

# Speech Communication through the Skin: Design of Learning Protocols and Initial Findings

Jaehong Jung<sup>1</sup>, Yang Jiao<sup>1</sup>, Frederico M. Severgnini<sup>1</sup>, Hong Z. Tan<sup>1</sup>,  
Charlotte M. Reed<sup>2</sup>, Ali Israr<sup>3</sup>, Frances Lau<sup>3</sup>, and Freddy Abnoui<sup>3</sup>

<sup>1</sup> Purdue University, West Lafayette, IN 47907, USA

{jung137;jiao12;fmarcoli;hongtan}@purdue.edu

<sup>2</sup> Massachusetts Institute of Technology, Cambridge, MA 02139, USA  
cmreed@mit.edu

<sup>3</sup> Facebook Inc., Menlo Park, CA 94025, USA  
{aliisrar;flau;abnoui}@fb.com

**Abstract.** Evidence for successful communication through the sense of touch is provided by the natural methods of tactual communication that have been used for many years by persons with profound auditory and visual impairments. However, there are still currently no wearable devices that can support speech communication effectively without extensive learning. The present study reports the design and testing of learning protocols with a system that translates English phonemes to haptic stimulation patterns (haptic symbols). In one pilot study and two experiments, six participants went through the learning and testing of phonemes and words that involved different vocabulary sizes and learning protocols. We found that with a distinctive set of haptic symbols, it was possible for the participants to learn phonemes and words in small chunks of time. Further, our results provided evidence of the memory consolidation theory in that recognition performance improved after a period of inactivity on the part of a participant. Our findings pave the way for future work on improving the haptic symbols and on protocols that support the learning of a tactile speech communication system in hours as opposed to the much longer periods of time that are required to learn a new language.

**Keywords:** haptic speech communication, memory consolidation, phoneme.

## 1 Introduction

For a very long time, individuals with hearing impairments have utilized various visual and touch-based methods for speech communication such as *sign language* and *finger-spelling*. Individuals who are both deaf and blind, however, use a natural (i.e., not assisted by man-made devices) method called *Tadoma* that relies solely on the sense of touch. In this method, the “listener” places the hand on the speaker’s face, and in the absence of any visual or auditory information, feels the articulatory processes (e.g., mouth opening, air flow, muscle tension and presence/absence of laryngeal vibration) associated with speech production (**Fig. 1**). The Tadoma method is a living proof that

speech communication through the skin alone is entirely achievable, as established by past research on Tadoma and other natural speech communication methods [1]. Inspired by the Tadoma method, electromechanical devices have been developed to study and replicate the cues produced by a talking face that are crucial for speech reception by Tadoma users [2]. Experimental results on one such device, the Tactutor, demonstrate an information transmission rate of 12 bits/s, the same rate that has been established for speech communication by Tadoma users [1, 3]. Despite such promising results in the laboratory, however, there are still no commercially-available communication devices that can be used by people who are deaf or deaf-and-blind for speech reception at a level that is comparable to that shown by Tadoma users. The challenges are manifold. First, a talking face is a rich display and require actuators that can deliver a wide range of sensations including movement, vibration, airflow, etc. Second, such communication devices should be wearable, or at least portable. Third, the input to such devices should ideally be processed, as opposed to raw, acoustic signals for consistent conversion of sound to touch. Fourth, the mapping between speech sounds and haptic symbols needs to be designed such that it is easy to learn and retain. Last but not the least, a significant effort is expected of any individual who wishes to learn to use such a communication device.



**Fig. 1.** The Tadoma method of speech communication. Shown are two individuals who are both deaf and blind (on the left and right, respectively) conversing with a researcher (center) who has normal vision and hearing. (Photo courtesy of Hansi Durlach)

With the recent development in haptic display and speech recognition technologies, now is the time to give it another try to develop a wearable system for speech communication through the skin. In our research, we use an array of wide-bandwidth actuators for presenting “rich” haptic signals to the skin on the forearm. The display portion of our system is wearable but still tethered to equipment that has yet to be miniaturized. We assume that speech recognition technologies are available to extract phonemes from oral speech in real time, and therefore use phonemes as the input to our system. We have designed and tested a set of distinct haptic symbols representing phonemes of the

English language. The present work focuses on the exploration of training protocols that facilitate the learning of phonemes and words in hours as opposed to days or weeks. Our initial results with six participants have been very promising. In the rest of this paper, we present the background for our approach followed by methods and results from one pilot study and two experiments. We conclude the paper with guidelines for effective learning of speech communication through the skin via man-made systems.

