a male participant with normal vision and hearing who attained 88% – 95% accuracy recognizing 34 spoken words with his palm based on the vibrations of a speaking tube after 45.5 hours of flashcard training [15]. Geldard (1957) demonstrated that 3 participants achieved 75% – 100% accuracy in identifying 26 letters presented in vibrotactile format on their chest after 12 hours of flashcard training [16] and 1 participant reached 90% accuracy in sentence comprehension after 65 hours of flashcard training [17]. A female participant identified 250 words with 75.6% accuracy after 80.5 hours of flashcard training using a different device (Queen's Tactile Vocoder) which divides the acoustic waveform into multiple channels that are independently processed to activate specific vibrators on a haptic display worn on the ventral forearm [5, 7]. Although these initial attempts were encouraging, the limited sample size and long training times were barriers to the widespread adoption of TLC in people with normal vision and hearing.

Efforts to remove these barriers progressed slowly, but momentum has increased with the widespread adoption of mobile and wearable devices, and the desire to offload the heavily taxed visual and auditory modalities. In one study, 12 participants were able to identify the meanings of 9 vibrotactile patterns with close to 80% accuracy after 45 min of self-guided learning [14]. Moreover, with 30 minutes of flashcard training, 24 participants could identify 26 letters and 10 digits with 85.9% and 88.6% accuracy, respectively, using a spatial-temporal vibrotactile device (EdgeVib) [24]. One drawback of mapping letters to vibrotactile patterns is that it cannot be applied to languages that do not use letters, like Chinese. In contrast, a phonemic approach is more generalizable and has higher speech-to-touch information transfer rate (i.e., on average, 3.34 phonemes per word [23] vs. 5 letters per word [4]).

In a recent work, Zhao and colleagues proposed a "phonemic" approach, in which each phoneme is paired with an unique vibrotactile pattern based on place-voice-manner of articulation, that demonstrated 83% recall accuracy for 10 words after 28 minutes of self-guided and flashcard training [34]. Retention and generalization were encouraging, achieving 55% recall accuracy for new words, and after a short review of the

---

SAP '18, August 10–11, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5894-1/18/08…$15.00
https://doi.org/10.1145/3225153.3225174

word list, the scores increased to 88%. Acoustic coding algorithms, used in the Queen's Tactile Vocoder [5–7] and vibrotactile vests [12, 29], had poorer recognition and retention of learned words and worse generalization to new words [32].

The most commonly used learning approach in TLC was a traditional *flashcard* technique; where the correct answer was provided immediately after learners responded to each vibrotactile pattern [5, 7, 16, 24, 34]. Yet, studies have shown that compared to traditional flashcard techniques, *incremental rehearsal*, a method that intersperses 90% known and 10% unknown words in a flashcard method, gave better acquisition, retention, and generalization [28, 31]. Based on these promising results, we incorporated *incremental rehearsal* in TLC training.

Another training approach that has been reported in TLC is *self-guided* learning, where the learners initiated the playback of the corresponding vibrotactile patterns using their own pace [14, 34]. *Self-guided* learning shifts the learning responsibility to the learners themselves, making them more motivated and engaged, and are more likely to appreciate the learning materials [8, 35]. However, the success of *self-guided* learning largely depends on learners' ability to monitor their own learning process, such as what to study, how to learn, and how to allot the time, etc. [21, 22]. Furthermore, *self-guided* learning is better for people with greater experience, whereas for novices, *guided* learning is more effective, leading to less errors and higher performance [10].

Additionally, learning through *video game play* has drawn a lot of attention recently. Its motivational, interactive, and multimodal elements make it fun and engaging, so learners are more likely to spend time mastering the skills [11]. Studies also showed that video games promoted language acquisition and non-native speech sound categorization [14,19]. Hence, we included *incremental rehearsal* and *video game play* with our *guided* learning paradigm. Based on aforementioned literatures, there is a reason to believe such design is more helpful for beginners to learn TLC than the traditional *flashcard* and *self-guided* learning approaches.

In this paper, we generated a set of vibrotactile patterns that are perceptually distinct and easily recognizable. These patterns are associated with phonemes based on the place-voice-manner of articulation [34]. We compared different learning approaches, in terms of recall and generalization accuracy. Generalization indicates the ability to identify new words based on learned knowledge. This is critical for reducing learning time. We first compared two learning paradigms: guided learning, which consisted of incremental rehearsal, flashcards, explore mode, and a video game in fixed time and order, analogous to a curriculum; and self-guided learning, which had flashcards and explore mode, where participants could navigate freely at their own pace. Finally, we tested if being consciously aware of the coding rule – place, voice, and manner of articulation – facilitated learning.

2 Methods

2.1 Haptic Stimuli

The haptic stimuli consisted of rough and smooth single taps at the wrist, elbow, and upper-arm; and slow and fast motion cues in different directions on the forearm. The rough and smooth percepts were respectively generated by intermittent 15-ms-long bursts of 250 Hz sinusoidal waves every 27 ms and 250 Hz continuous sinusoidal waves. We limited the single taps to the wrist and elbow because past studies showed that tactors placed closer to body landmarks were localized more accurately [9]. 267-ms-long slow and 120.6-ms-long fast motion cues, both including 10% ramp time, were respectively generated by four 102-ms-long tactor activations with 55 ms stimulus onset asynchrony (SOA) and four 51.6-ms-long tactor activations with 23 ms SOA.

