Punctuation Input on Touchscreen Keyboards: Analyzing Frequency of Use and Costs

Sana Malik
Department of Computer Science | HCIL
University of Maryland
College Park, MD 20740
maliks@cs.umd.edu

Leah Findlater
College of Information Studies | HCIL
University of Maryland
College Park, MD 20740
leahkf@umd.edu

ABSTRACT
Non-alphanumeric symbols are rarely considered in text input research even though some punctuation is more frequent than the least common English letters. In this paper, we first evaluate punctuation frequency in two contrasting sources (Twitter and Google N-Grams). We then present a controlled study to compare existing techniques for ten-finger punctuation input on touchscreens: (1) the status quo, where punctuation symbols are stored on an alternate keyboard layer, and (2) an approach where users draw punctuation symbols atop the Qwerty keyboard itself [3]. Our findings show patterns in punctuation use (*e.g.*, comparing mobile and desktop input), and highlight the cost of mode-switching to enter punctuation marks.

Author Keywords
Touchscreen, text input, gestures

ACM Classification Keywords
H.5.2 [Information interfaces and presentation]: User interfaces - input devices & strategies.

INTRODUCTION
Despite the substantial research devoted to touchscreen text input in recent years, almost no attention has been paid to understanding and supporting use of non-alphanumeric symbols. Instead, advances have concentrated primarily on alphabetic input, such as incorporating language models to improve typing accuracy [5], calculating optimal keyboard layouts to reduce movement time [13], and proposing alternatives to standard tapping on Qwerty keyboards (*e.g.*, [7, 9, 11]). Non-alphanumeric symbols are rarely considered, even though some punctuation marks are more frequent than some English letters [13]. Underscoring this oversight, Mackenzie and Soukeroff's set of 500 phrases [10]—arguably the most popular phrase set for text input evaluations—contains no punctuation.

In this paper, we explore two aspects of non-alphanumeric symbol input on touchscreens. First, to assess how punctuation marks are being used at perhaps the extreme of punctuation-heavy touchscreen text input, we evaluate the frequency of punctuation symbols in a corpora of mobile Twitter data and compare that to desktop tweets and to the Google N-gram corpus [9]. Previous analyses of character frequencies were prior to the current mobile computing revolution (*e.g.*, [13]) and we were interested in how these frequencies may differ for a common mobile text input task.

Second, since punctuation symbols are typically difficult to access on touchscreen keyboards due to space constraints, we conducted a controlled lab study that allows us to precisely characterize the cost of entering these symbols with two very different input techniques. Both techniques were implemented on a keyboard large enough for ten-finger typing, such as those found on tablets: (1) a gestural technique where users draw punctuation symbols atop the keyboard itself [3] (Figure 1), and (2) a status quo alternate-keyboard technique where the user switches between alternate keyboard layouts (Figure 2). While a previous preliminary evaluation of these two techniques suggested that overall text input performance between the two was similar, no detailed analysis was conducted of the costs of switching between letter and punctuation input modes. Such an analysis could allow us to identify how to better support punctuation input in future touchscreen keyboard designs

RELATED WORK
Previous work on alternate (non-QWERTY) mobile text entry can be divided into two types: key-based and gesture-based input. Many key-based approaches involve rearranging keys from the QWERTY layout to optimize speed and accuracy. MacKenzie and Zhang design and evaluate a soft keyboard layout with a predicted 35% performance improvement over QWERTY [11].

Many gesture-based keyboards have been made to augment the QWERTY keyboard. Goldberg and Richardson [4] present a set of gestures for alphabetic text entry to facilitate "eyes-free" typing with a stylus. Cirrin [12] is an alphabetic keyboard with a circular layout and users insert words by dragging their fingers to connect letters. Bi et al. [2] present a similar gesture keyboard over the QWERTY layout and allow bimanual entry.

All of the mentioned keyboards focus on alphabetic text entry and do not include punctuation entry.

ANALYZING PUNCTUATION SYMBOL FREQUENCY
Previous work by Zhai *et al.* [13] compared punctuation between informal and formal settings (chat room logs and online news articles), and showed a higher usage of punctuation in informal writing. However, with the increased use of touchscreen mobile devices, we were interested, first, in how the device itself impacts text input compared to a traditional computer, and, second, how newer forms of communication (*e.g.*, social networks like Twitter and Facebook, SMS) affect punctuation use

Letters				Punctuation Marks			
Symbol	Twitter Mobile	Twitter Desktop	Google N-gram	Symbol	Twitter Mobile	Twitter Desktop	Google N-gram
e	9.34	9.52	11.58	.	1.694	1.748	1.151
a	9.15	9.25	7.52	@	1.221	1.258	0.000
o	7.09	7.36	7.07	!	0.940	0.813	0.013
t	7.04	6.82	8.57	'	0.550	0.446	0.200
i	6.52	6.44	7.08	_	0.527	0.499	0.001
n	6.15	6.02	6.74	,	0.401	0.532	0.000
s	5.19	5.26	6.15	:	0.381	0.344	0.087
h	4.60	4.51	4.71	#	0.377	0.350	0.000
l	4.38	4.35	3.82	?	0.338	0.362	0.032
r	4.24	4.37	5.86	"	0.205	0.110	2.284
m	3.18	3.15	2.38	-	0.185	0.193	0.217
d	3.12	3.14	3.55	)	0.181	0.228	0.140
u	3.10	3.17	2.55	<	0.095	0.100	0.001
y	2.74	2.64	1.55	>	0.094	0.106	0.002
g	2.60	2.41	1.75	(	0.089	0.087	0.140
c	2.02	2.09	3.13	*	0.075	0.072	0.008
k	2.00	2.00	0.52	&	0.055	0.044	0.005
w	1.95	1.86	1.55	;	0.048	0.051	0.096
b	1.85	1.75	1.40	/	0.042	0.046	0.019
p	1.64	1.72	2.00	^	0.017	0.023	0.003
f	1.42	1.48	2.23	=	0.016	0.025	0.002
v	0.80	0.87	0.99	~	0.013	0.020	0.001
j	0.57	0.54	0.16	$	0.010	0.012	0.005
z	0.27	0.28	0.09	|	0.007	0.007	0.001
x	0.27	0.29	0.22	\	0.005	0.003	0.001
q	0.09	0.15	0.11	+	0.005	0.006	0.001
				%	0.004	0.004	0.006
				]	0.002	0.001	0.010
				{	0.002	0.001	0.000
				}	0.002	0.003	0.000
				[	0.001	0.002	<0.001

