2019 IEEE World Haptics Conference (WHC)
Tokyo, Japan
9-12 July 2019

Haptic Speech Communication Using Stimuli Evocative of Phoneme Production

Maurício Fontana de Vargas[1] Antoine Weill–Duflos[2] Jeremy R. Cooperstock[2]

***Abstract*— The tactile sense can be used as a channel for general communication, especially in contexts where the visual and auditory modalities are occupied with other tasks or compromised. We propose a new method for communicating generic words through the sense of touch that relies on delivering vibration patterns, representing the phonemes composing the words, to the user's skin through two vibrotactile transducers worn on the forearm. The novelty of this technique is that vibration patterns are created from the audio of the corresponding English phoneme, resulting in vibration patterns that resemble physical characteristics when uttering the phoneme during normal speech. After 100 minutes of training, participants were able to recognize 50 words rendered haptically with an average accuracy of 94.4%. Results support the possibility of using the proposed apparatus in real-world applications.**

I. Introduction and Related Work

The idea that it is possible to transmit language through the sense of touch is not new. Gault [1] describes the idea of "distinguishing the feel of one word from the feel of another and associating meanings with different feels".

Geldard [2] introduced the method of vibratese for communication through touch and showed that it was possible to understand words transmitted by vibration at a rate of $38\,\mathrm{words/min}$. Encouraged by these results, researchers in the area of assistive technologies for the deaf and the deaf-blind [3] have created tactile aids known as vocoders, in which the live acoustic speech signal is processed in real time to provide temporal, intensity, or spectral information as vibration stimuli. Sorgini [4] provides a review of such systems.

A different approach for haptic speech communication—in which this works falls within—assumes a speech recognition module at the front end of the device responsible to convert the speech audio into a text string, which is delivered haptically as a series of discrete stimuli according to a pre-defined mapping. While this mapping can be alphabet-based, the latest research in the field tends to focus instead on multi-actuator communication systems to represent phonemes [5]–[7]. The drawback of these systems, however, is that they impose non-trivial hardware requirements. We take note of the Tadoma language [8], which demonstrated that one could obtain a sufficient understanding of speech (of a known language) through the feeling of lip movements and vibrations of the throat. In a similar manner, Zhao et al. [5] also found that an articulation-based mapping significantly helped participants recognizing words, compared to a random mapping. These findings inspired our own approach, leveraging our innate understanding of the mapping of phonemes to these physical manifestations of speech, which make the haptic representation inherently more easily learnable. In this regard, our objective is not only to achieve rich, efficient communication of language using haptics, but to do so while minimizing the learning curve. As such, it can be used by the general population, with no prior knowledge of Braille, Morse code, or other deaf-blind tactile languages.

[1]Maurício Fontana de Vargas is with the School of Information Studies, McGill University, Montreal, Canada mauricio.fontanadevargas@mail.mcgill.ca

[2]Antoine Weill–Duflos and Jeremy R. Cooperstock are with the Department of Electrical and Computer Engineering, McGill University, Montréal, Canada antoine.weill-duflos@mcgill.ca, jer@cim.mcgill.ca

II. Apparatus

Our apparatus for communicating speech through touch uses only two vibrotactile transducers, worn on the forearm. As such, this simple design minimizes hardware demands, and thus, facilitates mobility.

Unlike prior work based on a discrete mapping, in which the mappings from phonemes to vibration patterns are determined by the designer, we rely primarily on the audio waveform of the corresponding English phoneme to generate the associated patterns. Prior research has similarly exploited the locus of phoneme articulation as a rough guide to determine the spatial location of some of the corresponding stimuli [6] [5]. However, our mapping strategy relies on additional elements related to the vocalization of each phoneme. This offers the important benefit that the vibration patterns resemble physical characteristics——such as the place in the mouth where the phoneme is articulated, the vibration caused by air friction, and the vibration intensity of the vocal chords—when uttering the associated phoneme during normal speech. In turn, this provides a more natural, and easily learnable mapping, allowing users to generalize to an understanding of new words, rendered haptically.

In the remainder of this section, we elaborate on our hardware and the strategy used to represent language haptically.

A. Hardware

The device is composed of two Haptuator Mark-II (Tactile Labs, Montreal) voice-coil transducers, attached to armbands worn near the wrist and elbow, as illustrated in Fig. 1. The transducers are driven by a two-channel audio amplifier (SURE Electronics AA-AB013V120) connected to a microcomputer through its standard audio output. The left audio channel is mapped to the wrist-mounted transducer, while the right channel is mapped to one mounted near the elbow. Vibration patterns are stored as standard stereo audio files.