13 distinctive vibrotactile patterns were designed to map to 13 phonemes (Fig. 1A), including 7 consonants /m/, /n/, /d/, /b/, /t/, /k/ and /s/, and 6 vowels /eɪ/ (**A**te), /ɛ/ (s**E**t), /i/ (**EA**t), /ɪ/ (**I**t), /aɪ/ (h**I**de) and /u/ (m**OO**n). The patterns are associated with phonemes based on the articulation of the phonemes. Consonants are mapped to single-tap vibrations and vowels to moving patterns. Voiced consonants /b/ and /d/ are represented as rough vibrations and unvoiced consonants /t/ and /s/ as smooth vibrations. The location of constriction within the mouth during the production of the consonant is mapped to the location on the arm. The labials /m/, /b/ are mapped closer to the wrist and velars /k/ are mapped on the upper arm. The manner consonants are produced is mapped to the side of the arm. Nasals /m/ and /n/ are presented on the volar side of the arm while stops and fricatives (/b/, /s/, etc.) are presented on the dorsal side. In vowels, short and long vowels are presented as fast and slow motion patterns. /ɛ/ (s**E**t) and /ɪ/ (**I**t) are represented by a faster version of /eɪ/ (**A**te) and /i/ (**EA**t), respectively. Finally, /u/ (m**OO**n) is designed as rotation around the wrist mimicking the shape of the mouth during articulation.

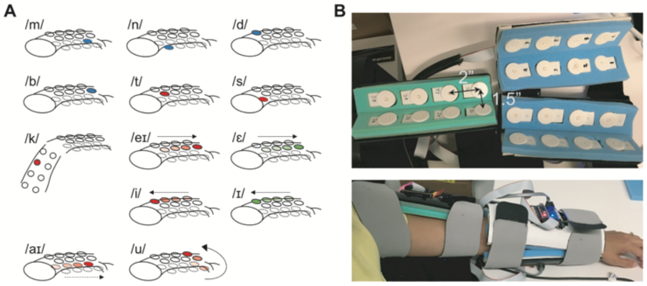


Figure 1: *(A) Phonemes and corresponding vibrotactile patterns based on* place-voice-manner of articulation. *Blue dots represent intermittent bursts of 250 Hz sinusoidal waves, red and green dots represent continuous 250 Hz sinusoidal waves,* with green dots moving at 2x the speed of red dots. *(B) Haptic device used in the study to display vibrotactile patterns.*

Using these 13 phonemes, we generated 100 words with 1 to 3 phonemes per word and separated them into 5 lists (Table 1). The time interval between the phonemes within a word was 200 ms. The study consisted of 3 1-hour sessions, 1 day apart. To reduce the effect of word lists on learning and generalization performance, the two-word lists in the same session had similar: average word frequency, occurrence of each phoneme, and number of words of a given phoneme.

Table 1: *Word lists generated from 13 phonemes*

List 1	aim, day, deed, do, I, mine, moon, my, name, need
List 2	die, dine, knee, made, main, may, me, mood, new, nine
List 3	beam, bee, bet, bike, buy, cake, debt, did, men, miss, say, sick, sit, soon, tune
List 4	base, bed, bit, boom, came, in, key, kick, mean, mess, neck, set, site, sue, tide
List 5	ate, bake, bay, bean, been, bid, bite, boot, case, date, dead, deck, deem, died, dim, eat, it, kid, kiss, kit, mate, meat, met, mice, might, neat, net, nice, night, noon, said, same, seat, see, seek, seem, side, sigh, sign, suit, take, tea, team, teen, ten, tie, tight, time, tube, two

