Unipad: Single Stroke Text Entry With Language-based Acceleration

I. Scott MacKenzie  
Dept of Computer Science and Engineering, York University  
Toronto, Ontario, Canada M3J1P3  
smackenzie@acm.org

Javier Chen  
Dept of Computer Science and Engineering, York University  
Toronto, Ontario, Canada M3J1P3  
yu256619@cse.yorku.ca

Aleks Oniszczak  
Dept of Computer Science and Engineering, York University  
Toronto, Ontario, Canada M3J1P3  
aleks@cse.yorku.ca

ABSTRACT
A stylus-based text entry technique called Unipad is presented. Unipad combines single-stroke text input with language-based acceleration techniques, including word completion, suffix completion, and frequent word prompting. In a study with ten participants, entry rates averaged 11.6 wpm with 0.90% errors after two hours of practice. In follow-on sessions to establish the expert potential, four users entered “the quick brown fox” phrase repeatedly for four blocks of 15 minutes each. Average rates on the last block ranged from 17.1 to 35.1 wpm, with peak rates reaching 48 wpm.

Author Keywords
Mobile text entry, word completion, word prediction, stylus input.

ACM Classification Keywords
H.5.2 User Interfaces Input devices and strategies

INTRODUCTION
The push toward mobility in computing has rekindled interest in one of the oldest areas in office automation: text entry. This is due in no small measure to the huge popularity of SMS messaging on mobile phones, where it is estimated that one trillion SMS messages were sent in 2005 alone (mda-mobiledata.org). Related services on other mobile devices are also gaining popularity. In view of this, researchers and companies are ambitiously searching for improvements to mobile text entry techniques; improvement that are easy to learn, easy to use, and capable of high speed entry of machine-readable alphanumeric text.

Viable entry methods depend on context. Two broad categories are key-based and stylus-based devices. The former include most mobile phones and pagers, the latter include personal digital assistants (PDAs) tablet PCs and portable gaming devices. Stylus-based input is further divided between handwriting with automatic recognition and tapping on soft or virtual keyboards. Our focus here is stylus-based text entry using handwriting recognition.

This paper presents the design and evaluation of Unipad. Unipad combines single-stroke handwriting recognition with language-based acceleration. Features promoting high-speed entry are movement minimization and reduced attention demand. These are explained in the discussions that follow, after which an evaluation is presented.

Entry Speed
Mobile text entry speeds are unlikely to rival those for desktop touch typing, where rates of 60+ wpm are common. However, such rates are rarely achieved without years of practice.

Numerous text entry techniques exist, either commercially or in research prototypes. Although provocative, comparing entry rates among techniques is misleading unless framed in the proper context, such as the type of task or users’ prior computing experience or practice with the technique. Soft keyboard entry rates, for example, have been measured at 40+ wpm [11] and even 70+ wpm (www.fitaly.com), but these rates were achieved following considerable practice. When users first confront a soft keyboard, entry rates are much lower, on the order of 18-28 wpm for a Qwerty layout [3, 11, 13], or as low as 5-7 wpm for an unfamiliar layout or method [12, 13, 16].

Handwriting speeds are in the 13-22 wpm range [1]. However, when coupled with recognition, the interaction is quite different, since users must attend to and respond to the ongoing process of recognition and error correction. As well, most implementations impose constraints on users’ writing style to match the requirements of recognition (e.g., stroke sequence and/or direction, or stylized strokes). Achieving higher speeds with handwriting recognition, therefore, necessitates new strategies, such as using a simplified stroke alphabet or reducing the number of required strokes using word completion. Unipad includes both these features, as now explained.

Simplified Strokes
An 1895 report on typewriting indicated that “the strain of sight typing on the eyes and mind increases with speed, until a point is reached where it cannot be kept up for any length of time” (quoted from [5]). And so, the skill of “sight typ-
“parts”, “parted”, “parting”, or “partly” can be entered with just four strokes (Note: “part” appears in the candidate list after two strokes):

```
·
```

The desired word appears in the candidate list after the third stroke. Tapping on it, selects the word and delivers it to the edit buffer with a terminating SPACE.

Size Precedent Sort Order

Although the most obvious way to present the candidate list is sorted by word probability, we take a slightly different approach. The words are sorted first by size, then lexically within size. Thus, more information is inherently encoded in the ordering. The expected effect – although difficult to quantify – is to reduce the visual scan time to find the desired word.