978-1-5386-9461-9/19/$31.00 © 2019 IEEE

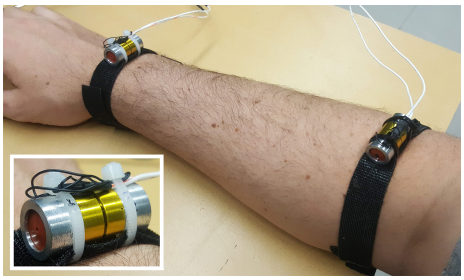

Fig. 1: The two vibrotactile transducers responsible for delivering the vibration patterns to the user's skin.

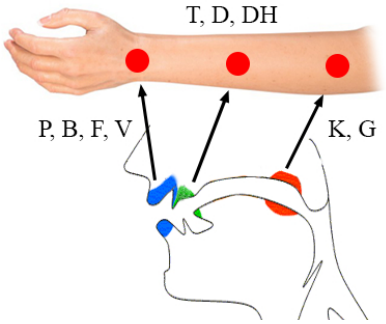

Fig. 2: Mapping between the vocal tract location where a consonant is articulated and the approximate position in the forearm where the corresponding vibration pattern is felt.

B. Covered Phonemes

Our present design supports 24 English phonemes (15 consonants, 5 vowels, and 4 diphthongs), selected based on their frequency of use in casual conversation [9]. Most of the excluded phonemes have a similar phoneme in the supported set that may potentially be used as substitutes with minimal impact on word-level comprehension.

C. Consonants mapping strategy

The haptic stimuli for 15 consonant phonemes (P, B, T, D, K, G, F, V, DH, S, Z, M, N, L, R) were constructed using the following strategy:

1) Audio of the isolated phonemes was obtained from recordings of a native English speaker (`www.jbdowse.com/ipa`).

2) The audio signal was processed to highlight the phoneme's features that are most salient during normal speech, resulting in a stimulus that resembles those features when rendered haptically through the vibrotactile transducers. These features are inherent to how speech organs are involved when producing the sound, and thus, phonemes produced in the same manner are subject to similar signal processing operations. For example, a high-pass filter was selected to emphasize the high-frequency, turbulent sound of the sibilant phonemes (F, V, S, Z), produced by forcing the air through a narrow gap between the lips or teeth. On the other hand, low frequencies were boosted on the nasal phonemes (M, N) to emphasize the characteristic low-frequency resonant sound produced by the air escaping through the nose. Tab. I describes the signal processing applied to all consonants covered.

3) The interaural level difference of the two channels was adjusted to effect a spatial (front/back) panning, indicative of the phoneme's locus of articulation within the vocal tract: phonemes produced towards the front of the mouth (lips, teeth) are biased to the left channel, with the vibration delivered primarily to the wrist; phonemes produced towards the back of the mouth are biased towards the right channel, resulting in vibrations near the elbow, as illustrated in Fig. 2.

4) Finally, a low-pass filter (700 Hz) was applied to remove audible frequencies not detected by skin receptors.

D. Vowels and diphthongs mapping strategy

The construction of the vibration patterns representing 5 vowels (IY, EH, AH, UW, AA) and 4 diphthongs (EY, AY, AW, OW) followed the same principles used for the consonants. The differences are explained below:

1) The audio of the phoneme was obtained from computer synthesized speech using the Praat software. Computer synthesized speech was chosen over human speech recordings due to the flexibility of generating speech sounds for each vowel with specific lengths and constant formant frequency, i.e., characteristic peaks in the frequency spectrum that differentiate vowels from each other. All vowels are 750 ms in length and were synthesized with a fundamental frequency (F0) of 140 Hz. The formant frequencies (F1 and F2) used in the syntheses were provided by the software tools, using average values obtained from real speech [10]. Diphthongs were constructed by varying F1 and F2 linearly between the values of the frequencies of the two associated vowels. Thus, all diphthongs are 1500 ms in length.

2) During normal speech, vowels are produced with the same manner of articulation, without obstructions of airflow, and are all voiced (i.e., vocal chords are used). Thus, they do not present distinct characteristics that can be highlighted with additional signal processing. Rather, they are naturally distinguished by their resonant frequencies in the oral cavity, i.e., their formants, which are also perceived when rendered through the vibrotactile transducers. To increase distinguishability between vowels and consonants, a volume ramping that starts at -70 dB, increases linearly to 0 dB at half-phoneme length, and decreases back to -70 dB at the end of the phoneme, was applied to all vowels, creating a fade-in and fade-out effect unique to the vowels.

