Passive Haptic Learning for Computer Stenography

A thesis presented to the Academic Faculty
by Timothy J. Aveni

In partial fulfillment of the requirements
for a Bachelor of Science in Computer Science
in the College of Computing

Georgia Institute of Technology
May 2019
Dedicated to Mirabai Knight and the Open Steno Project
Abstract

Passive haptic training can be used to teach motor skills using repeated tactile cues applied to the body. We explored the use of passive haptic training to teach participants how to produce various commonly-used phonemes with the stenotype keyboard, a phonetic chorded text input mechanism typically used by highly-trained stenographers. Using passive haptic training, we taught participants four common beginnings and four common endings to English words, then tested participants on thirteen combinations of these endings that produce full monosyllabic words in stenography. We found that, with an effective primer on the basics of the stenotype keyboard, PHL was able to teach not only how to write full words in stenography, but how to write individual reusable parts of words to be used as building-blocks for yet-unknown words.

Introduction

Stenography: Learning

Computer stenography enables text input at the rate of dictation. Stenographers find employment in court reporting and realtime captioning, lucrative professions that require a great deal of skill. Typically, stenographers train at for-profit stenography schools, where students spend thousands on hardware, software, and tuition to learn stenography. Reaching acceptable speeds to graduate from these schools can take anywhere between 1 and 6 years of intensive daily practice [1]. The process takes self-discipline that is often hard for students to maintain. Open-source software and hardware have begun to provide expanded access to the technology in contrast to proprietary stenography tools. Stenography schools, despite a national average dropout rate of 85%, are still the main resource for learning stenography [1].

Backed by open-source software and increasingly accessible stenography education, those without the opportunity to attend expensive stenography schools have begun to learn to use the technology, which can be useful beyond just professional opportunity. Since the stenotype keyboard is designed to be used in real-time, at the rate of dictation, it can be used in place of speech for people with speech impediments or Deaf individuals [1]. Stenography can be suitable for blind or low-vision people, who can use the technology to work as remote live-captioners without leaving their homes [1]. Since stenography is a more rapid text entry method than typing, even everyday electronic correspondence may be made more efficient using
stenography. Many people are reluctant to adopt complex new text entry methods, however, because of their steep learning curves. The barriers in the current stenography ecosystem exclude all but the most determined and privileged.

**Stenography: Technique**

On the stenotype keyboard, each key represents an individual English sound, rather than an alphabet letter [14]. Typing is performed in chords: pressing multiple keys at once to form one or two phonetic syllables at a time. Most fingers are responsible for two or more keys, and a single finger may have to press multiple keys at once.

The monosyllabic word “straps” combines six individual sounds from keys sweeping left-to-right across the keyboard.

![Figure 1. The keys associated with the word “straps”.

The keyboard does not include a separate key for each phoneme in the English language, so some phonemes are produced using combinations of keys. For example, the “M” sound on the left side of the keyboard is produced with the “P” and “H” keys together, so that the word “map” is produced by pressing the keys “P” (on the left), “H,” “A,” and “P” (on the right). Most stenotype keyboards do not have visible legends on the keys, which can reduce confusion in these nontrivial cases. Real-time computer stenography uses a translation engine that decodes each chord of keys, looks up the chord in a customizable dictionary, and outputs English text. In this
research, we use the free software Plover and its default dictionary to translate chords to written text [15].

In computer stenography, each chord, representing a single syllable, consists of beginning, middle, and end sounds. Beginning sounds are typed using the left half of the keyboard, middle sounds are typed by the thumbs, and end sounds are typed using the right half of keys. Beginning and end sounds are consonants, and middle sounds are vowels (e.g. “str-a-ps,” “c-a-t”).

Chords can be broken down into what refer to as “subchords” that can be reused to construct new chords. A stenographer who can type “straps” will also be able to type “traps,” “caps,” “maps,” and “laps,” provided the stenographer knows how to type the different beginning sounds. This enables a reuse and recombination of learned information, so trained stenographers can produce many unfamiliar words without having learned them explicitly. We introduced a model that trains stenography by subchord so that users can learn small, digestible portions of chords and then use this knowledge to discover new chords phonetically.