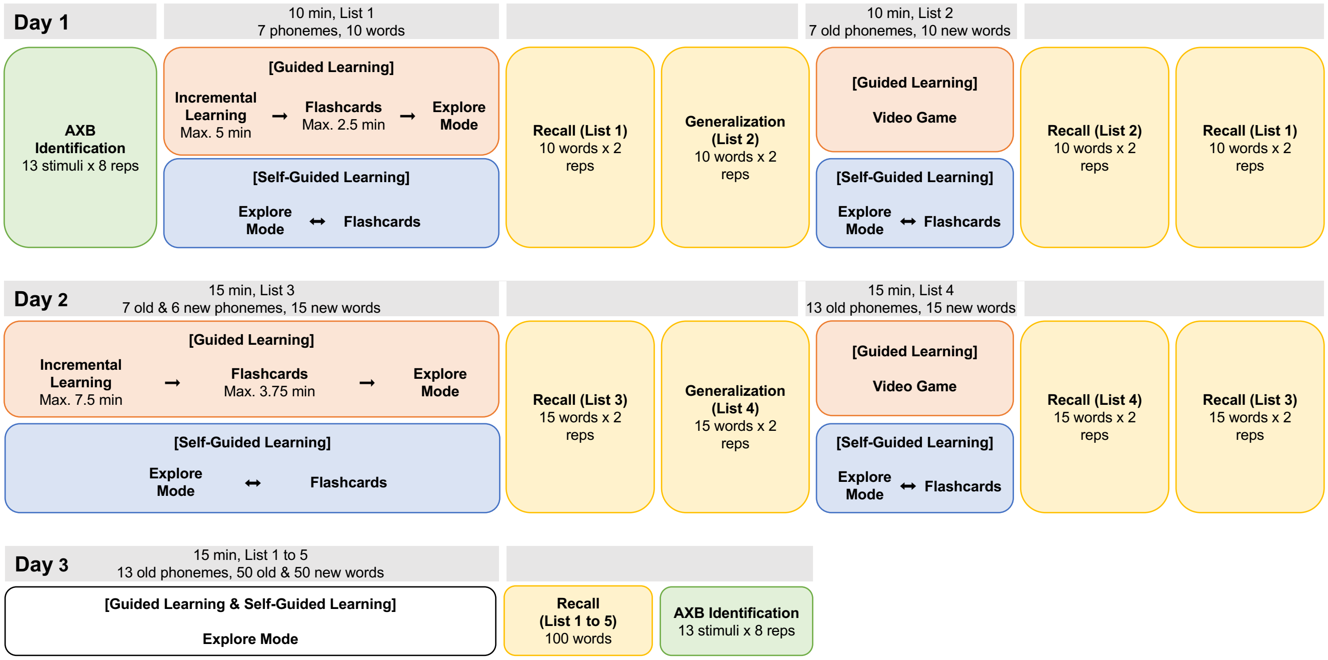


Figure 2: *Illustration of 65 minutes of learning paradigms spread across 3 days.*

2.2 Apparatus

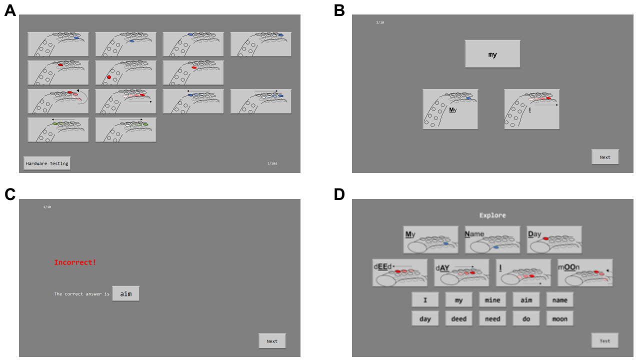
The haptic stimuli were controlled by Max/MSP software (https://cycling74.com/) using a 24-channel audio interface (model 24A0, MOTU Inc., Cambridge MA) and a custom audio-amplifier board, and then were actuated via three custom-designed 8-tactor vibrotactile displays (Fig. 1B). The tactors were voice coils (model: TEAX13C02, Tectonic Elements, UK) surrounded by biocompatible 3-D printed cases, embedded in a medical grade foam sheet, and arranged in a 4 x 2 grid with 1.5" lateral and 2" longitudinal spacing between the tactors. The tactors and foam were attached to a rigid thermoformed arced surface to make a wearable display. The displays were placed on the dorsal and volar sides of the forearm, and the dorsal side of the upper arm. They were cleaned with alcohol wipes between participants.

The system was connected to a medical-grade isolation transformer and the voltage to the wearable device was regulated to 12 V. An emergency stop button was incorporated in the hardware, while the temperature and current were monitored, limited, and alarmed.

2.3 Procedure

We obtained informed consent from 19 naïve participants with normal vision and hearing, with no active skin conditions. As shown in Fig. 2, they were randomly assigned to 2 groups – the guided learning group (6 males, 3 females; 7 native English speakers; average age 27.33 ± 4.74 (SD) years; age range 21 – 37 years old) and the self-guided learning group (7 males, 3 females; 7 native English speakers; average age 31.80 ± 6.60 years; age range 22 – 43 years old). To motivate participants to learn the TLC system, they were rewarded with an extra $25 monetary gift if they reached over 90% accuracy in the last 100 words identification test.

2.3.1 AXB Identification Task. In each trial of AXB identification task, participants were presented with 3 vibrotactile patterns sequentially with a 200 ms inter-stimulus-interval (ISI), and then were required to identify the second vibrotactile pattern (X). The first (A) and last (B) vibrotactile patterns were used to mask the target (X), simulating the identification of a vibrotactile pattern in the middle of a haptic word. All participants performed an AXB identification task at the beginning and at the end of the 3-day training to evaluate the effect of familiarity and experience on the identifiability of vibrotactile patterns. During the AXB identification task, participants had 5 minutes to get familiar with all 13 vibrotactile patterns and their illustrations by initiating the playback of each pattern (Fig. 3A), followed by 3 practice trials with correct answer feedback. Each vibrotactile pattern was tested 8 times in random order, making a total of 104 test trials with a 400 ms inter-trial interval (ITI). No feedback was provided during the test trials.