3) The interaural level difference (ILD) of the two channels were adjusted to effect a spatial panning indicative of the horizontal position of the tongue when producing the sound: vowels produced toward the front of the mouth (e.g., IY as in meet), are biased to the left channel resulting in vibrations delivered primarily to the wrist, while vowels produced towards the back (e.g., AA as in pot) are felt near the elbow. In the case of diphthongs, the ILD varies progressively between the values of the two vowels, so the perceptual illusion where the vibration is felt moves progressively from the location of the first vowel to that

TABLE I: Overview of the vibration patterns for consonant phonemes. (W = Wrist, C = Center of forearm, E = Elbow)

Phoneme	**Distinct characteristic being reproduced**	**Signal processing applied to highlight characteristics**	**Subjective impression**	**Duration** (ms)	**Location**
P	Strong puff of air	Gain change (13 dB)	Strong, short burst	20	W
B	Weak puff of air	Gain change (−8 dB)	Weak, short burst	20	W
T	Strong puff of air	Gain change (0 dB)	Strong, short burst	23	C
D	Weak puff of air	Gain change (−12 dB)	Weak, short burst	30	C
K	Strong puff of air	Gain change (4 dB)	Strong, short burst	35	E
G	Weak puff of air	Gain change (−7 dB)	Weak, short burst	35	E
F	Airflow btwn lip and teeth	Hi-pass filter (300 Hz, Q=5.0)	Weak blow	530	W
V	Vibration btwn lower lip and upper teeth and in vocal chords	Hi-pass filter (300 Hz, Q=5.0)	Strong buzzing	350	W
DH	Vibration btwn tongue and teeth	–	Weak, short burst followed by weak, reverbing vibrations	80	C
S	Airflow btwn tongue and teeth	HP filter (260 Hz, Q= 1.0)	Weak hissing	480	E
Z	Airflow btwn tongue and teeth with vocal chords vibration	HP filter (200 Hz, Q = 4.3)	Strong buzzing	430	E
M	Resonant vibration in the nasal cavity	Low-frequencies (<110 Hz) boost (8.8 dB, Q = 0.7)	Muffled, low-frequency shaking	550	W
N	Resonant vibration in the nasal cavity	Low-frequencies (<110 Hz) boost (8.8 dB, Q = 0.7)	Muffled, low-frequency shaking	550	E
L	Resonant vibration	Pitch shift (-10 semitones)	Constant, low-frequency vibration	240	W
R	Resonant vibration	Pitch shift (-10 semitones)	Constant, low-frequency vibration	240	E

of the second vowel, thus mimicking the movement of articulation that occurs when speaking diphthongs.

III. User Study

A user study was conducted on the haptic rendering of phonemes and words using the apparatus described. The experiment structure, training activities, and testing protocol were the same as the one used in previous work in the topic [7], with minor adjustments in the phoneme sets made as necessary due to differences in the phonemes covered and the mapping strategy so as to allow for a fair comparison between results. In addition, two novel, more challenging activities were added to the testing protocol, i.e., Sequential Post-Test and Open Answer Final Test, as detailed in Section III-E.

A. Participants

Fourteen participants (8 male, 22-43 years of age, $\mu = 29, \sigma = 6$) were recruited through McGill University's email lists. Only one participant was a native English speaker and five reported having some prior knowledge of phonetics. All participants provided informed consent of the experiment protocol, following Research Ethics Board guidelines, and received compensation of CAD $60.

B. Experimental environment

The experiment was held in a laboratory setting. Participants sat in front of a computer, which ran the experiment software. Instructions were displayed on the computer's monitor and participants interacted with the software through a mouse. Participants wore the arm bands near the wrist and elbow on the right arm, with the vibrotactile transducers facing up, and over-ear headphones with pink noise played at a comfortable level to block out exterior noise and sound produced by the vibrotactile transducers.

C. Training and testing protocol

Participants performed 100 minutes of self-training over 16 activity sessions, spread across four consecutive days, as described in Fig. 3. Activities were designed to teach phonemes progressively according to the sets described in Tab. II, allowing participants to focus on learning a small subset of phonemes at a time, while reviewing and reinforcing the phonemes previously learned.