**Passive Haptic Learning**

Passive learning occurs when users acquire knowledge from ambient stimuli in the periphery of their attention. Most work on this topic uses audio or visual stimuli, but haptic stimuli can also provide passive training. Passive haptic learning (PHL) (or “passive tactile learning/ training”) can be used to train motor skills using repeated tactile stimuli, even while learners are distracted by other tasks [6]. Discrete motor actions, such as those used in typing, are effective targets for passive haptic training [16].

In this method, the skill (such as typing different groups of buttons) is translated into tactile cues that can be applied to the body (e.g. a vibrotactile “tap” on the fingers that type each button). Performance and knowledge of the skill improves by repeatedly applying these tactile cues, even while the user is focused on unrelated tasks (making learning “passive”) [17].

The hardware and software used in PHL are accessible and easily replicable, so passive haptic learning would be a welcome addition to the growing, open-source learning toolkit for stenographers. We believe that PHL may be useful for both beginning stenographers and
experienced, professional stenographers. Beginners to the input method must learn to write hundreds of basic phonemes and words before they can get started transcribing sentences, and it is difficult to stay motivated before gaining the ability to produce intelligible output [1]. Since PHL can teach input methods to people who don’t have any experience, passive training might help steno students overcome that initial hurdle and learn enough to get hooked. Professional stenographers need to keep their translation dictionaries updated so that they can write phrases related to current events, and staying on top of these additions and changes requires constant practice. If we can turn active practice into passive practice, we save time even for seasoned stenographers.

**Literature Review**

Haptic guidance was an early form of haptic learning, developed in the early 2000s to teach people motor skills [2, 3]. Early studies in haptic guidance were promising, so the idea expanded over time into the modern field of haptic training. In particular, passive haptic learning is a technique which uses repeated tactile stimuli to train complex motor skills passively, even concurrently with a distracting primary task.

Prior studies have created a framework for the domains in which passive haptic learning can be applied and the mechanisms by which they function most effectively [16]. In 2008, the Contextual Computing Group at Georgia Tech pioneered a system called Mobile Music Touch that teaches piano skills using gloves embedded with tactile vibration motors in which each finger is stimulated with a single motor [4]. When we apply a tactile ‘tap’ to the subject’s fingers in a particular sequence repeatedly over the course of a session and synchronize these stimuli with corresponding audio, users learn the sequence of actions required to play the tune without wearing the glove.

Later work used passive tactile stimuli to train discrete actions and their associated meanings, teaching users more than a fixed sequence of actions. One such study trained users on Morse code, a rhythmic text entry method [5]. The Morse study demonstrated that discrete actions can be trained passively, while the Mobile Music Touch study trained a series of actions in a sequence. It also suggested that explicit information encoded by the haptic stimuli can be extracted by the user, e.g., to read and write Morse code.
Passive haptic training can also teach chorded input methods, which involve simultaneous actions (pressing more than one key or button at once for a particular input). Braille text entry takes place on a six-key keyboard — one for each dot in the six-dot grid — and a character is written by pressing the appropriate key for each dot in the character’s Braille representation. Passive haptic training has been used to teach this Braille input method [6]. Early inquiry in this study found that the fingers are not able to distinguish stimuli when more than one finger is vibrated at once, so Braille is taught by administering the stimuli of a chord in sequence. After learning passively, subjects are able to read and write Braille, which continues to demonstrate that explicit information can be extracted from the tactile training from the learning session.

The haptic interface used for training must be designed to indicate particular inputs, like on a computer keyboard number pad, where a finger can press the top key, bottom key, or middle key [7]. This method allows for each finger in an input method to have multiple responsibilities, distinguished spatially, and for these responsibilities to be taught with passive haptic learning.

**Goals**

Although faster, more ergonomic text entry methods than the standard computer keyboard have existed for some time, their learning curve, paired with inaccessible or inadequate training methods, have served to slow their adoption. By creating a training mechanism for the stenotype keyboard, we hope to eliminate some of the barriers preventing rapid adoption of this input method.