Incorrect!

The correct answer is aim

Figure 3: *User interface during (A) AXB identification task and participants' familiarity with vibrotactile patterns, (B) incremental rehearsal, (C) flashcards, and (D) Explore mode.*

2.3.2 Guided Learning Paradigm. Both groups learned 100 words across 3 days. Participants in the guided learning group learned the 7 phonemes and 10 words from List 1 on Day 1 and the 6 new phonemes and 15 new words from List 3 on Day 2 through incremental rehearsal, flashcards, and explore mode for 10 minutes on Day 1 and 15 minutes on Day 2. They learned 10 new words from List 2 on Day 1 and 15 new words from List 4 on Day 2 through a video game for another 10 minutes on Day 1 and 15 minutes on Day 2. In each trial of the incremental rehearsal, participants were presented with a 600 ms fixation point, followed by one of the words from List 1 (or List 3 on Day 2) and visual illustrations of the corresponding vibrotactile patterns. The participants had up to 30 seconds to replay the vibrotactile patterns as a word or as individual phonemes as many times as they wanted (Fig. 3B). To implement incremental rehearsal in the learning process, the presentation order of the words was fixed as "I", "my", "aim", "name", "day", "deed", "need", "do", and "moon" (or "buy", "bike", "bee", "beam", "men", "bet", "debt", "did", "miss", "say", "cake", "sick", "sit", "soon", and "tune" on Day 2), so only one new phoneme was introduced during each trial.

Next was the flashcard recall task, in which participants were presented with one of the words from List 1 (or List 3 on Day 2) and were required to type the word. The order of the words was presented randomly. Feedback was provided after the response, and participants could replay the word for up to 15 seconds (Fig. 3C). For the remainder of the 10 minutes (or 15 minutes on Day 2), the participants could playback words and phonemes from List 1 (or List 3 on Day 2) in explore mode (Fig. 3D).

After the guided learning, participants took a recall test of List 1 (or List 3 on Day 2) and a generalization test of List 2 (or List 4 on Day 2). The generalization test documented how well participants could generalize the phonemes they had learned to 10 new words from List 2 (or 15 new words from List 4). The procedure for the recall and generalization tests was like the flashcards, except that no feedback was provided after each response. Each haptic word was tested 2 times in random order, making a total of 20 trials (30 trials on Day 2) in each recall and generalization test.

Subsequently, participants spent 10 minutes learning List 2 via a video game, in which participants' task was to save robots from being dropped into a recycle site by choosing the correct word from the vibrotactile pattern that was paired with each robot (Fig. 4A). There were 10 robots (or 15 robots on Day 2) to be saved at each level. Each vibrotactile pattern was replayed every 2.3 seconds until the participants made a response or the robot was dropped into the recycle site (Fig. 4B). The corresponding spoken word cue was also presented after the first and third playback of each vibrotactile pattern at level 0. As participants progressed to higher levels, the probability of presenting the spoken cue was reduced, making the game more challenging and guiding the participants to rely more on the vibrotactile pattern to make a correct response. At level 10, the highest level, no spoken word cue was presented. Correct answer feedback was always provided when the participants made an incorrect response. At the end of each level: the participants progressed to the next level if accuracy was 70% or above; replayed the same level with if accuracy was 50%-70%; and was levelled down if accuracy was below 50%. The game stopped after 10 minutes (or 15 minutes on Day 2) or when the participant achieved 90% or higher accuracy at the highest level.

Recall tests of List 2 and List 1 (or List 4 and List 3 on Day 2) were administered after stopping the video game using the same procedure described above.

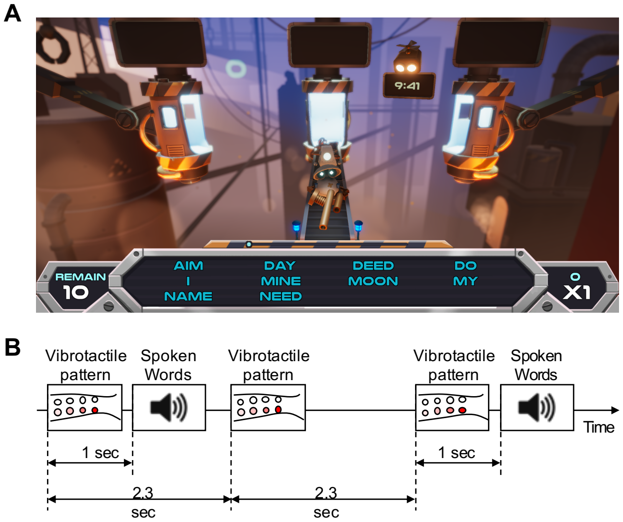


Figure 4: *(A) Video game user interface. (B) Illustration of a learning trial (for a given robot) in the video game.*

2.3.3 Self-Guided Learning Paradigm. The learning procedure on the first two days for the self-guided learning group was like the guided learning group; except that there was no incremental rehearsal and participants could navigate between explore mode and the flashcards described earlier without any time limitations at each trial. The video game from guided learning was replaced with a new explore mode and flashcards for List 2 (or List 4 on Day 2).