Testing activities were performed following the training sessions each day, but did not provide feedback, and were thus not considered as training time.

TABLE II: Phoneme sets used during training and testing.

Set	Added Phonemes	Phonemes	Size
C1	Plosives	{P,B,T,D,K,G}	6
C2	Nasals, Approximants	C1 + {M,N,L,R}	10
C3	Fricatives	C2 + {V,F,DH,Z,S}	15
V1	Most distinct vowels	{IY,AH,AA}	3
V2	Intermediate vowels	V1 + {EH,UW}	5
V3	Diphthongs	V2 + {EY,AY,AW,OW}	9

D. Training activities

1) Introduction to phonemes: A brief explanation of the phonemes and phonemic transcription is given, including examples presenting the place of articulation and the phonemic transcription of some common words.

2) Free-play: Participants can click on buttons representing phonemes to feel the corresponding vibration patterns on their arm. The software displays a diagram with the

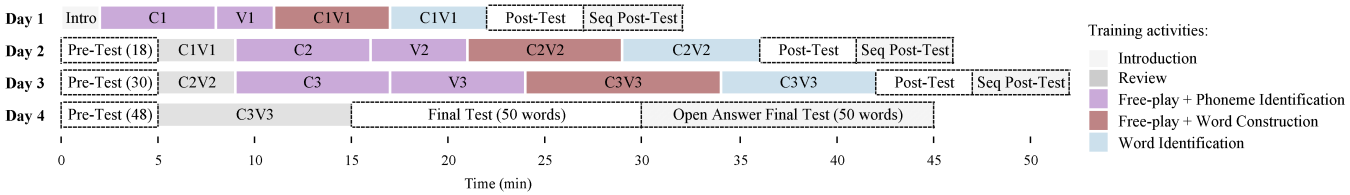

Fig. 3: Training and testing protocol. All activities are time-limited, apart from the Pre-Test, Final Test, and Open Answer Final Test, which have a fixed number of trials, indicated in parentheses.

configuration of speech organs when uttering the phoneme, a visual representation of the vibration pattern, and the approximate location on the arm where it is delivered.

3) Phoneme Identification Quiz: This activity is composed of a self-administered quiz with correct-answer feedback. In each trial, a phoneme is rendered and the participant is requested to identify it among all the alternatives in the set. If the response is wrong, participants can compare the haptic rendering of the correct alternative with their answer.

4) Word Construction Quiz: The goal of this activity is to train participants on how to combine phonemes to form words. The phonemes composing a word are rendered individually and advanced under the control of the participants, and participants are requested to identify them (as in the *Phoneme Identification Quiz*). Once participants have advanced through all phonemes within the word, they are requested to identify the word from a multiple-choice list. Correct-answer feedback is provided.

5) Word Identification Quiz: Same as the *Word Construction Quiz*, but participants identify only the rendered words rather than individual phonemes. Phoneme advance is still controlled by participants.

6) Phoneme Review with Pre-Test: As the first activity on days 2-4, participants are subjected to a pre-test on all the phonemes they have learned so far and are able to review them through the *Free-play* and *Phoneme Identification Quiz* activities. Pre-test accuracy scores are displayed besides each phoneme so they can practice accordingly.

E. Testing activities

1) Post-Test and Final Test: In each trial, a random word is rendered with phoneme advance controlled by the participant. Once all phonemes within the word have been rendered, participants are requested to identify it from a multiple-choice list containing 12 randomly selected words (except on Day 1, due to the word set size), as described in Section III-F.

2) Sequential Post-Test: In each trial, a random word is rendered with an Interstimulus Interval (ISI) i.e., the temporal gap between two consecutive phonemes, of 1 s. Participants are requested to identify the word from a multiple-choice list containing 12 random words.

3) Final Test: The Final Test is similar to the *Post-Test*, but without a time limit to finish the activity. Instead, participants must identify a fixed number (50) of words.

4) Open Answer Final Test: In each trial, a random word is rendered haptically with an ISI of 1 s, and participants are requested to type the word in a text box. Partial responses are also accepted. This test evaluated the ability of participants to recognize a word without any visual cueing of possible choices.