Prior work has shown passive haptic training may help to train discrete actions with their associated meanings, simultaneous actions in chorded input, and spatial tasks. Computer stenography brings each of these training challenges together into a single complex input method. By combining the findings from the studies exploring text input on a Braille keyboard and a computer keypad, we intend to demonstrate the applicability of passive haptic learning to the acquisition of computer stenography skills. In addition to these challenges addressed in past research, we will train users to type words that use interchangeable subchords and to type words that use differing spatial actions for each finger within a chord. This higher-order inquiry will combine multiple established techniques in passive training, suggest new capabilities for the
training method, and propose a complex input method as a new use case for passive haptic learning.

Since stenography is a complex input method in which combining multiple distinct inputs is nontrivial, we seek to demonstrate that, having learned individual phonetic components, PHL learners are able to compose entire words.

**Methods**

We have performed a few pilot studies in pursuit of this goal. Initial results were underwhelming, but through persistent iteration, we have found an effective setup for training stenography.

**Apparatus**

For the initial prototype, gloves similar to those used in the study that trained users to type a sequence on a computer keypad were built to administer stimuli to the top or bottom of each finger, in sequence.

An in-depth study was conducted in 2015 to understand the optimal placement for tactile motors for learning, alongside the optimal motor to use [8]. This study found that stimuli were better recognized on the dorsal side of the hand by a small measure and that stimuli were better recognized closer to the palm of the hand. In our stenography studies, to balance the stimuli vertically, the “bottom” stimulus (the one closer to the palm of the hand) is on the ventral side of the finger, and the top stimulus is on the dorsal side. This should also help distinguish the stimuli better than placing both vibration motors on the same side of the finger.

**Study 1**

We attempted to train six words (stick, stein, lick, lay, nay, nine) over three twenty-minute sessions. We did this with a simple sequential stimulus pattern, vibrating each motor (on the top or bottom of each finger) in series. We tested the learned words, alongside three “combined” words (stay, line, nick) that weren’t taught explicitly, but were comprised of the same parts used in the learned words. In the testing session, the word was shown on the screen and played through headphones, and the participant had three chances to type each word (with
all-or-nothing feedback on each word entry). We used a video game without linguistic content and without audio as a distraction task.

We also included a “training” session at the beginning of the study to acclimate users to the input method. In this session, users were presented with increasingly complex stimulus patterns and asked to replay them immediately into the stenotype machine. This helped participants understand, before starting the PHL sessions, how to press down entire chords at once without pressing extra buttons, which can be difficult to get right at first.

Discussion

We did not see evidence of learning in this study (p > 0.5 for many of the tested hypotheses). In general, participants seemed overwhelmed with the stimuli. Even though we administered stimuli sequentially as in the Braille study, there was just too much going on at once and participants seemed frustrated. After this, we made a few simplifying changes:

- When vibrating a full word, we added temporal space between the three sections of the keyboard (the left, middle (thumbs), and right). This we did to avoid sending too much information at once.
- In some trials, we chunked both the audio and stimuli by keyboard section. To teach “straps,” for example, we played audio for “str,” followed by the appropriate vibration, then the audio for “a,” with the corresponding vibration, then “ps,” with the remaining vibrations.
- Some participants took a long time on the initial training session, or failed to complete it at all. This indicated that, independently of PHL, the core stimuli were not effective in indicating the correct keypresses. Some participants also mentioned that they had trouble distinguishing between the top and bottom stimulus on a particular finger. We added a difference between the “feel” for the top and bottom keys so they are better distinguished, because the vertical distance between motors on the fingers isn’t very large. When indicating a keypress in the top row, the gloves issue one long vibration, and a keypress in the bottom row is indicated by two short vibrations. To indicate a keypress with both of a finger’s keys, we apply a “jitter” vibration to both the top and the bottom motors.
Study 2

After applying these changes, we tried a simpler study with only two words (straps and prong). Each participant was randomly assigned to one of these words to learn with chunked audio (using audio for “str,” “a,” and “ps” separately, for example), and they learned the other word without chunked audio (but still with lightly-separated tactile stimuli). The order of the words was also randomized, for a total of four conditions. The improved stimuli seemed to be successful in teaching stenography chords, with both the chunked and unchunked stimuli (though with marginally better performance in the chunked condition).

In the first study, many participants failed to use the full keyboard to type a word with a beginning, middle, and end. In the two-word study, since the PHL sessions had separated audio for each word chunk, we also tested each chunk separately before having the participant combine each sub-chord into the larger word. We added visuals to the testing interface to steer focus onto a particular part of the keyboard. We also removed accuracy feedback from the test to ensure that any learning was happening because of the PHL session, and not because of reinforcement from the testing procedure.