## 2 Related Work

There is a long history of research on the development of synthetic tactile devices as speech-communication aids for persons with profound hearing loss (e.g., see reviews [4-6]) that continues to the present day [7-9]. From a signal-processing point of view, many devices have attempted to display spectral properties of speech to the skin. These displays rely on the principle of frequency-to-place transformation, where location of stimulation corresponds to a given frequency region of the signal. Another approach to signal processing has been the extraction of speech features (such as voice fundamental frequency and vowel formants) from the acoustic signal prior to encoding on the skin. For both classes of aids, devices have included variations in properties such as number of channels, geometry of the display, body site, transducer properties, and type of stimulation (e.g., vibrotactile versus electrotactile).

A major challenge in the development of tactile aids lies in encoding the processed speech signals to match the perceptual properties of the skin. Compared to the sense of hearing, the tactual sensory system has a reduced frequency bandwidth (20-20,000 Hz for hearing compared to 0-1000 Hz for touch), reduced dynamic range (115 dB for hearing compared to 55 dB for touch), and reduced sensitivity for temporal, intensive, and frequency discrimination (see [10]). The tactual sense also lags behind the auditory sense in terms of its capacity for information transfer (IT) and IT rates [11]. For example, communication rates of up to 50 words/min are achieved by experienced operators of Morse code through the usual auditory route of transmission, compared to 25 words/min for vibrotactile reception of these patterns [12]. Taking these properties of the tactual sense into account, certain principles may be applied to create displays with high IT rate. One such principle is to include as many dimensions as possible in the display, while limiting the number of variables along each dimension.

Another challenge is present in the need to provide users with adequate training in the introduction of novel tactile displays. Compared to the structured training and many years of learning associated with the Tadoma method, most tactile aids have been evaluated within the context of relatively limited exposure in a laboratory setting. Recent advances arising out of the literature in language-learning [13] and memory consolidation [14] offer insight into improved approaches for training that may be applied to the use of a novel tactile speech display. Results from a multi-modal, game-based approach to language learning have shown that observers are able to learn to categorize auditory speech sounds when they are associated with the visual stimuli needed to perform the task. In addition, studies of learning suggest that following exposure to a new task, the initial memories associated with this task may be further consolidated by activations of

these memories during periods of wakefulness and sleep. Thus, learning can occur between the laboratory sessions with no explicit training involved.

Based on the literature, we explored several strategies of facilitating learning and training such as breaking up training time into smaller intervals, sounding a phoneme/word out while feeling the corresponding haptic symbols, and keeping the task doable but challenging at all times.

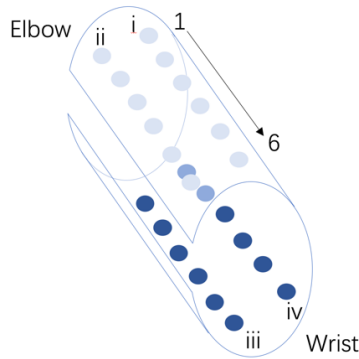
### 3 General Methods

#### 3.1 Participants

A total of six participants (2 females; age range 20-30 years old) took part in the present study. All were right handed with no known sensory or motor impairments. The participants came from diverse language backgrounds. While all participants spoke English fluently, only one was a native English speaker. Other languages spoken among the participants included Korean, Chinese, Tibetan, Hindi and German. Most of the participants also received early childhood music training including piano, clarinet, violin and percussion.

#### 3.2 Apparatus

The experimental apparatus consisted of a 4-by-6 tactor array worn on the non-dominant forearm. The 24 tactors form four rows in the longitudinal direction (elbow to wrist) and six columns (rings) in the transversal direction (around the forearm). As shown in **Fig. 2** below, two rows (i and ii) reside on the dorsal side of the forearm and the other two (iii and iv) on the volar side. The tactor positions were adjusted so that the rows form straight lines and the columns are evenly distributed from the elbow to the wrist. The tactors were attached to a sleeve via adjustable Velcro strips. The sleeve was then wrapped around the forearm with a snug fit to ensure good contact between the tactors and the skin.