Suffix Completion

Reduced keystroke suffix completion in languages such as English have a long history in assistive technologies, where common suffixes are appended typically with a single keystroke [6]. With our system, suffix completion mode is entered when a candidate word is selected with a top-left to bottom-right stroke (¢). The candidate word list is replaced with a set of popular suffixes. The size of the list is programmable, however our tests use a list size of twelve: “s”, “es”, “er”, “est”, “ly”, “able”, “ful”, “ion”, “ive”, “ment”, “ness”. Tapping a suffix appends it and adds a terminating SPACE.

As an example, “parts”, “parted”, “parting”, or “partly” can be entered with just four strokes (Note: “part” appears in the candidate list after two strokes):
Suffix completion reduces keystrokes two ways. The first was just demonstrated, but, additionally, words containing suffixes are eliminated from the dictionary (with counts added to the base word count). The ancillary benefit is that the candidate list contains a richer set of possibilities. This also reduces keystrokes since a given word stem expands to a more diverse set of choices.

Although suffix completion can save keystrokes, our experience suggests that the primary benefit is in simplifying the interaction primitives. Straight-line strokes are fast (<200 ms [4, 7]) and users quickly assimilate the “ṣ·” pattern using chunking or unitization [18].

**Frequent Word Prompting**

At the beginning of each word the input pad is blank, as there is no word stem upon which to generate candidates. This void is filled with a list of the most frequent words in the dictionary. We call this the “frequent word prompt”. The size of the list is programmable, however our tests use a list size of twelve. This coupled with our default dictionary yields “for”, “the”, “you”, “and”, “was”, “that”, “of”, “a”, “to”, “is”, “in”, “it”. A tap selects the word, delivering it to the edit buffer with a terminating SPACE.

**Other Features**

Unipad includes several other features supporting general purpose text entry and editing. Uppercase and caps lock modes are implemented using a bottom to top (1) stroke, in a manner that mimics commercial systems, such as Graffiti on Palm PDAs. A right to left stroke (←) serves as a backspace.

Several soft buttons are included, for example to clear the edit buffer (CLEAR) or enter a carriage return (ENTER). There is a button to bring up a soft keyboard (SOFTKEY) for inputting punctuation and other symbols. Dedicated character-delete (CHAR<>) and word-delete (WORD<>) buttons are implemented to facilitate users’ editing strategies.

**Unipad Interface**

The general operation of Unipad and its language-based acceleration techniques are now illustrated.

Figure 2 is a screen snap of Unipad working in concert with our text entry evaluation software. Phrases are presented to the user for input. During entry, performance measures are gathered and saved for follow up analyses. In the example, the user is entering “hours”. After two strokes, the desired word has yet to appear in the candidate list, so the user proceeds to enter the third stroke. The digital ink trails accompanying input appear directly on top of the candidate list.

In Figure 3, the user is about to enter “of”. Since this word appears in the frequent word prompt, a simple tap suffices.

Figure 4 shows interaction using suffix completion. To enter the word “parking” the user has a few options. The plain single stroke method requires eight strokes:

```
 parking
```

Using suffix completion however, only five strokes are required:

```
par· (s = tap entry in suffix list)
```

The fourth stroke is a TOP-LEFT TO BOTTOM-RIGHT gesture beginning on top of the entry “park”, which appears in the candidate list with the stem “par”. This brings up the suffix completion list (see figure). A final tap on “ing” completes the word and adds a terminating SPACE.
Keystrokes Per Character

Although isolated examples such as those above serve to explain interaction features, they fail to characterize or quantify the overall benefits in terms of interaction primitives. For this, we use Keystrokes Per Character (KSPC) [8]. KSPC is the number of keystrokes (in this case, “stylus strokes”) required, on average, to generate each character of text using a given interaction technique in a given language. For plain single stroke text entry, KSPC = 1, since each stroke generates a character. For techniques such as multitap on mobile phones, KSPC > 1, since more than one key press is required to produce each character. However, KSPC < 1 if language-based acceleration techniques are used, as just discussed.

Ideally, the lowest possible KSPC in a word completion text entry system occurs where each keystroke generates a word. With our default dictionary, the mean word size is 4.59 characters. Adding 1 for a terminating SPACE character, the lowest possible value for KSPC is 1 / 5.59 = 0.179. (Prediction can also work at the phrase level [2, 15], but this is not explored here.)

We computed KSPC using our default dictionary and under numerous settings for Unipad’s acceleration features. Figure 5 shows results for various list sizes for the candidate list (“CL”) and frequent word prompt. The calculations do not include the keystroke reduction due to suffix completion.