Figure 2. The testing interface used during the second study.

Discussion

This study showed more promising results, with some tests reaching statistical significance at the $\alpha = 0.05$ level. We speculate that the most important changes that made this study
successful were the improved top-bottom distinction in the tactile stimuli and the separately-tested word chunks, which reinforced the core stenography principle of typing phonemes left-to-right.

**Study 3**

Following this research, we chose to dive into a study with a control that we can use to demonstrate learning in our technique versus the condition with no tactile stimuli at all. Now that we had a setup for effective learning, we decided to expand the set of learned words and attempt again to have the participant mix-and-match different pieces of words.

To determine which phonemes we should teach participants to maximize utility, we analyzed public word frequency data from Mirabai Knight, the founder of the Open Steno Project [9]. We split each word into its beginning sounds, on the four fingers of the left hand, and ending sounds, on both hands' thumbs and the right hand's fingers. We computed each subchord's frequency by considering the frequency of each word for which the subchord was a component. We identified a set of common beginning and ending inputs that can be mixed and matched to create English words: to begin words, we chose the “s,” “w,” “m” and “f” sounds, and to end words, we chose “-all,” “-ore,” “-ock,” and “-un”. Many of the combinations of these sounds, like “wall,” “mock,” and “fun,” are English words that can be written in a consistent fashion in stenography.

We taught these stimuli in individual sections, then tested each phoneme individually, combining the stimuli into words during the testing session. In the testing phase, we completely removed any visuals containing information about the tested word or which part of the keyboard should be used to write it. By focusing the user on the audio, we emphasized the phonetic nature of stenography and avoided confusion with some words (like “for,” “wore,” and “one”) that are homophones or sound similar to other words.

We showed a diagram during the study so that participants who remember the stimuli they experienced can easily map these stimuli to actions. This we derived from the Braille study, which uses a similar diagram during reading and writing tests [6]. We also showed a diagram reminding users of the left-middle-right ordering in stenography.
Distraction task

During passive learning, the participant is instructed to play SpikeDislike2 as a distracting, unrelated task [13]. Participants are told to focus only on the game and to ignore the background stimulation. This game was chosen based on a few characteristics used in past PTL research, including its fast-paced nature, which serves to avoid letting the user rest to focus on the passive stimuli [6].

This game requires no audio and minimal textual content and is played using just one finger operating the spacebar on the standard computer keyboard. No part of the game contains
explicit or implicit information about stenography. The highest score attained in the game is recorded for each session.

**Results**

The Uncorrected Error Rate, or UER, of a chord is a number in the range [0, 1] related to the number of correct and incorrect keys pressed in a chord [10]. An incorrect key is one that was pressed and should not have been pressed, or a key that was not pressed that should have been.

\[
UER = \frac{\# \text{ incorrect keys}}{\# \text{ incorrect keys} + \# \text{ correct keys}}
\]

Table 1 shows the average error for each subchord and full word for the test that is administered immediately after the word is learned. The final column shows the p-value for a one-tailed two-population t-test on these error rates. Every individual subchord, and every full word that was taught explicitly, shows statistical significance at the \( \alpha = 0.05 \) level. Between the experimental and control groups, the average error across all of these chords (indicated in the final row of the table) also shows a statistically significant difference.

Nine “unfamiliar” words were never taught explicitly but were tested immediately after both of their constituent subchords were taught. For each of the words in Table 2, one of the two parts of the word had already been taught by the time the participant started the pre-test. For example, the test surrounding “wall” tested “wore,” because this learning session teaches “w” and a prior learning session had already taught “-ore.” Here, we tested the participant’s ability to combine old knowledge with new knowledge, applying the rules of stenography to the learned information.

The UER difference is statistically significant for many of the unfamiliar words. The words with the least statistical significance (“mall,” “fall,” and “one”) use subchords that showed less effective learning in the individual tests.
A one-tailed t-test comparing the maximum score attained in the distraction task between the experimental and control groups did not demonstrate statistical significance, indicating a lack of evidence to suggest that the experimental condition has an impact on distraction task performance.