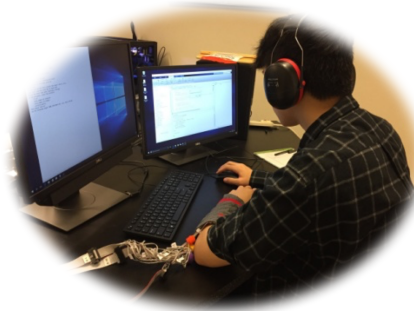


**Fig. 2.** Illustrations of tactor layout in the experimental apparatus

A wide-bandwidth tactor (Tectonic Elements, Model TEAX13C02-8/RH, Part #297-214, sourced from Parts Express International, Inc.) was used as the actuator. It has a flat frequency response in the range 50 Hz to 2 kHz with a resonant peak close to 600 Hz. A MOTU audio device (MOTU, model 24Ao, Cambridge, MA, USA) was used for delivering 24 channels of audio waveforms to the 24 tactors through custom-built stereo audio amplifiers. A Matlab program running on a desktop computer generated the multi-channel waveforms corresponding to the haptic symbols for phonemes, presented a graphic user interface for running the experiments, and collected responses from the participants.

With this setup, the tactors can be driven independently with programmable waveforms and on-off timing. The stimulus properties include amplitude (specified in dB sensation level, or dB above individually-measured detection thresholds), frequency (single or multiple sinusoidal components), waveform (sinusoids with or without modulation), duration, location, numerosity (single tactor activation or multiple tactors turned on simultaneously or sequentially), and movement (smooth apparent motion or discrete salutatory motion varying in direction, spatial extent, and trajectory).

The participants sat comfortably in front of a computer monitor. They wore a noise-reduction earphone to block any auditory cues emanating from the tactors. The participants placed their non-dominant forearm on the table with the volar side facing down. The elbow-to-wrist direction was adjusted to be parallel to the participant's torso. The participants used their dominant hand to operate the computer keyboard and mouse (Fig. 3).



**Fig. 3.** Experimental setup

### 3.3 Phoneme codes

English words are pronounced by a sequence of sounds called phonemes [15]. **Table 1** shows the IPA (international phonetic alphabet) symbols of the 39 English phonemes used in the present study and example words that contain the corresponding phonemes. The list consists of 24 consonants and 15 vowels.

Vibrotactile patterns using one or more of the 4-by-6 tactors were created, one for each phoneme. The mapping of the phonemes to haptic symbols incorporated the articulatory features of the sounds, balanced by the need to maintain the distinctiveness of the 39 haptic symbols. For example, place of articulation was mapped to the longitudi-

**Table 1.** The thirty-nine (39) English phonemes used in the present study.

IPA symbol	example word	IPA symbol	example word	IPA symbol	example word	IPA symbol	example word
/i/	meet	/U/	hood	/k/	key	/ʒ/	azure
/ei/	mate	/ʌ/	hut	/p/	pay	/tʃ/	chew
/u/	mood	/OU/	boat	/t/	tea	/dʒ/	jeep
/aI/	might	/OI/	boy	/b/	bee	/h/	he
/ae/	mat	/aU/	pouch	/g/	guy	/r/	ray
/ɑ/	father	/d/	do	/f/	fee	/l/	lie
/O/	bought	/m/	me	/ʃ/	she	/j/	you
/E/	met	/s/	see	/θ/	think	/n/	new
/ɜ/ /ə/	bird	/w/	we	/v/	voice	/ŋ/	sing
/I/	bid	/ð/	the	/z/	zoo		

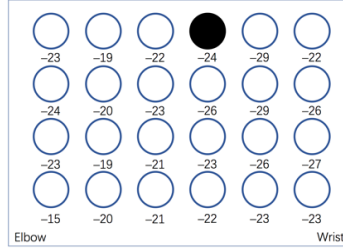
nal direction so that the wrist corresponds to the front of the mouth and the elbow the back of the mouth. Therefore, the consonant /p/ was mapped to a 100-ms 300-Hz pulse delivered near the wrist (front of the mouth) whereas the consonant /k/ was mapped to the same waveform delivered near the elbow (back of the mouth). Their voiced counterparts, /b/ and /g/, were mapped to the 100-ms 300-Hz pulse modulated by a 30-Hz envelope signal delivered near the wrist and elbow, respectively. The modulation resulted in a “rough” sensation that signified voicing. Details of the phoneme mapping strategies and the resultant haptic symbols can be found in [16].

