Training Behavior of Successful Tacton-Phoneme Learners*

Genna Angello, Guanlong Zhao, Akhil Babu Manam
and Ricardo Gutierrez-Osuna, *Senior Member, IEEE*

Abstract— **Sixteen subjects were trained with a user-driven interface to recognize sixteen tacton-phoneme pairs. To improve distinctiveness, tactons were created from vibrotactile stimuli that varied based on four forearm locations: dorsal/ventral, wrist/elbow; and four vibration patterns: continuous, short, fast, and frequency-modulated. To improve memorability, vibration patterns were assigned to similar-sounding phonemes. Subjects were asked to select the phoneme associated with each played tacton following a 30-minute training session consisting of interleaved learning and practice tests, during which they were free to choose the number of tactons and the difficulty of each practice test. Subjects who performed best on the final test chose to receive (i) more repetition during the practice tests and (ii) more difficult practice tests that resembled the final test. These two strategies gave successful learners greater opportunity for strengthening the association between tactons and phonemes.**

I. INTRODUCTION

To quickly learn speech communication through the skin, the individual tactile stimuli must be distinctive and memorable. Tactile stimuli are *distinctive* when they are easily discriminated from each other and *memorable* when their association with the corresponding speech unit is intuitive. There has been some success in designing distinctive vibrotactile stimuli along the dimensions of spatial location, duration, and frequency. However, the meaning of these stimuli has been difficult to learn and recognize after short training sessions, and learning often requires reduced presentation rates.

To address this issue, we present a training protocol that uses multiple strategies to facilitate rapid learning of phoneme-tactile associations. First, it uses a subset of 16 phones in American English across 4 categories: vowels, fricatives, stops and nasals/liquids. Second, each phone category is associated with a distinctive vibrotactile pattern to increase memorability. Third, tactors are placed near anatomical landmarks to facilitate discrimination. Finally, the protocol allows users to adjust the difficulty of each session, and interleave training and testing to exploit the so-called *testing effect*.

II. RELATED WORK

Gault [1] was the first to transmit speech onto the skin by amplifying the speech signal and projecting it onto deaf participants' fingertips. After 100 hours of practice, some participants could discriminate among vowel sounds and recognize a 58-word vocabulary. A major downside of this method is that it required slowing down the speech to around 4 seconds/word [2]. When a cochlear model was created to present traveling wave patterns to the forearm, 32 hours of training were required for word identification [3].

Since then, a number of speech-tactile encodings have been tested by assigning speech units to tactile stimuli that varied along as many as three of several tactile dimensions: spatial location, waveform/roughness, rhythm, vibration frequency, and amplitude [4]. When tactile stimuli differ only along one dimension, discrimination scores are reduced [5]. While the number of dimensions used is important for discrimination, the total number of tactons associated with speech units is also critical, with greater numbers impairing learning [6, 7]. Yet, even when as few as 9 tactons are used, recognition rates remain low, around 70% [8].

Practice testing is known to improve future memory performance more than passively studying information, a phenomenon known as the *testing effect* [9]. However, the difficulty of the practice test is also a crucial factor in improving memory [10]. By introducing variation or unpredictability into the practice test, learners are forced to improve their retrieval methods. Furthermore, difficult practice tests can reveal to learners exactly what they do not know.

III. METHODS

Sixteen participants received monetary incentives upon successful recognition of isolated phonemes associated with tactile stimuli presented with four C3 tactors (Engineering Acoustics Inc.). We pre-selected four distinct tactile stimuli (see Fig. 1), each associated with one phoneme class. Vowels were assigned a continuous vibration since they are generally sustained sounds; stops were assigned a short vibration since they involve a sudden and short release of airflow; fricatives were assigned a fast vibration since their energy is at high frequency; and liquids/nasals were assigned a frequency-modulated vibration that resembled their smooth characteristics. In addition, we selected four distinct arm locations: dorsal and ventral positions at the wrist and upper forearm [11]; see Fig. 2a.

During learning, participants could play the 16 tactons in any order; see "Learn" tab in Fig. 2b. Once participants felt ready to practice, they chose a subset of phonemes and whether to randomly rearrange their position on the grid –to avoid relying on cues that would not be available during the final test. During this practice test, participants were presented with a random tacton from the subset and were asked to click the corresponding button; phonemes in the chosen subset were highlighted in yellow. If incorrect, their selection was highlighted in red and the correct phone in

---

*Research supported by Facebook Inc.

G. Angello, G. Zhao, A.B. Manam and R. Gutierrez-Osuna are with the Computer Science & Engineering Dept. at Texas A&M University, College Station, TX 77843-3112 (corresponding author: rgutier@tamu.edu).

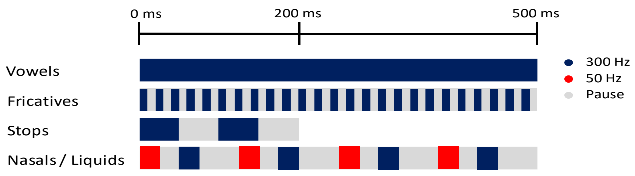

Figure 1. Vibrotactile stimuli for each phone category.

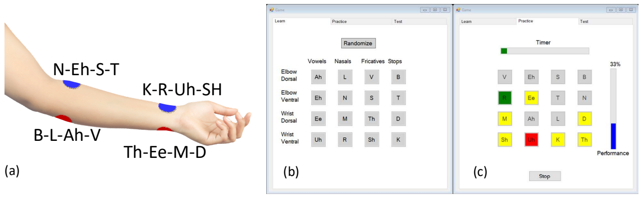

Figure 2. a: Phoneset and tactor location; b,c: user interface.

green; see "Practice" tab in Fig. 2c. Participants practiced until they felt ready to take the final test. The final test consisted of 80 trials over all 16 phonemes, delivered on a separate "Test" tab that only showed cumulative performance without highlighted phoneme feedback.