Table 1. Immediate post-test UER for explicitly-taught words.

<table>
<thead>
<tr>
<th></th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.0003</td>
</tr>
<tr>
<td>OR</td>
<td>0.00</td>
<td>0.33</td>
<td>0.00</td>
<td>0.67</td>
<td>0.56</td>
<td>1.00</td>
<td>0.67</td>
<td>1.00</td>
<td>0.75</td>
<td>0.85</td>
<td>1.00</td>
<td>1.00</td>
<td>0.0126</td>
</tr>
<tr>
<td>SOR (sore)</td>
<td>0.08</td>
<td>0.50</td>
<td>0.00</td>
<td>0.75</td>
<td>0.25</td>
<td>0.86</td>
<td>1.00</td>
<td>1.00</td>
<td>0.72</td>
<td>0.85</td>
<td>0.94</td>
<td>1.00</td>
<td>0.0035</td>
</tr>
<tr>
<td>W</td>
<td>0.33</td>
<td>1.00</td>
<td>0.00</td>
<td>0.17</td>
<td>0.33</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.0005</td>
</tr>
<tr>
<td>AUL</td>
<td>0.50</td>
<td>0.74</td>
<td>0.28</td>
<td>0.27</td>
<td>0.87</td>
<td>0.80</td>
<td>1.00</td>
<td>0.72</td>
<td>1.00</td>
<td>1.00</td>
<td>0.67</td>
<td>1.00</td>
<td>0.0143</td>
</tr>
<tr>
<td>WAUL (wall)</td>
<td>0.43</td>
<td>0.71</td>
<td>0.00</td>
<td>0.13</td>
<td>1.00</td>
<td>0.73</td>
<td>1.00</td>
<td>0.83</td>
<td>1.00</td>
<td>1.00</td>
<td>0.87</td>
<td>1.00</td>
<td>0.0094</td>
</tr>
<tr>
<td>PH</td>
<td>0.67</td>
<td>0.69</td>
<td>0.00</td>
<td>0.67</td>
<td>1.00</td>
<td>0.04</td>
<td>1.00</td>
<td>1.00</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.0109</td>
</tr>
<tr>
<td>OBG</td>
<td>0.69</td>
<td>0.68</td>
<td>0.04</td>
<td>0.00</td>
<td>0.33</td>
<td>0.24</td>
<td>1.00</td>
<td>0.97</td>
<td>0.98</td>
<td>0.89</td>
<td>1.00</td>
<td>0.85</td>
<td>0.0003</td>
</tr>
<tr>
<td>PHOOGM (mock)</td>
<td>0.57</td>
<td>0.61</td>
<td>0.00</td>
<td>0.33</td>
<td>0.50</td>
<td>0.00</td>
<td>0.83</td>
<td>1.00</td>
<td>0.70</td>
<td>0.69</td>
<td>0.96</td>
<td>0.88</td>
<td>0.0011</td>
</tr>
<tr>
<td>TP</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
<td>0.00</td>
<td>0.83</td>
<td>0.00</td>
<td>1.00</td>
<td>0.89</td>
<td>0.72</td>
<td>0.75</td>
<td>0.44</td>
<td>1.00</td>
<td>0.0013</td>
</tr>
<tr>
<td>UPB</td>
<td>0.06</td>
<td>0.41</td>
<td>0.69</td>
<td>0.22</td>
<td>0.68</td>
<td>0.59</td>
<td>0.85</td>
<td>0.88</td>
<td>0.75</td>
<td>0.88</td>
<td>1.00</td>
<td>0.33</td>
<td>0.0190</td>
</tr>
<tr>
<td>TPUPB (fun)</td>
<td>0.00</td>
<td>0.29</td>
<td>0.57</td>
<td>0.39</td>
<td>0.29</td>
<td>0.00</td>
<td>0.83</td>
<td>0.95</td>
<td>0.75</td>
<td>0.82</td>
<td>0.88</td>
<td>0.67</td>
<td>0.0001</td>
</tr>
<tr>
<td>average</td>
<td>0.28</td>
<td>0.51</td>
<td>0.13</td>
<td>0.30</td>
<td>0.55</td>
<td>0.44</td>
<td>0.93</td>
<td>0.94</td>
<td>0.85</td>
<td>0.89</td>
<td>0.90</td>
<td>0.89</td>
<td>5.9 x 10^-7</td>
</tr>
</tbody>
</table>

Table 2. Immediate post-test UER for unfamiliar words.