### 3.4 Intensity calibration

In order to control the perceived intensity of vibrotactile signals at different frequencies and different locations on the forearm, signal amplitudes were calibrated in two steps in Exp. I. First, individual detection thresholds were taken at 25, 60 and 300 Hz for the tactor on the dorsal side of the forearm near the center (row i, column 4 in **Fig. 2**). A one-up two-down adaptive procedure was used and the resulting detection threshold corresponds to the 70.7 percentile point on the psychometric function [17]. Signal amplitudes were then defined in sensation level (SL); i.e., dB above the detection threshold at the same frequency. In the present study, signal amplitudes were set to 30 dB SL for a clear and moderately-strong intensity.

Second, the perceived intensity of the 24 tactors was equalized using a method of adjustment procedure. A 300-Hz sinusoidal signal at 30 dB SL was sent to the tactor used in the detection threshold measurements (see the black tactor in **Fig. 4**). The participant selected one of the remaining 23 tactors, say the upper-left tactor in **Fig. 4**, and adjusted its vibration amplitude until the vibration felt as strong as that of the black tactor. This was repeated for all the tactors. The equalization results for one participant are shown in **Fig. 4**. The numbers below each tactor indicate the signal amplitudes in dB relative to the maximum amplitude allowed in the Matlab program for a 300-Hz vibration at 30 dB SL. For example, this participant’s detection threshold at 300 Hz was −54 dB relative to the maximum allowable amplitude. The amplitude for the black reference tactor was therefore at −24 dB for a 30 dB SL signal. The number −23 below the tactor in the upper-left corner indicates that its amplitude needed to be 1 dB higher than that of the black tactor to match its strength for a signal at 30 dB SL. In other words,

the skin near the elbow under this factor was slightly less sensitive than the skin under the black factor. Generally speaking, the skin on the dorsal side was more sensitive than that on the volar side and the wrist area was more sensitive than the elbow area.



**Fig. 4.** Tactor intensity equalization

### 3.5 Data analysis

The experiments reported in this paper consisted of learning and testing of phonemes through individual haptic symbols and words through sequences of haptic symbols. Test results were organized as stimulus-response confusion matrices where each cell entry is the number of times a haptic symbol was recognized as a phoneme label. **Table 2** below shows an example of a confusion matrix for a 6-phoneme stimulus set. As was typical of most tests, a majority of the trials fall on the main diagonal cells (i.e., correct answers). Therefore, the results could be well captured by the percent-correct scores ( $48/50 = 96\%$ ) and the error trials  $/i/ \rightarrow /ei/$  and  $/i/ \rightarrow /u/$ . Therefore, in the rest of the paper, we report the percent-correct scores and error trials for each test.

**Table 2.** An example confusion matrix for a 6-phoneme recognition test. Each cell represents the number of times a haptic symbol is recognized with a phoneme label

		Phoneme Labels (Responses)					
		/m/	/d/	/s/	/ei/	/u/	/i/
Haptic Symbols (Stimuli)	/m/	9	0	0	0	0	0
	/d/	0	5	0	0	0	0
	/s/	0	0	7	0	0	0
	/ei/	0	0	0	5	0	0
	/u/	0	0	0	0	13	0
	/i/	0	0	0	1	1	9

## 4 Pilot study: Learning of 6 Phonemes and 24 Words

The purpose of the pilot study was to gain initial experience and insight into the learning process. One participant (P1) practiced and learned phonemes and words over a period of 21 days and took detailed notes. Within the 21 days, 4 days fell on a weekend, there was a break of 3 days after the 5<sup>th</sup> learning day, and a break of 2 days after the 11<sup>th</sup> learning day. Thus, there were a total of 12 experimental sessions.

*Learning Materials.* The materials included 6 phonemes and 24 words made up of the phonemes. The six phonemes were /i/, /ei/, /u/, /d/, /m/, and /s/. The words consisted of 10 CV (consonant-vowel) words (e.g., may, see) and 14 CVC (consonant-vowel-consonant) words (e.g., moose, dude).

*Time Constraints.* The learning and testing was self-paced. The participant took a break whenever needed.

*Procedure.* The following list shows the tasks performed by P1 over the 12 learning days. For each task, he practiced first with an individual phoneme (or word), then with a random list of phonemes (or words), followed by a recognition test. The numbers in parentheses indicate the total time spent on each learning day.