F. Words used

During Training and Post-Test activities on Days 1–3, words are randomly selected from a subset of a 150-word list containing only words composed by phonemes so far learned. These subsets contained 7, 38, and 105 words, respectively. The entire list includes 91 words from Dunkelberger et al. [7], plus an additional 59 words, and ranges from one to six phonemes ($\mu = 3.1, \sigma = 0.98$). In the Final Test, a subset with 50 of these words, including 36 words from Dunkelberger et al. [7] was used, also ranging from one to six phonemes ($\mu = 3.1, \sigma = 1.09$). In the Open Answer Final Test, another 50-word set was used, which includes 25 words ($\mu = 3.1$ phonemes/word) from the 150-word list and 25 completely new words ranging from two to six phonemes ($\mu = 3.1$).

IV. Results

The results for all test activities are shown as a box plot in Fig. 4, and Fig. 5 expands across all participants the Post-Test (PT) and Final Test (FT) accuracy scores, calculated as the number of correctly identified words divided by the number of trials. As can be seen, the average score in the

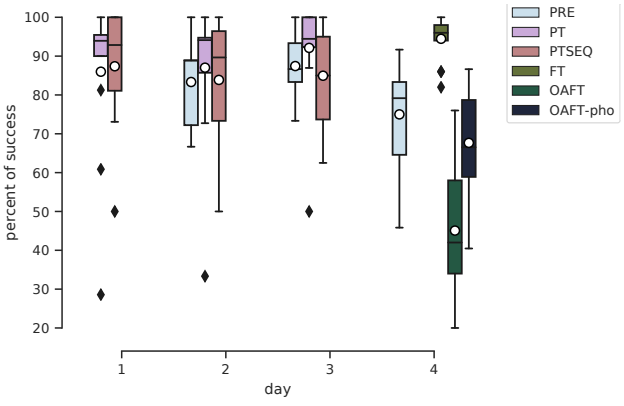

Fig. 4: Testing results. White circles represent averages and black diamonds represent outliers.

Final Test was 94.4%, excluding the results of P10, who was non-compliant with the experiment instructions, and instead randomly answered training and testing quizzes as quickly as possible. These results compare favorably against the state of the art results in haptic communication of language

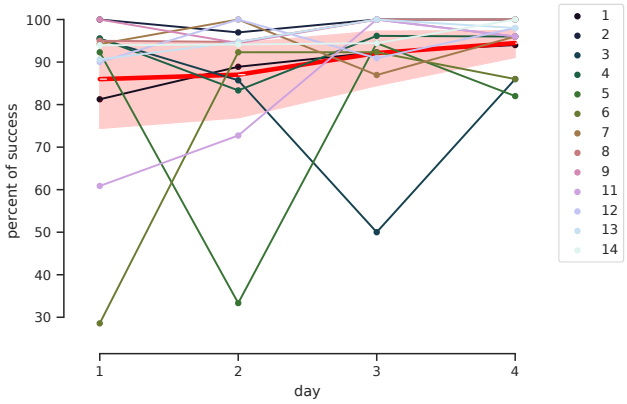

Fig. 5: Partial and Final test results. Average scores and standard deviation across all participants are displayed as the red line and the red-shaded area, respectively.

intended for the general population, for which Dunkelberger et al. obtained 86.6% recognition accuracy [7] with the same training and testing protocol as we used. We note that the results from similar research on this topic cannot be compared directly due to substantial differences in the training protocol and phoneme coverage. For example, Jiao et al. [11] reported 80.6% word recognition accuracy using a word set containing 100 words with two or three phonemes after 100 minutes of training spread over 10 days.

For the more challenging Open Answer Final Test, responses were analyzed based on their phonetic transcription. In other words, answers with the same pronunciation as the rendered word (e.g., "eye" and "I") were considered correct. Average word accuracy score (OAFT) was 45% (51% for words from the training sets and 39% for new words). Overall phoneme identification accuracy (OAFT-pho) averaged 68%.

The distribution of phoneme presentation rate, defined as the average time between the rendering of consecutive phonemes within a word on Post-Tests and Final Test, is presented in Fig. 6. The average value was 3.1 s (σ =2.9 s), slightly faster than observed under similar conditions (3.5 s) in the study by Dunkelberger et al. [7].