<table>
<thead>
<tr>
<th></th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WOR (wore)</td>
<td>0.08</td>
<td>0.67</td>
<td>0.00</td>
<td>0.60</td>
<td>0.50</td>
<td>0.60</td>
<td>1.00</td>
<td>0.95</td>
<td>0.72</td>
<td>0.88</td>
<td>1.00</td>
<td>1.00</td>
<td>0.0011</td>
</tr>
<tr>
<td>SOBG (sock)</td>
<td>0.40</td>
<td>0.43</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.40</td>
<td>1.00</td>
<td>1.00</td>
<td>0.91</td>
<td>0.81</td>
<td>0.86</td>
<td>0.67</td>
<td>0.0001</td>
</tr>
<tr>
<td>PHOR (more)</td>
<td>0.40</td>
<td>0.88</td>
<td>0.00</td>
<td>0.67</td>
<td>0.88</td>
<td>0.67</td>
<td>1.00</td>
<td>1.00</td>
<td>0.71</td>
<td>0.89</td>
<td>1.00</td>
<td>0.83</td>
<td>0.0261</td>
</tr>
<tr>
<td>PHAUL (mall)</td>
<td>0.89</td>
<td>0.90</td>
<td>0.33</td>
<td>0.75</td>
<td>0.89</td>
<td>0.46</td>
<td>1.00</td>
<td>0.87</td>
<td>0.83</td>
<td>0.83</td>
<td>0.85</td>
<td>1.00</td>
<td>0.0509</td>
</tr>
<tr>
<td>WOOG (wok)</td>
<td>1.00</td>
<td>0.71</td>
<td>0.00</td>
<td>0.00</td>
<td>0.55</td>
<td>0.89</td>
<td>0.80</td>
<td>1.00</td>
<td>1.00</td>
<td>0.67</td>
<td>1.00</td>
<td>0.86</td>
<td>0.0408</td>
</tr>
<tr>
<td>TPAUL (fall)</td>
<td>0.75</td>
<td>0.78</td>
<td>0.33</td>
<td>0.33</td>
<td>0.85</td>
<td>0.67</td>
<td>1.00</td>
<td>0.75</td>
<td>0.63</td>
<td>0.92</td>
<td>0.67</td>
<td>1.00</td>
<td>0.0504</td>
</tr>
<tr>
<td>TPO (for)</td>
<td>0.00</td>
<td>0.55</td>
<td>0.40</td>
<td>0.67</td>
<td>0.77</td>
<td>0.50</td>
<td>0.80</td>
<td>0.88</td>
<td>0.72</td>
<td>0.80</td>
<td>0.86</td>
<td>1.00</td>
<td>0.0053</td>
</tr>
<tr>
<td>WPB (one)</td>
<td>0.40</td>
<td>0.60</td>
<td>0.67</td>
<td>0.13</td>
<td>1.00</td>
<td>0.83</td>
<td>0.80</td>
<td>0.91</td>
<td>0.88</td>
<td>0.91</td>
<td>1.00</td>
<td>0.60</td>
<td>0.0531</td>
</tr>
<tr>
<td>SUPB (sun)</td>
<td>0.00</td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
<td>0.60</td>
<td>1.00</td>
<td>0.73</td>
<td>0.88</td>
<td>0.78</td>
<td>1.00</td>
<td>0.60</td>
<td>0.0079</td>
</tr>
<tr>
<td>average</td>
<td>0.44</td>
<td>0.64</td>
<td>0.19</td>
<td>0.35</td>
<td>0.76</td>
<td>0.62</td>
<td>0.93</td>
<td>0.90</td>
<td>0.81</td>
<td>0.83</td>
<td>0.92</td>
<td>0.84</td>
<td>0.0010</td>
</tr>
</tbody>
</table>
We saw patterns across participants in some of the word parts that showed less consistent learning. In the PH chord, for example, many participants pressed the T and P keys, a one-key shift from the correct chord. In AUL, much of the error resulted from a shift to the left in the right hand, where many participants used their middle instead of their ring finger. Another common error was a vertical shift in the W chord, where many participants pressed the left P key instead. The heatmaps shown in Figure 5 are generated from the keys pressed by every member of the experimental group for every test of the associated chord after the learning session containing the chord.