The procedure on Day 3 was the same for both guided learning and self-guided learning groups. They were given 15 minutes to explore 50 new words from List 5 along with the 50 old words they had learnt on Day 1 and 2. And then they took a recognition test for the 100 words, in which after the presentation of a vibrotactile pattern, the correct answer and 11 other words that were randomly selected from 100 words appeared on the screen. Participants had to type their answers.

2.3.4 Introduction of Articulation Map. 9 naïve participants (6 males, 3 females; average age 33.78 ± 9.38 years) were invited to only learn List 1 and List 2 using the same materials and procedures as the guided learning group on Day 1; except that the experimenter explained how the phonemes and vibrotactile patterns were designed based on place-voice-manner of articulation to test whether introducing the mnemonic aid could help learning. At the end of the session, participants were asked to rate the helpfulness of the articulation map on a 7-point scale.

2.4 Analyses

We were conservative in estimating the response accuracy. If the response did not completely match the answer, even if misspelling was suspected, then it was counted as an incorrect response. To compare the learnability and generalizability of guided learning and self-guided learning paradigms across time, test results were entered into a mixed ANOVA with learning paradigm (guided learning vs. self-guided learning) as between-subject factor and time of assessment (List 1 recall, List 2 generalization, List 2 recall and the second recall of List 1 on Day 1; List 3 recall, List 4 generalization, List 4 recall and the second recall of List 3 on Day 2; and recognition of 100 words) as within-subject factor. Bonferroni adjustments were applied to post-hoc multiple comparisons. Since 100 words recognition test provided 12 choices in the background, the test accuracy was compared against 8.33% chance level using one-sample t-test. In addition, Pearson chi-square analysis was performed to compare if the number of participants that achieved > 90% accuracy in the 100 words recognition test was different between the two learning paradigms.

We subjected the identification accuracy of the AXB identification task to another mixed ANOVA with learning paradigm (guided learning vs. self-guided learning) as between-subject factor and time of assessment (at the beginning of the study on Day 1 vs. at the end of the study on Day 3) as within-subject factor to examine how familiarity and experience contribute to vibrotactile pattern identification and check whether vibrotactile pattern identification accuracy was a confounding factor between the guided learning and self-guided learning paradigms. Subsequently, we were interested in the degree to which identifiability of individual vibrotactile patterns contributed to 100 words recognition accuracy. We first estimated word recognition accuracy using a stochastic model in which the correct identification of each constituent phoneme was assumed to be independent with probability equal to the accuracy seen in the identification study. We performed hierarchical multiple regression by entering the estimated accuracy of 100 words recognition based on identification accuracy on Day 3 first, followed by estimated accuracy based on identification accuracy on Day 1. Paired samples t-test was also performed to compare the estimated and actual 100 words recognition accuracy.

The effectiveness of place-voice-manner of articulation as a mnemonic device was tested by entering test results during guided learning on Day 1 into a mixed ANOVA with availability of articulation map (available vs. not available) as between-subject factor and time of assessment (List 1 recall, List 2 generalization, List 2 recall and the second recall of List 1) as within-subject factor. Pearson correlation coefficients were calculated to assess the relationship between perceived helpfulness of articulation and test results.

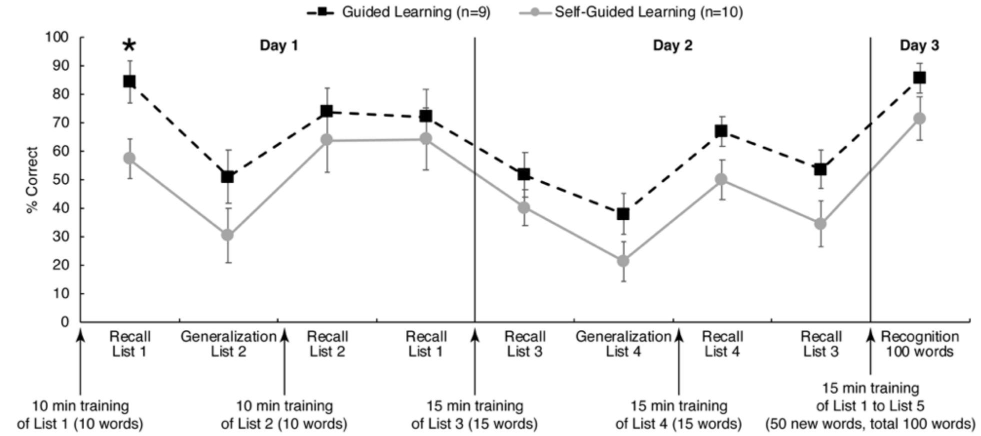


Figure 5: *Test accuracy at different phases from Day 1 to Day 3. Error bars represent standard error of the mean.*