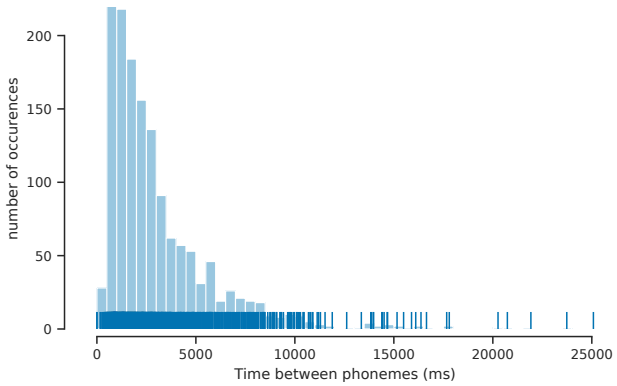

Fig. 6: Distribution of phoneme presentation rate, averaged by word.

To further analyze individual phoneme identification performance, we generated a confusion matrix (Fig. 7) with scores obtained on the retention test (Pre-Test) on Day 4. We also calculated whether the responses belong to the correct phoneme group, finding 99.4% for plosives, 92.3% for nasals and approximants, 70.0% for fricatives, 75.4% for vowels, and 95.2% for diphthongs.

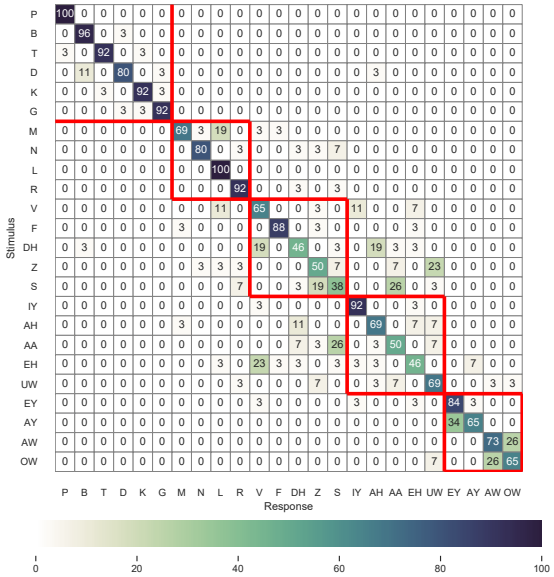

Fig. 7: Phoneme retention scores from Pre-Test, Day 4, plotted as a confusion matrix. Phoneme groups are highlighted with red borders.

V. Discussion

The high scores achieved on the Final Test demonstrate the feasibility of rendering words as a sequence of vibration patterns, constructed from the phonemes composing the word. The phoneme mapping strategy we adopted allowed participants to learn 24 phonemes and to achieve, to the best of our knowledge, the highest scores to date (μ = 94.4%, σ = 5.9%) on word identification accuracy (Final Test), considering similar training and testing conditions.

Although the list of answer options supported participants identifying the correct answer (as discussed below), the high accuracy on phoneme retention obtained on the Pre-Tests strongly suggests that participants made use of this ability to identify words based on their constituent phonemes.

We also compared word identification accuracy between the condition of participant-controlled phoneme advance (Post-Test) vs. that with a fixed ISI (Sequential Post-Test). Results indicate only a slight decrease in accuracy (7% on Day 3, not statistically significant on a Student's t-test) for words rendered with a fixed ISI (of 1 s), and no apparent change on days 1 and 2. These data are therefore inconclusive as to whether the 1 s ISI was uncomfortably short for most participants, despite the small fraction of instances (17%) in which participants requested a faster phoneme advance on the Final Test. We surmise that the lack of response clues contributed more to the increased difficulty of the OAFT, as reflected in the drop in accuracy compared to the FT scores.

The results demonstrate encouraging performance results on word identification, even in the absence of a visual list of candidate words from which to choose. Despite the fact that participants received no training in this condition, nor on word identification with a fixed ISI, they were able to identify a significant number of words correctly.

OAFT performance may have been affected by difficulties in identifying particular phonemes. Due to our mapping approach based on speech production, it was expected that when participants mistake a phoneme, their response would involve a similarly spoken one (i.e., same manner of articulation) to the correct stimulus, which would potentially facilitate word comprehension. Although this was observed for plosives, vowels, nasals, and approximants, some fricatives and vowels were often mistaken for very different phonemes (e.g., Z - UW; S - AA; EH - V). This type of error potentially impacted the word identification accuracy on OAFT, since they are more prone to transform the sequence of phonemes into a meaningless, nonsensical sound.

VI. Conclusions and Future Work

We presented a novel apparatus for haptic communication in which vibration patterns resemble physical characteristics when uttering the associated phoneme during normal speech. After 100 minutes of self-guided training, participants achieved an average accuracy of 94.4% when identifying 50 words haptically, as compared to the 86.6% accuracy attained by Dunkelberger et al. [7] following a similar training process and experimental protocol.