- Day 1-3 (20, 10, 15 min): 6 phonemes;
- Day 4-7 (5, 5, 5, 17 min): 10 CV words;
- Day 8-11 (5, 24, 5, 30 min): 14 CVC words;
- Day 12 (24 min): test with all 24 words.

*Results.* Participant P1 achieved a performance level of 200/200 (correct-trials/total-trials) with the 6 phonemes on Day 3, 198/200 with the 10 CV words on Day 7, 200/200 with 7 of the 14 CVC words on Day 9, 200/200 with the remaining 7 CVC words on Day 11, and 198/200 with all 24 words on Day 12.

*Insight Gained.* The results of the pilot study indicate clearly that participant P1 was able to learn the 6 phonemes and 24 words almost perfectly after 165 min. He intuitively progressed from easier to harder tasks, each time learning and testing himself before moving onto more difficult tasks. Since the task was highly demanding, it was challenging for P1 to maintain a high level of concentration after about 20 min. Therefore, it was necessary and more productive to spread the learning and testing over many days instead of spending a long time continuously. Furthermore, P1 found that his performance did not deteriorate after a 3-day gap between Day 5 and Day 6.

Encouraged and informed by the results of the pilot study, two experiments were conducted. Experiment I tested four new naïve participants with 10 phonemes and 51 words. Experiment II tested one more naïve participant with the full set of 39 phonemes and tested the memory consolidation theory explicitly.

## 5 Experiment I: Learning of 10 Phonemes and 51 Words

### 5.1 Methods

Four new naïve participants (P2-P5) took part in Exp. I. Each participant spent a total of 60 min to learn 10 phonemes and 50 words made up of the 10 phonemes.

*Learning Materials.* The ten phonemes included the six used in the pilot study and four more: /w/, /ð/, /k/, /a/. The words consisted of the 24 words used in the pilot study plus 27 additional words (13 CV and 14 CVC words).

*Time Constraints.* The learning time was capped at 10 min on each day, with no break. The design ensured that each participant could maintain full concentration during the time spent learning the phonemes and words, and took advantage of memory consolidation by spreading the one-hour learning period over multiple days.



*Procedure.* The following list shows the tasks performed by each participant over the six learning days. On each day, the participant practiced for 5 min, followed by a test with trial-by-trial correct-answer feedback for another 5 min.

- Day 1 (10 min): 6 phonemes;
- Day 2 (10 min): 24 words made up of the 6 phonemes;
- Day 3 (10 min): 4 new phonemes learned, all 10 phonemes tested;
- Day 4 (10 min): 27 new words;
- Day 5-6 (10, 10 min): all 51 words.

## 5.2 Results

The results in terms of percent-correct scores are shown in **Fig. 5**. Due to the fact that the participants reached near-perfect performance level on Day 3 (10 phonemes) and Day 6 (51 words), we do not report the few error trials. The results are organized by day (and cumulative training time in min). Data for the four participants are shown in different color patterns.

Several observations can be made. First, it was relatively easy for the participants to learn the phonemes. Performance was near perfect on Day 1 (6 phonemes) after 5 min of learning and on Day 3 (10 phonemes) after 5 min of learning 4 new phonemes. Second, the transition from phoneme to words took some getting used to, as seen comparing the results from Day 1 (6 phonemes) and Day 2 (24 words made up of the 6 phonemes). This indicates that additional learning was required in order to process phonemes delivered in a sequence. Third, despite the initial “dip” in performance on Day 2 and Day 4 when the participants transitioned from phonemes to words, word-recognition improved quickly as seen in the rising performance from Day 4 to Day 6. The most significant improvement occurred with P5 who reached 62.5%, 77.5% and 97.5% correct from Day 4 to Day 6, respectively. Finally, regardless of individual differences among the four participants in earlier days, all participants succeeded in identifying the 51 words with very few errors by the end of the 60-min period.

Compared with the pilot study, participants in Exp. I learned more phonemes and words in less time. This is probably due to the strict control of learning time per day in order to maintain a high level of concentration on the part of the participants. In addition, the mapping from phoneme to haptic symbols was improved based on the feedback from P1 in the pilot study. The new haptic symbols were more distinct and easier to learn than those in the pilot study. The continued improvement from Day 4 to Day 6 for all participants, especially P5, led us to speculate that memory consolidation may have played a big role in the learning process. We therefore designed Exp. II to explicitly test the effect, if any, of memory consolidation. We also used Exp. II to test whether all 39 phonemes could be learned and how long the learning process would take.