3 RESULTS AND DISCUSSION

3.1 Guided Learning vs. Self-Guided Learning Paradigm

The guided learning group significantly outperformed the self-guided learning group at the beginning of the study in recalling List 1 [$t(17) = 2.57$, $p = .02$; Fig. 5], but no significant differences were observed in the rest of the phases. Both groups performed significantly above 8.33% chance level ($t(8) = 13.76$ and $t(9) = 8.30$ for guided and self-guided learning paradigms, respectively, $p < .01$). After 65 minutes of learning spread across 3 days, although the average performance was similar between the two learning paradigms in 100 words recognition test (average accuracy 85.71 ± 5.62 (SEM) and 71.5 ± 7.62 for guided and self-guided learning paradigms, respectively), significantly more people achieved 90%

or better accuracy in guided learning (66.7%) than self-guided learning (20%) paradigms ($x^2(1) = 4.23$, $p = .04$), implying that the guided learning paradigm was better at forming learning blocks, and was beneficial to the majority of the participants. This finding is in line with a previous study; where guided learning to a new technology is more helpful for a novice because it is less overwhelming and distracting [10]. Furthermore, this is the first study to introduce guided learning for haptic words. The 9 naïve participants in the guided learning group demonstrated the fastest learning rate of haptic words (1.54 words/min), which is 4 − 124 times faster than previous studies [5, 14, 15, 34], independent of coding algorithms, without compromising performance.

3.2 Changes in Vibrotactile Pattern Identification Accuracy

Both groups significantly improved identification accuracy of vibrotactile patterns after the 3-day training ($F(1,17) = 45.66$, $p < .01$; Fig. 6A), suggesting that familiarity and experience contributed to the identification of vibrotactile patterns. Moreover, identification accuracy of vibrotactile patterns on Day 3 was a strong predictor of 100 words recognition performance ($F(1,18) = 38.97$, $p < .001$; Fig. 6B). It explained 70% of the variance in 100 words recognition performance. Identification accuracy of vibrotactile patterns on Day 1 did not further increase the predictability (R^2 Change = .02, $p = .29$). Interestingly, the identification task was rated easier than the word recognition task. Presumably because participants were only required to identify the second vibrotactile pattern in the identification task, while in the word recognition task, they had to identify all the patterns, associate each pattern to its corresponding phoneme, and piece together the word based on the phonemes. From a bottom-up perspective, using identification accuracy of vibrotactile patterns alone should overestimate 100 words recognition performance because it simplifies the word recognition process, but it actually significantly underestimated the performance ($t(18) = 5.32$, $p < .001$), suggesting top-down knowledge plays a critical role in the word recognition and learning process. Similar results were also found by Zhao and her colleagues. They found a shorter training time was required for phonemes within the context of words compared to individual phonemes alone [34]. Therefore, the learning of TLC should be implemented in the context of words to make the learning faster and closer to the real-world setting where context facilitates comprehension.

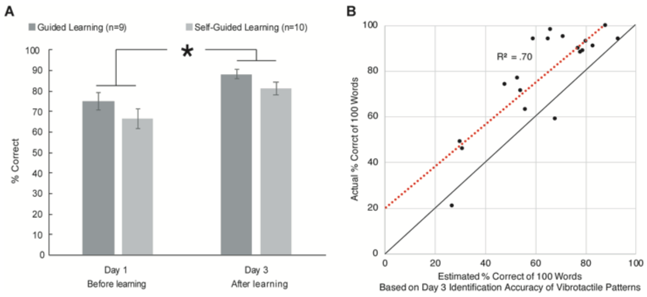


Figure 6: *(A) Identification accuracy of vibrotactile patterns in both groups on Day 1 and Day 3. Error bars are standard errors. (B) Identification accuracy of vibrotactile patterns after 65 minutes of training spread across 3 days predicted 70% variation of 100 words recognition accuracy. Red dotted line represents the best-fit line. Diagonal black line represents the exact predictor.*

3.3 Effectiveness of Articulation Map

Whether or not the articulation map was introduced did not affect learnability or generalizability ($F(1,16) = 002$, $p = .90$; Fig. 7), suggesting that the articulation map did not facilitate or interfere with memorizing the association of vibrotactile patterns and phonemes in the context of this study. Although 5 out of 9 participants found the articulation map to be helpful for learning, the other 4 participants disagreed. In addition, the perceived helpfulness of the articulation map was not correlated to word recall or generalization accuracy (r = .10, .06, .06, and .04; p = .80, 0.88, .88, and .92; for List 1 recall, List 2 generalization, List 2 recall, and the second recall of List 1; respectively). Nevertheless, previous studies have suggested that some mnemonic devices require extensive practice before being used effectively [2, 20], especially when the mnemonic devices are unfamiliar to the participants [3]. Indeed, when speaking, articulation flows automatically without much attention to place-voice-manner of articulation of individual phonemes unless someone wants to improve his or her speech production [13, 26]. Therefore, making the articulation map an effective mnemonic device may require more training to first get familiar with the articulation map itself.

An alternative explanation for the lack of significance of the articulation map is implicit learning meaning that participants who were not introduced to the articulation map, still acquired the rule without conscious awareness. Like the artificial grammar learning paradigm, where unbeknown to the learners, they were asked to observe or remember a series of symbol strings that were made up of complicated arbitrary rules. Later, they were able to distinguish strings that were "grammatically" correct or not [19, 30]. Previous work also showed participants performed better if the phonemes were mapped to the articulation map rather than random assignment [34]. Therefore, it is important to map the phonemes to vibrotactile patterns based on a certain rule, but it may take time for people to learn the rule and use it.