IV. RESULTS

Participants correctly recognized phonemes on the final test on an average of 73% (SD = 24.7%) of trials[1]. Fig. 3 shows the sequence of training events for two representative subjects, one who achieved a 97% recognition rate (S5), and a second who achieved 41% (S13) despite having selected similar phonemes to practice. S5 used more difficult practice tests than S13. First, S5 practiced on larger phoneme subsets than S13. S5 practiced with 4 phonemes only on 10% of the trials, whereas S13 used 4 phonemes on 54% of the trials. Unlike S13, S5 always randomized the location of phoneme answer choices (see Fig. 2c) when practicing all 16 phonemes. This better prepared S5 for how phonemes would be arranged during the final test, and strengthened tacton-phoneme associations. Finally, S5 was more successful at managing practice time, allocating extra practice to previously tested phoneme locations and increasing the number of response alternatives. S13, however, spent too many trials practicing only on the elbow (49%) or wrist locations (31%), practicing with both locations combined only on 20% of trials.

Correlation coefficients were computed between training choices and final test performance. More time spent learning tacton-phoneme associations was associated with poorer test performance (-0.4), and better performance was associated with more time spent practicing with random arrangement of phonemes (0.3), more tactons (0.4), and more practice trials (0.4). The strongest predictor was a greater number of practice trials tested over more tactons (0.7). Some successful subjects (i.e., S5) worked up to all 16 phonemes in practice, while others practiced with all 16 phonemes throughout the majority of their training session. Given the freedom to leave practice at any time, more difficult practice tests could have increased the temptation to switch to learning to resolve confusion. However, the behavior most predictive of future success on the final test was to remain in difficult practice sessions long enough to strengthen memory for tacton-phoneme pairs.

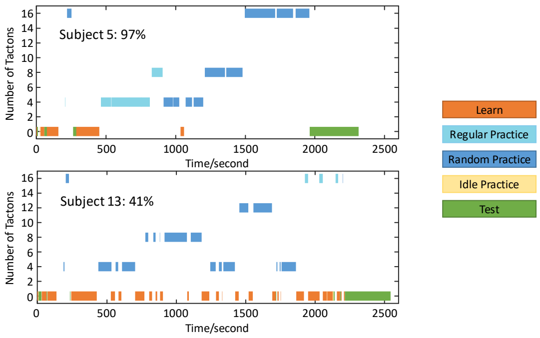

Figure 3. Learning strategies of two sample participants, including number of tactons practiced as a function of training time.

V. DISCUSSION

Optimal tacton-speech training should include repetition that ultimately leads to greater retrieval difficulty, forcing learners to create strong mental associations between each tactile stimulus and its associated speech meaning. Recognizing words by interpreting sequences of tactons requires quickly retrieving phonemes to keep up with the presentation rate. Future work will need to test whether this rapid-paced, repetitive tacton-phoneme training can transfer to word recognition tasks.

ACKNOWLEDGMENT

The authors would like to acknowledge Ali Israr and Frances Lau for feedback on study design and procedure.

REFERENCES

[1] R. H. Gault, "Progress in experiments on tactual interpretation of oral speech," *J. Abnorm. Soc. Psychol.*, vol. 19, pp. 155-159, 1924.

[2] J. H. Kirman, "Tactile communication of speech: A review and an analysis," *Psychol. Bull*, vol. 80, pp. 54-74, 1973.

[3] W. Keidel, "Electrophysiology of vibratory perception," in *Contrib Sens Physiol.* vol. 3, W. Neff, Ed., ed New York: Academic, pp. 1-79, 1968.

[4] K. E. MacLean, "Haptic interaction design for everyday interfaces," *Rev. Hum. Factors Ergon.* vol. 4, pp. 149-194, 2008.

[5] M. Clements, L. Braida, and N. Durlach, "Tactile communication of speech: Comparison of two computer-based displays," *J. Rehabil. Res. Dev.*, vol. 25, pp. 25-44, 1988.

[6] L. M. Brown, S. A. Brewster, and H. C. Purchase, "Multidimensional tactons for non-visual information presentation in mobile devices," in *Proc 8th Conf on Mobile HCI*, pp. 231-238, 2006.

[7] I. Politis, S. Brewster, and F. Pollick, "Speech tactons improve speech warnings for drivers," in *Proc 6th Intl Conf on Automot User Interfaces Interact Veh Appl.* pp. 1-8, 2014.

[8] M. Enriquez, K. MacLean, and C. Chita, "Haptic phonemes: Basic building blocks of haptic communication," in *Proc 8th Intl Conf on Multimodal interfaces*, pp. 302-309, 2006.

[9] H. L. Roediger III and J. D. Karpicke, "The power of testing memory: Basic research and implications for educational practice," *Perspect. Psychol. Sci.* vol. 1, pp. 181-210, 2006.

[10] R. A. Bjork, "Memory and metamemory considerations in the training of human beings," *Metacognition: Knowing about knowing*, pp. 185-205, 1994.

[11] R. W. Cholewiak and A. A. Collins, "Vibrotactile localization on the arm: Effects of place, space, and age," *Atten. Percept. Psychophys.* vol. 65, pp. 1058-1077, 2003.

---

[1] A large performance gap was found between the top performers (n: 8; mean: 89%; std: 6%) and the rest (n: 8; mean: 47%; std: 16.6%).