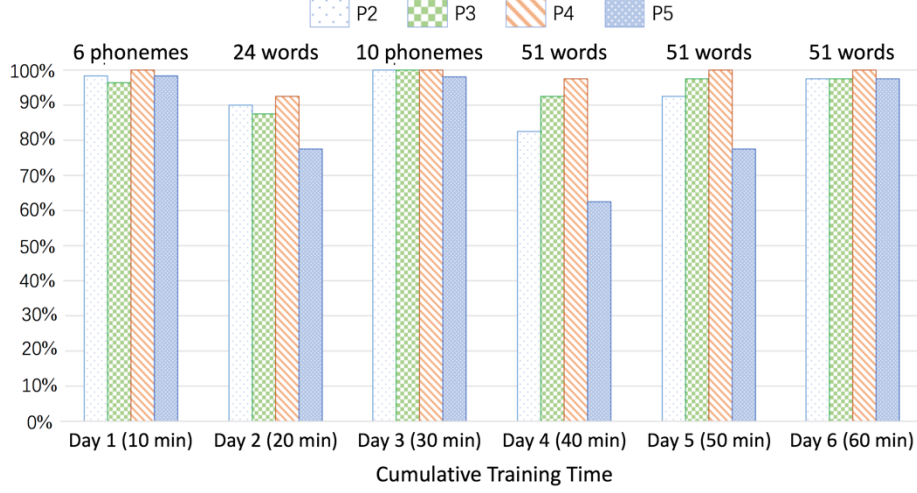


Fig. 5. Results of Exp. I (10 phonemes and 51 words)

## 6 Experiment II: Learning of 39 Phonemes

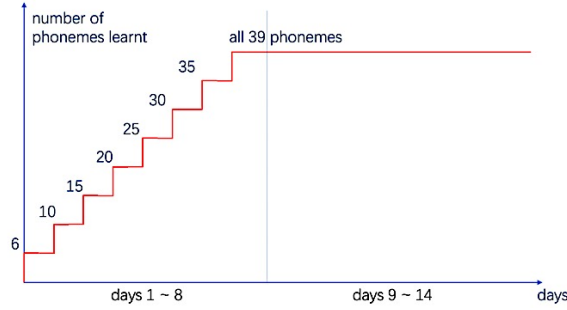
### 6.1 Methods

The objectives of Exp. II were to (1) test memory consolidation explicitly, (2) gain experience and insight into learning all 39 phonemes, and (3) record the time it takes to learn the phonemes and the attainable performance level. One new naïve participant (P6) took part in Exp. II for a total of 14 consecutive days, including the weekends.

*Learning Materials.* All 39 phonemes shown in **Table 1** are included in Exp. II. The 39 phonemes were divided into 8 groups. In addition to the first two groups (containing 6 and 4 phonemes, respectively) that were used in Exp. I, an additional 6 groups of phonemes were created with 5 new phonemes per group except for the last group that contained 4 new phonemes.

*Time Constraints.* As in Exp. I, the learning time was capped at 10 min on each day, with no break. The participant took detailed daily notes on his observations afterwards.

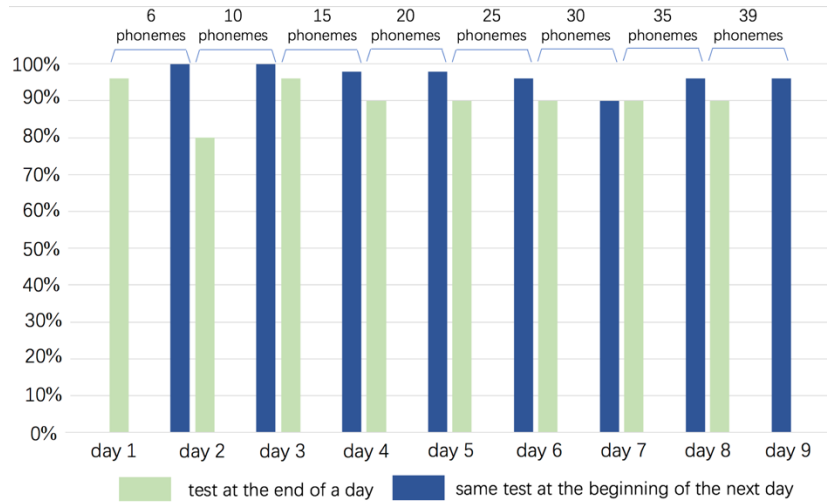
*Procedure.* For testing the memory consolidation theory, the participant always ended a day and began the next day with the *same* test, except for Day 1 when there was no test at the beginning of the day. The participant then spent 3-4 min on learning new phonemes and the rest of the 10 min on testing all phonemes learned so far, with trial-by-trial correct-answer feedback. In addition, the participant sounded out a phoneme during the learning phase. The learning plan is shown below, with the total number of phonemes learned/tested each day clearly marked (**Fig. 6**). The participant had to achieve a percent-correct score of 90% or higher before he could move on to the next group of new phonemes. As shown below, the participant was able to learn one group of phonemes per day during the first 8 days, and was tested with all 39 phonemes from Day 9 to 14.