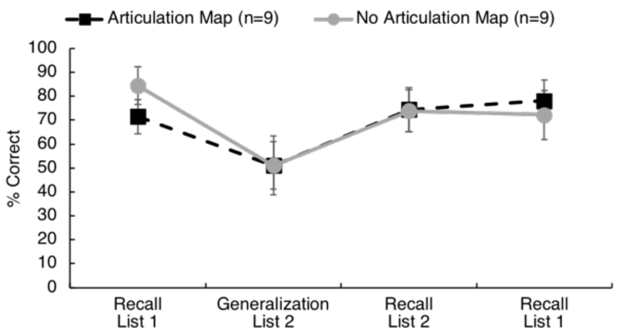


Figure 7: *Participants performance with or without knowing the vibrotactile patterns were coded based on articulation. Error bars are standard error of the mean.*

4 CONCLUSIONS

In the guided learning paradigm, participants were guided to learn the association of vibrotactile patterns and phonemes through incremental rehearsal, flashcards, explore mode, and video game by following a curriculum. In contrast, participants in self-guided learning had complete freedom to allot time between explore mode and flashcards. We found that the guided learning paradigm was better than the self-guided learning paradigm to form learning blocks at the beginning when participants were first introduced to the TLC system. Furthermore, it helped more people achieve > 90% accuracy in recognizing 100 words, showing by far the fastest learning rate and comparable performance to previous studies. In addition, familiarity and experience greatly improved identification of individual vibrotactile patterns, which in turn, facilitated 100 words recognition. Although phonemes were mapped to vibrotactile patterns based on place-voice-manner of articulation, introduction of the articulation map did not facilitate or interfere with the learning of haptic words. To make the articulation map a more effective mnemonic device, we may need to include a more comprehensive training of the articulation map in our future learning paradigm. Next steps also involve decreasing the size of the device, validating the design of the entire set of vibrotactile patterns for all the phonemes, and examining words per minute in the context of sentence comprehension.

REFERENCES

[1] Alcorn, S. 1932. The Tadoma Method. *The Volta Review.* 34, (1932), 195–198.

[2] Bellezza, F.S. 2012. *Imagery and Related Mnemonic Processes.* Springer New York.

[3] Bellezza, F.S. and Reddy, B.G. 1978. Mnemonic devices and natural memory. *Bulletin of the Psychonomic Society.* 11, 5 (1978), 277–280. DOI:https://doi.org/10.3758/BF03336829.

[4] Bochkarev, V. V. et al. 2015. The average word length dynamics as an indicator of cultural changes in society. *Social Evolution and History.* 14, 2 (2015), 153–175.

[5] Brooks, P.L. et al. 1985. Acquisition of a 250-word vocabulary through a tactile vocoder. *The Journal of the Acoustical Society of America.* 77, 4 (Apr. 1985), 1576–1579. DOI:https://doi.org/10.1121/1.392000.

[6] Brooks, P.L. et al. 1986. Continuing evaluation of the Queen's University tactile vocoder II: Identification of open set sentences and tracking narrative. *Journal of rehabilitation research and development.* 23, 1 (1986), 129–138.

[7] Brooks, P.L. and Frost, B.J. 1983. Evaluation of a Tactile Vocoder for Word Recognition. *Journal of the Acoustical Society of America.* 74, 1 (1983), 34–39. DOI:https://doi.org/10.1121/1.389685.

[8] Chang, M.-M. 2005. Applying Self-Regulated Learning Strategies in a Web-Based Instruction — An Investigation of Motivation Perception. *Computer Assisted Language Learning.* 18, 3 (2005), 217–230. DOI:https://doi.org/10.1080/09588220500178939.

[9] Cholewiak, R.W. and Collins, A.A. 2003. Vibrotactile localization on the arm: Effects of place, space, and age. *Perception & Psychophysics.* 65, 7 (Oct. 2003), 1058–1077. DOI:https://doi.org/10.3758/BF03194834.

[10] Debowski, S. et al. 2001. Impact of Guided Exploration and Enactive Exploration on Self-Regulatory Mechanisms and Information Acquisition Through Electronic Search. *Journal of Applied Psychology.* 86, 6 (2001), 1129–1141. DOI:https://doi.org/10.1037//0021-9010.86.6.1129.

[11] Dondlinger, M.J. 2007. Educational Video Game Design : A Review of the Literature. *Journal of Applied Educational Technology.* 4, 1 (2007).

[12] Eagleman, D. 2014. Plenary talks: A vibrotactile sensory substitution device for the deaf and profoundly hearing impaired. *2014 IEEE Haptics Symposium (HAPTICS)* (Feb. 2014), xvii–xvii.

[13] Ecroyd, D.H. 1966. Making the separate sounds of speech. *Voice and Articulation: A Handbook.* Scott Foresman & Co. 63–87.