The point at the top-left in Figure 5 shows the result if language-based acceleration is not used: KSPC = 1. The effect of adding language-based acceleration is seen moving across and down the figure. The point at the bottom-right has KSPC = 0.428. Adding suffix completion reduces this by a very small amount; however, as noted above, the primary benefit with suffix completion is in simplifying the stroke primitives.

The KSPC figure just cited represents a best-case scenario. That is, Unipad’s acceleration features afford English text entry with just under 0.5 keystrokes per character. However, several important issues remain. First, it is not clear that users will actually use the acceleration features effectively. Missed opportunities tend to push KSPC up, to the point where users are interacting essentially in plain single-stroke text entry mode. Clearly, learning is important as well. With practice, a user’s KSPC may fall as features are learned and exploited at the earliest opportunity. And finally, it is not clear that reductions in KSPC yield a corresponding increase in throughput, since the acceleration features also add cognitive demands to the interaction. To test for the potential benefits of Unipad, a user study is required. This is presented in the next section.

METHOD
Participants
Ten paid volunteer participants (6 male, 4 female) were recruited from the local university campus. Participants ranged from 18 to 35 years (mean = 26.1, sd = 5.9). All were daily users of computers, reporting 3 to 8 hours usage per day (mean = 5.6, sd = 2.2). Self-assessed typing speeds ranged from 35 to 52 words per minutes (mean = 42.2, sd =
4.7). Two indicated prior experience with Graffiti, describing their expertise as “novice” or “intermediate”.

Apparatus
The experiment was conducted in a research lab using a standard desktop computer system equipped with a 13.3 inch Wacom PL-400 tablet for stylus entry. The PL-400 tablet is both a digitizer for input and a 1024 x 768 LCD colour screen for output.

The experimental software was an in-house Java application for text entry evaluation. The candidate list appeared at the first stroke. Other settings were as follows:

- Candidate list size: 5
- Frequent word prompt list size: 12
- Suffix list size: 12

Upon launch, the program reads series of 500 phrases ranging from 16 to 43 characters (mean = 28.6) [9]. There were 2712 total words, including 1163 unique words. Words ranged from 1 to 13 characters (mean = 4.46). The correlation between the letter frequencies in the phrase set and those in our reference corpus was r = .9541. During execution, phrases are selected randomly and presented to the participant for input.

Procedure
Participants completed a pre-test questionnaire soliciting demographic and computer usage information (results cited above) and a post-test questionnaire on their subjective impressions of the methods (discussed later).

Each participant completed two one-hour sessions. Plain single stroke text entry was used in the first session. This served to bring participants up to speed on learning the single stroke alphabet (see Figure 1). In the second session, word-completion was introduced.

Prior to collecting data, the experimenter briefly explained the task and demonstrated the software including the method to correct errors and to terminate entry at the end of a phrase. The single-stroke alphabet was displayed on a wall chart. Acceleration features were explained and demonstrated before the second session.

The participants were instructed to enter the phrases “as quickly and accurately as possible”. The task was described as “similar to sending an email message to a friend” – the message should be understandable, but not necessarily perfect.

The goal of the instructions was to elicit typical text input behaviours and therefore improve the external validity of the experiment. However, a side effect is that there are two types of errors: those corrected, and those remaining in the transcribed text. The former are reflected in the measured KSPC as additional strokes to backup and correct mistakes. The latter are measured by comparing the presented and transcribed text phrases and computing the error rate. The error rate was computed using the minimum string distance method [17].

Participants entered a few warm-up phrases during which time they could ask questions about the procedure. Data collection began with the first stroke for each phrase and ended with a tap on the ENTER soft-key. Participants were allowed to rest at their discretion between phrases.

Each session was divided into 4 blocks of 15-minutes each. Hence blocks 1-4 refer to data from session one (plain single stroke entry method), while blocks 5-8 are for session two (single stroke method with word completion). Participants were encouraged to take a short break between blocks.

The software recorded a timestamp for each stroke as well as a variety of statistics for follow-up analyses.

RESULTS AND DISCUSSION
In all, participants entered 754 phrases using the single-stroke method (blocks 1-4) and 937 phrases using the single-stroke method with word completion, suffix completion, and frequent word prompting (blocks 5-8).

Outliers
In word completion mode, an anomalous behaviour was observed. The ENTER soft key, to terminate entry of a phrase, was occasionally tapped to select a word. This prematurely terminated the phrase, resulting in an abnormally high error rate. We classified these trials as outliers and removed them. In all, 30 of the 937 phrases (3.2%) were eliminated by applying this criterion.