![Heatmaps for the PH and AUL chords during post-tests.](image)

**Discussion**

Our results suggest that passive stimuli can help beginners learn stenography, and we hope to address the learning crisis in the industry. Computer stenography is a fundamentally more complex input method than those taught in prior PHL research. We have shown that, with an effective apparatus, we may use passive tactile training to help teach an input method that:

- Uses chords, not individual keys, for text entry
- Constructs these chords as combinations of exchangeable, phonetic subchords
- Assigns more than one key to each finger

Each learned primitive in stenography can be applied to new words. Using our teaching structure, which taught components used commonly in English, users were able to combine learned subchords to type unfamiliar words, which is fundamental to the modular nature of stenography. After using passive haptic training, participants were able to write new
stenography words that they never learned explicitly, which suggests that knowledge gained through PHL can be composed, beyond simple muscle-memory replay.

Even in these cases when error was high (as in the heatmaps shown in Figure 5), participants frequently had the right idea about which keys to press. Often, the error was a one-key vertical or horizontal shift in one or more of the user's fingers. This suggests that much of the error in learning came not from general confusion, but because of specific misunderstood tactile stimuli. We believe that, by iterating on the specifics of the tactile interface, we can improve stimulus localization and increase learning accuracy. Past PHL research can inform these changes. Stimuli administered by gloves are most effective when the haptic actuators fit snugly to a narrowly-localized region on the hand [16].

Not every error is attributable to the learning mechanism. The stenography dictionary included with Plover, which is aggregated from a number of sources, contains a number of entries for “misstrokes,” which are dictionary entries that do not have a strict basis in theory — and would not typically be considered valid chords — but are included because users often accidentally input them while attempting to enter a particular chord [11]. The dictionary maps the misstroke to the correct word, like a deterministic autocorrect mechanism. A frequent mistake for participants typing the PH chord was to use TP instead, which is a single left shift from the correct keys. This misstroke is not exclusive to PHL learners. Indeed, the Plover dictionary includes misstroke entries for many words starting with the “m” sound, but which were typed by pressing TP or TPH instead of PH [12].

One example is a misstroke entry for “mosquito.” If a Plover user slips and inputs the chords for “fosquito,” the default dictionary will still translate the word correctly to “mosquito.” Because of these misstroke entries, we conclude that errors in entering PH may be partially caused by the difficulty of the input method, which can affect even professional stenographers, rather than a flaw in our training apparatus.

**Future Work**

Results suggest that PHL may be a practical mechanism for teaching some amount of stenography skill. Is it the most effective learning method? Future work should compare
PHL-guided stenography education against traditional education to determine whether PHL is suitable to act as a replacement, rather than just a supplement, for stenography education. Some amount of basic theoretical education will always be required to get started with such an exotic input method, but much of stenography training past the initial theory hurdle requires simple memorization, which PHL is well-suited to provide.

The stimulus apparatus potentially still has room for improvement. Error in those chords in which the participants’ button presses were simply shifted horizontally or vertically might occur because of difficulty telling between the stimuli in the fingers. It is worth continuing to look for improvements in the stimulus apparatus to minimize ambiguity.

Further research can examine teaching every English phoneme with a wearable, tactile system. As users learn more phonemes, the number of words that can be typed by combining these phonemes increases dramatically. Teaching every phoneme with PHL could bootstrap beginners who want to get started writing many real words with stenography.

Some prior research suggests that passive stimulation can help improve or maintain speed in motor skills [7]. This could be instrumental in developing a streamlined, wearable system to aid stenographers with learning and practice.

Conclusion

A well-designed haptic interface shows promise in training computer stenography skill. This skill expands past the simple knowledge of how to write explicitly-trained words. By training and reinforcing words using interchangeable subchords, the haptic interface is able to help the user grasp more general principles of the technique behind stenography. An effective passive training interface for stenography need not train every word in the English language; by training reusable word segments, the interface enables the user to type even words that were never trained explicitly.
References


[12] “main.json,”