[14] Enriquez, M. et al. 2006. Haptic phonemes: basic building blocks of haptic communication. *Proceedings of the 8th international conference on Multimodal interfaces.* (2006), 302–309. DOI:https://doi.org/10.1145/1180995.1181053.

[15] Gault, R.H. 1924. Progress in experiments on tactual interpretation of oral speech. *Journal of Abnormal Psychology and Social Psychology.* 19, 2 (1924), 155–159. DOI:https://doi.org/10.1037/h0065752.

[16] Geldard, F.A. 1957. Adventures in tactile literacy. *American Psychologist.* 12, 3 (1957), 115–124. DOI:https://doi.org/10.1037/h0040416.

[17] Geldard, F.A. 1960. Some Neglected Possibilities of Communication. *Science.* 131, 3413 (1960).

[18] Gonzalez-pardo, A. 2013. Game-like language learning in 3-D virtual environments. *Computers & Education.* 60, 1 (2013), 210–220. DOI:https://doi.org/10.1016/j.compedu.2012.07.001.

[19] Hulstijn, J.H. 2005. Theoretical and Empirical Issues in the Study of Implicit and Explicit Second-Language Learning: Introduction. *Studies in Second Language Acquisition.* 27, 02 (Jun. 2005), 129–140. DOI:https://doi.org/10.1017/S0272263105050084.

[20] Kliegl, R. et al. 1990. On the Locus and Process of Magnification of Age Differences During Mnemonic Training. *Developmental Psychology.* 26, 6 (1990), 894–904. DOI:https://doi.org/10.1037/0012-1649.26.6.894.

[21] Kornell, N. and Bjork, R.A. 2008. Optimising self-regulated study: The benefits and costs of dropping flashcards. *Memory.* 16, 2 (2008), 125–136. DOI:https://doi.org/10.1080/09658210701763899.

[22] Kornell, N. and Bjork, R.A. 2007. The promise and perils of self-regulated study. *Psychonomic Bulletin & Review.* 14, 2 (2007), 219–224.

[23] Lamel, L.F. et al. 1989. Speech Database Development: Design and Analysis of the Acoustic-Phonetic Corpus. *Speech Input/Output Assessment and Speech Databases.* 2 (1989), 2161–2170.

[24] Liao, Y.-C. et al. 2016. EdgeVib: effective alphanumeric character output using a wrist-worn tactile display. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16.* (2016), 595–601. DOI:https://doi.org/10.1145/2984511.2984522.

[25] Lim, S. joo and Holt, L.L. 2011. Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization. *Cognitive Science.* 35, 7 (2011), 1390–1405. DOI:https://doi.org/10.1111/j.1551-6709.2011.01192.x.

[26] Lindblom, B. 1996. Role of articulation in speech

perception: Clues from production. *The Journal of the Acoustical Society of America.* 99, 3 (1996), 1683–1692. DOI:https://doi.org/10.1121/1.414691.

[27] Liu, R. and Holt, L.L. 2011. Neural Changes Associated with Nonspeech Auditory Category Learning Parallel Those of Speech Category Acquisition. *Journal of Cognitive Neuroscience.* 23, 3 (2011), 683–698. DOI:https://doi.org/10.1162/jocn.2009.21392.Neural.

[28] MacQuarrie, L.L. et al. 2002. Comparison of retention rates using traditional, drill sandwich , and incremental rehearsal flash card methods. *School Psychology Review.* 31, 4 (2002), 584–595.

[29] Novich, S.D. and Eagleman, D.M. 2015. Using space and time to encode vibrotactile information: toward an estimate of the skin's achievable throughput. *Experimental Brain Research.* 233, 10 (2015), 2777–2788. DOI:https://doi.org/10.1007/s00221-015-4346-1.

[30] Reber, A.S. and Allen, R. 2000. Individual differences in implicit learning: Implications for the evolution of consciousness. *Individual differences in conscious experience: Advances in consciousness research.* R.G. Kunzendorf and B. Wallace, eds. John Benjamins Publishing Company. 227–247.

[31] Roberts, M.L. et al. 1991. Differential effects of fixed instructional ratios on students' progress in reading. *Journal of Psychoeducational Assessment.* 9, 4 (1991), 308–318.

[32] Turcott, R. et al. 2018. Efficient Evaluation of Coding Strategies for Transcutaneous Language Communication. *Proceedings of Eurohaptics* (2018).

[33] Wade, T. and Holt, L.L. 2005. Incidental Categorization of Spectrally Complex Non-Invariant Auditory Stimuli in a Computer Game Task. *Journal of the Acoustical Society of America.* 118, 4 (2005), 2618–2633. DOI:https://doi.org/10.1121/1.2011156.

[34] Zhao, S. et al. 2018. Coding Tactile Symbols for Phonemic Communication. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '18)* (2018).

[35] Zimmerman, B.J. 1990. Self-Regulated Learning and Academic Achievement : An Overview Self-Regulated Learning and Academic Achievement : An Overview. *Educational Psychologist.* 25, 1 (1990), 3–17. DOI:https://doi.org/10.1207/s15326985ep2501.