Speed, Accuracy, and KSPC
The results for speed are shown in Figure 6a. The first block rate of 6.6 wpm is close to the 7 wpm novice rate for Graffiti reported by Fleetwood et al. [3]. As expected, entry speed increased significantly with practice ($F_{1,9} = 40.7, p < .0001$). There was also a significant increase in speed by entry mode ($F_{3,27} = 58.1, p < .0001$). The average entry rate for block 8 (single stroke method with word completion) was 12.8 wpm, with rates for individual participants ranging from 8.0 to 18.2 wpm. The block $\times$ entry mode interaction was not significant, however ($F_{3,27} = 0.45, ns$).

Averaged by block, error rates were below 3% throughout the experiment. As seen in Figure 6b, error rates improved during blocks 5-8, dropping to just under 1% by the end of the experiment. The lower error rates during blocks 5-8 are likely due to a cognitive shift to “word-level attention” when using word completion. More accurate entry is also a likely by-product of the reduced KSPC when using word completion. Participants varied quite a bit in their disposition for accuracy. The main effect was not statistically significant for block ($F_{1,9} = 2.7, p > .05$) or entry mode ($F_{3,27} = 0.54, ns$).
Figure 6. Results for the single stroke method (blocks 1-4) and the single stroke method with word completion (blocks 5-8). (a) speed (b) error rate (c) keystrokes per character

Figure 6c clearly shows two trends for KSPC. During blocks 1-4, KSPC was relatively stable, with means of 1.417 in block 1 and 1.301 in block 4, representing a drop of 8.2%. While using language-based acceleration (blocks 5-8), KSPC was markedly less – and below 1.000, as conjectured. An improvement of 6.0% occurred with KSPC dropping from 0.793 in block 5 to 0.746 in block 8. These improvements were statistically significant, as demonstrate in the F-statistics for entry mode ($F_{1,9} = 857.3$, $p < .0001$) and block ($F_{3,27} = 9.40$, $p < .0005$). The block x entry mode interaction was not significant, however ($F_{3,27} = 1.44$, $p > .05$).

Errors followed by corrective actions tend to push the observed KSPC up, as do unrecognized strokes, or characters erroneously inserted and left in the transcribed text. During blocks 5-8, KSPC is also pushed up by participants failing to capitalize on opportunities, such as selecting a word when it first appears in the candidate list.

The most significant observation in Figure 6 is the failure of Unipad’s language-based acceleration features to yield a significant increase in text entry throughput – an increase consistent with the decrease in keystrokes afforded by word completion and the related acceleration features. The average observed KSPC for block 8 (0.746) is still (0.746 - 0.428) / 0.428 = 74.2% above the optimal value.

**Participant Questionnaire**

The post-test questionnaire (Table 1) solicited comments and responses to seven statements. The results are shown below. (Note: strongly disagree =1; disagree = 2; neutral = 3; agree = 4; strongly agree = 5).

<table>
<thead>
<tr>
<th>Question</th>
<th>Mean (sd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. I liked entering text using the plain single stroke method.</td>
<td>3.8 (0.79)</td>
</tr>
<tr>
<td>2. I liked entering text using the single stroke method with word completion.</td>
<td>3.2 (1.14)</td>
</tr>
<tr>
<td>3. Word completion is a useful addition to the single stroke method.</td>
<td>4.1 (0.88)</td>
</tr>
<tr>
<td>4. I entered text more quickly when word completion was added to the single stroke method.</td>
<td>3.9 (0.99)</td>
</tr>
<tr>
<td>5. I entered text more accurately when word completion was added to the single stroke method.</td>
<td>3.2 (1.23)</td>
</tr>
<tr>
<td>6. The following acceleration techniques were useful for word completion:</td>
<td></td>
</tr>
<tr>
<td>- Candidate list</td>
<td>4.1 (0.57)</td>
</tr>
<tr>
<td>- Frequent word prompt</td>
<td>4.2 (0.63)</td>
</tr>
<tr>
<td>- Suffix completion</td>
<td>3.8 (0.79)</td>
</tr>
<tr>
<td>7. Some fatigue was caused by the added attention required with word completion.</td>
<td>3.8 (0.79)</td>
</tr>
</tbody>
</table>

Eight participants gave additional comments. Two remarked on the similarity with Graffiti. Some indicated they needed more practice. One participant suggested using a dynamic word prompt and a richer suffix list.