**Fig. 6.** Learning plan for Experiment II

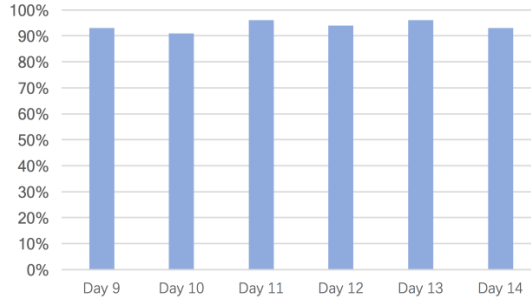
## 6.2 Results

The results are presented in two parts. We first show the percent-correct scores for phoneme recognition from Day 1 to Day 8, including the test conducted on Day 9 that repeated the last test on Day 8 (**Fig. 7**). It is clear that when the same test was conducted again the next day, the performance level either remained the same or improved. This provides a clear evidence for the memory consolidation theory in the sense that performance improved (with the exception of Day 6 to Day 7) after a period of activities not related to phoneme learning. A second observation is that the participant had no difficulty learning between 4 to 6 new phonemes a day, presumably due to the highly distinctive haptic symbols and the easy-to-learn mapping from phoneme to haptic symbols.



**Fig. 7.** Phoneme recognition performance from Day 1 to Day 9

Starting on Day 9, the participant was tested with 4 runs of 50 trials on all 39 phonemes for six consecutive days. The daily percent-correct scores from the 200 pooled trials are shown in **Fig. 8**. Overall, P6 was able to maintain a high performance level ( $93.8\% \pm 3.8\%$  correct).



**Fig. 8.** Daily phoneme recognition scores from Day 9 to Day 14

A stimulus-response confusion matrix was constructed from all 1200 trials ( $50 \text{ trials/run} \times 4 \text{ runs/day} \times 6 \text{ days}$ ) to examine the most-confused phoneme pairs. The following is the list, in order of descending number of confusions:

- /t/ with /k/ (9 times);
- /ae/ with /l/ (8 times);
- /b/ with /d/ (4 times);
- /g/ with /d/ (4 times);
- /i/ with /z/ (4 times);
- /t/ with /p/ (4 times);
- /u/ with /ai/ (3 times);
- /g/ with /m/ (3 times);
- /n/ with /h/ (3 times).

The rest of the errors occurred twice or less and are not listed. The confusion pattern served to guide further refinement of the design of haptic symbols.

## 7 Concluding Remarks

The present study offers evidence to support the claim that speech communication through the sense of touch is an achievable goal. The participants received speech information (phonemes and words) presented on their forearm through a sequence of haptic symbols encoding phonemes. The results of Exp. I show that four naïve participants were able to recognize 51 words in 60 min. The results of Exp. II show that all 39 English phonemes could be learned in 80 min. We demonstrated memory consolidation in Exp. II by showing an improvement in phoneme recognition performance when the participant P6 was tested a day after he learned the phonemes, over multiple days.

Several guidelines can be provided based on the experimental results and the insights we have gained from the present study. First, the learning time should be limited to 10 to 20 min per session. It was difficult for participants to maintain full concentration after 20 min. Second, it might be helpful for the participant to sound a phoneme out as the haptic symbol was delivered to the forearm although we did not collect data to prove that. Third, the task difficulty should be carefully managed so that the participant is challenged but able to make progress. Last but surely not the least, the results of Experiment II provide evidence for learning that occurred between laboratory sessions when the participant was not being trained. Therefore, we recommend spending a short period of time (e.g., 10 min) per day and taking advantage of memory consolidation for further improvement of learning outcome. Furthermore, the benefit of memory consolidation did not appear to be impacted when learning sessions were interrupted by one or more days.