Perhaps even more encouragingly, in the innovative test in which participants were not given any visual cues as to the word options list, they achieved results of 45% correct word recognition and 68% correct phoneme recognition, on average. Although further improvement is necessary, these results point to the possibility of using such a haptic communication apparatus for the rendering of language in real-world applications, e.g., for the deaf, or to facilitate speech understanding in a noisy environment or a crowd.

Our continuing work in this area seeks to achieve communication rates suitable for real-world applications. This involves investigation of the trade-off between recognition and transmission rates when using shorter stimuli for vowels and diphthongs, as well as intervals between phonemes within a word (ISI) tailored for the user based on his/her familiarity with the apparatus. Understanding how extensive training on words—which would potentially allow users to perceive a sequence of phonemes as a single chunk without processing each phoneme individually—impacts word recognition accuracy is also necessary.

Finally, we plan to study the effects on recognition of substituting similar phonemes to convey words whose actual phonemes are not produced by the apparatus, e.g., rendering "book" as (B - UW - K) instead of (B - UH - K). This is motivated by the fact that non-native English speakers are often unable to recognize and reproduce certain phonemes due to first-language phonological interference, yet do not encounter major difficulties in communicating. By understanding the representational accuracy needed for haptic communication, we can further optimize the set of phonemes produced by such systems in order to maximize language expressiveness while minimizing training time and cognitive demands.

References

[1] R. H. Gault, "'Hearing" through the sense organs of touch and vibration," *J. of the Franklin Institute*, vol. 204, no. 3, pp. 329–358, Sep. 1927.

[2] F. A. Geldard, "Adventures in tactile literacy.," *American Psychologist*, vol. 12, no. 3, p. 115, 1957.

[3] K. L. Galvin, G. Mavrias, A. Moore, R. S. Cowan, P. J. Blamey, and G. M. Clark, "A comparison of tactaid ii and tactaid 7 use by adults with a profound hearing impairment," *Ear and hearing*, vol. 20, no. 6, pp. 471–482, 1999.

[4] F. Sorgini, R. Caliò, M. C. Carrozza, and C. M. Oddo, "Haptic-assistive technologies for audition and vision sensory disabilities," *Disability and Rehabilitation: Assistive Tech.*, vol. 13, no. 4, pp. 394–421, May 2018.

[5] S. Zhao, A. Israr, F. Lau, and F. Abnousi, "Coding Tactile Symbols for Phonemic Communication," in *Human Factors in Computer Systems*, New York, NY, USA: ACM, 2018, 392:1–392:13.

[6] C. M. Reed, H. Z. Tan, Z. D. Perez, E. C. Wilson, F. M. Severgnini, J. Jung, J. S. Martinez, Y. Jiao, A. Israr, F. Lau, K. Klumb, R. Turcott, and F. Abnousi, "A Phonemic-Based Tactile Display for Speech Communication," *IEEE Trans. Haptics*, pp. 1–1, 2018.

[7] N. Dunkelberger, J. Sullivan, J. Bradley, N. P. Walling, I. Manickam, G. Dasarathy, A. Israr, F. W. Y. Lau, K. Klumb, B. Knott, F. Abnousi, R. Baraniuk, and M. K. O'Malley, "Conveying Language Through Haptics: A Multi-sensory Approach," in *Proc. Wearable Computers*, New York, NY, USA: ACM, 2018, pp. 25–32.

[8] 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," *J. of the Acoustical Society of America*, vol. 77, no. 1, pp. 247–257, Jan. 1985.

[9] M. A. Mines, B. F. Hanson, and J. E. Shoup, "Frequency of occurrence of phonemes in conversational english," *Language and speech*, vol. 21, no. 3, pp. 221–241, 1978.

[10] G. E. Peterson and H. L. Barney, "Control methods used in a study of the vowels," *Acoustical Society of America*, vol. 24, no. 2, pp. 175–184, 1952.

[11] Y. Jiao, F. M. Severgnini, J. S. Martinez, J. Jung, H. Z. Tan, C. M. Reed, E. C. Wilson, F. Lau, A. Israr, and R. Turcott, "A Comparative Study of Phoneme-and Word-Based Learning of English Words Presented to the Skin," in *Human Haptic Sensing and Touch Enabled Computer Applications*, Springer, 2018, pp. 623–635.