**EXTENDED SESSIONS**

To observe the effects of continued practice and to estimate the expert potential of Unipad, three subjects agreed to perform an extra one-hour session wherein they repeatedly entered just one phrase:

the quick brown fox jumps over the lazy dog

The second author, representing a point further along the learning curve, served as a fourth subject. The extended session was organized as before, in four 15-minute blocks. The test phrase has 43 characters; however, only 27 strokes are required:

```
FCFCU S CFCU U CFCU CUC
```

All four categories of taps are seen above: “F”, for tapping a word in the frequent word prompt list, “C” for tapping a candidate word, “U” for tapping the word produced by the single stroke method alone, and “S” for tapping a suffix.
The “quick brown fox” phrase is popular for this purpose since it includes every letter of the English alphabet. In fact, this trait makes it atypical of English, as reflected in the KSPC statistic, which is $27 / 43 = 0.628$, or about 47% higher than the average for English (0.428). Thus, the results for the extended sessions should be conservative compared with typical English.

The participants (second author excepted) were not informed of the optimal stroking pattern; thus, an additional motivation was to investigate the input strategy adopted by participants.

In all, 483 instances of the test phrase were entered. Of these, 31 were entered optimally and without errors; that is, with 27 strokes and 0% errors. The remaining phrases demonstrated a variety of behaviours, including non-optimal use of the acceleration features, errors, etc. The results for the fourth block are summarized in Table 2.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Error Rate (%)</th>
<th>KSPC</th>
<th>Percent Above Optimal</th>
<th>Speed (wpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>5.70</td>
<td>0.6852</td>
<td>9.11</td>
<td>30.0</td>
</tr>
<tr>
<td>#2</td>
<td>4.08</td>
<td>0.8463</td>
<td>34.8</td>
<td>20.3</td>
</tr>
<tr>
<td>#3</td>
<td>0.00</td>
<td>0.8074</td>
<td>28.9</td>
<td>17.1</td>
</tr>
<tr>
<td>JC</td>
<td>4.15</td>
<td>0.7446</td>
<td>18.6</td>
<td>35.1</td>
</tr>
</tbody>
</table>

Table 2. Results for 4th Block of Extended Session.

The entry rates for the fourth block ranged from 17.1 to 35.1 wpm. The slowest participant (#3) was also the most meticulous, choosing to correct all errors. The other participants were faster, but also committed more errors.

Examples of the progression in speed over the four blocks are shown in Figure 7 for two participants. The test phrase was entered 123 times by participant #1 (top), 180 times by JC (bottom). A clear and typical learning trend is seen in the figure for participant #1. However, JC’s performance trend was relatively flat, which is expected given the prior practice. His entry rate exceeded 40 wpm for 34 phrases, with 47.9 wpm as the highest entry rate.

Figure 8 shows the distribution of strokes by category. Most values are consistent with the optimal stroke pattern shown earlier. For example, 7.3% of the strokes were taps on an entry in the frequent word prompt list. The optimal value is $2 / 27 = 7.4%$. Similarly, 2.9% and 2.8% of the strokes were for showing the suffix list and tapping on a suffix, respectively. The optimal value in both cases is $1 / 27 = 3.7%$. The observed values are bit less, indicating that suffix completion was not used as often as possible.

CONCLUSIONS

We have presented Unipad, a stylus-based text entry technique combining a single-stroke text entry method with word completion and other acceleration aides. A key feature is that Unipad superimposes pen strokes on the acceleration aides, thus minimizing movement of the stylus. Entry rates ranged from 8.1 to 18.8 wpm after two hours of practice in an experiment with ten participants. In follow-on sessions aimed at establishing the potential expert performance, four users entered “the quick brown fox” phrase repeatedly over four blocks of 15 minutes each. Average rates on the last block ranged from 17.1 to 35.1 wpm, with peak rates reaching 48 wpm. Although these rates are promising for mobile text entry, the attention demand of language-based acceleration aides, such as word completion, remain an issue as there is both a time cost and cognitive load for users to monitor and utilized the on-going predictive features. Nev-
Nevertheless, with practice users acquire the ability to anticipate and quickly respond to the predictive aides, thus allowing entry rates to edge upward. Longer term testing of individual text entry acceleration techniques may warrant future study. Testing Unipad on a smaller device such as a PDA would also be worthy of future work.

ACKNOWLEDGEMENTS

We thank Xerox Palo Alto Research Center for the Unistrokes text entry technique.

REFERENCES