In the future, we will continue to improve the haptic symbols to further reduce errors in phoneme recognition. We will also compare different learning strategies such as reversing the order in which phonemes and words are introduced to the participants. Our ultimate goal is to develop a haptic speech communication system for people with all levels of sensory capabilities.

### Acknowledgments

This work was partially supported by a research grant from Facebook Inc. The authors thank all participants for their time and dedication, and Keith Klumb for project management.

### References

- [1] C. M. Reed, W. M. Rabinowitz, N. I. Durlach, L. D. Braida, S. Conway-Fithian, and M. C. Schultz, "Research on the Tadoma method of speech communication," *The Journal of the Acoustical Society of America*, vol. 77, pp. 247-257, 1985.
- [2] H. Z. Tan, N. I. Durlach, C. M. Reed, and W. M. Rabinowitz, "Information transmission with a multifinger tactual display," *Perception & Psychophysics*, vol. 61, no. 6, pp. 993-1008, 1999.
- [3] C. M. Reed and N. I. Durlach, "Note on information transfer rates in human communication," *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 5, pp. 509-518, 1998.
- [4] C. M. Reed, N. I. Durlach, and L. D. Braida, "Research on Tactile Communication of Speech: A Review," *American Speech-Language-Hearing Association (ASHA) Monographs Number 20*, 23 pp, 1982.
- [5] C. M. Reed, N. I. Durlach, L. A. Delhorne, W. M. Rabinowitz, and K. W. Grant, "Research on Tactual Communication of Speech: Ideas, Issues, and Findings," *Volta Review*, vol. 91, pp. 65-78, 1989.
- [6] J. H. Kirman, "Tactile communication of speech: A review and analysis," *Psychological Bulletin*, vol. 80, pp. 54-74, 1973.

- [7] H. Yuan, C. M. Reed, and N. I. Durlach, "Tactual display of consonant voicing as a supplement to lipreading," *The Journal of the Acoustical Society of America*, vol. 118, no. 2, pp. 1003–1015, 2005.
- [8] A. Israr, C. M. Reed, and H. Z. Tan, "Discrimination of vowels with a multi-finger tactual display," *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 17-24, 2008.
- [9] S. D. Novich and D. M. Eagleman, "Using space and time to encode vibrotactile information: Toward an estimate of the skin's achievable throughput," *Experimental Brain Research*, vol. 233, pp. 2777-2788, 2015.
- [10] R. T. Verrillo and G. A. Gescheider, "Perception via the sense of touch (Chap. 1)," in *Tactile Aids for the Hearing Impaired*, I. R. Summers (Ed.), London: Whurr Publishers, 1992, pp. 1–36.
- [11] C. M. Reed and N. I. Durlach, "Note on information transfer rates in human communication," *Presence: Teleoperators & Virtual Environments*, vol. 7, no. 5, pp. 509–518, 1998.
- [12] H. Z. Tan, N. I. Durlach, W. M. Rabinowitz, C. M. Reed, and J. R. Santos, "Reception of Morse code through motional, vibrotactile, and auditory stimulation," *Perception & Psychophysics*, vol. 59, no. 7, pp. 1004-1017, 1997.
- [13] T. Wade and L. L. Holt, "Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task," *The Journal of the Acoustical Society of America*, vol. 118, no. 4, pp. 2618-2633, 2005.
- [14] Y. Dudai, A. Karni, and J. Born, "The consolidation and transformation of memory," *Neuron*, vol. 88, pp. 20-32, 2015.
- [15] D. H. Ecroyd, M. M. Halfond, and C. C. Towne, *Voice and Articulation: A Handbook*, Glenview, IL: Scott, Foresman and Co., 1966.
- [16] C. M. Reed, H. Z. Tan, X. D. Perez, E. C. Wilson, F. M. Severgnini, J. Jung, J. S. Martinez, Y. Jiao, A. Israr, F. Lau, K. Klumb, and F. Abnoui, "A phonemic-based tactual display for speech communication," in preparation for *IEEE Transactions on Haptics*.
- [17] H. Levitt, "Transformed up-down methods in psychoacoustics," *The Journal of the Acoustical Society of America*, vol. 49, no. 2B, pp. 467-477, 1971.