Table 1: Letter and punctuation mark frequencies as percentage of total characters (minus spaces and commas). Each half of the table is ordered by frequency in the Twitter Mobile corpus.

compared to traditional English. To answer these questions, we compared three text corpora: (1) the Google N-gram corpus, which reflects traditional English spelling and grammar, and (2) mobile and (3) desktop tweets, which provide an informal, abbreviated style of text where some punctuation symbols play a unique role (*e.g.*, @, #).

Method

The Google N-gram corpus (Version 1) [9] contained 472,764,897 unigrams appearing in English books published between the years 1538–2008. We generated the Twitter corpora in June 2012 using the Twitter API, which uniformly samples 1% of the public tweet stream. Because the API's information about a user's language is not always reliable, we filtered tweets to those most likely to be in English by limiting allowed characters to ASCII 33–126. We then categorized touchscreen (mobile) versus desktop tweets using a manually created list of popular Twitter clients for Apple iOS and Google Android, which resulted in 57,662 desktop tweets and 173,876 mobile tweets. Finally, we removed URLs from the tweets, since these are often automatically generated or pasted in; had we kept the URLs, the gap seen below between the Google N-gram corpus and the Twitter data would have been bigger.

For each corpus, we counted the overall usage of English characters (letters and symbols, ASCII 33–126), and ranked the characters by frequency. Spaces made up 15.6% and 15.5% of the mobile and desktop Twitter data, respectively. Because spaces and commas were not present in the Google corpus, we excluded these two characters from the Twitter analysis (commas accounted for 0.4% and 0.3% of characters in the desktop and mobile corpora, respectively).

Results and Discussion

Punctuation and letter frequencies for the three corpora are shown in Table 1; numbers (0–9) made up an additional 1.1% in each of the Twitter corpora and 1.7% in the Google corpus. Overall, the Twitter corpora included substantially higher punctuation use than the Google corpus, comprising 7.5% of characters in the mobile tweets and 7.6% in desktop versus only 4.4% of characters in the Google corpus. With the Google corpus, only 6 punctuation symbols (. - ' ( ) ") appeared more frequently than the least frequent letter (*q*), whereas 11 and 14 symbols appeared more frequently than *q* in the desktop and mobile Twitter corpora, respectively.

That there is a difference between punctuation frequency in the traditional English versus Twitter corpora is not surprising. However, the magnitude of that difference—almost twice as much use with Twitter—emphasizes the need to consider punctuation when designing and evaluating new text input techniques.

Differences also emerged when comparing the mobile and desktop tweets, suggesting using a traditional keyboard versus a touchscreen keyboard influences how we type. Not only were more punctuation symbols frequent than the least common letter (q) with the mobile tweets, but the mobile tweets were also shorter than the desktop tweets (M = 49.9 *vs.* 54.9 characters).

EVALUATING METHODS FOR PUNCTUATION INPUT

In the second phase of this work, we conducted a detailed analysis of punctuation input on keyboards large enough for ten-finger typing (*e.g.*, tablets or tabletops), comparing two techniques: a standard alternate keyboard that allows users to toggle between multiple layouts to access punctuation, and a Qwerty keyboard augmented with gestural input, where users can draw the punctuation symbol atop the keyboard itself [3]. In a previous preliminary comparison of these techniques, six participants completed a short typing task (18 word pairs with punctuation) and no performance differences were found [3]. Here, we examine mode switching costs specific to punctuation input and we confirm the previous performance findings with a larger sample and longer task.

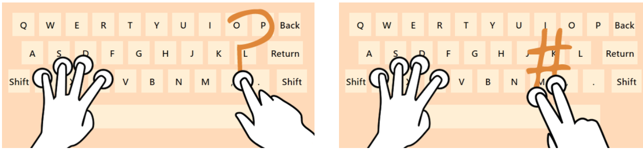

Figure 1. The gesture keyboard: Users enter gesture mode by placing four fingers on the keyboard. The keyboard changes color, and users draw the symbol they wish to enter. Some symbols (*e.g.*, #) allow single-touch and multi-touch gestures.

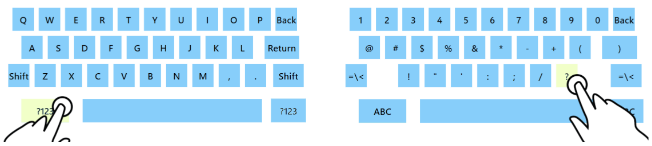

Figure 2: The alternate keyboard uses an alternate layer of keys for punctuation, based on Android keyboards.

The Alternate and Gestural Keyboard Interfaces

The keyboard interfaces were implemented for a Microsoft Surface 2 (now PixelSense) tabletop which has a vision-based 40" touchscreen. The *alternate keyboard* is an implementation of the status quo solution for punctuation input on touchscreens (Figure 2). It uses the same Qwerty layout as the gesture keyboard, but additionally has two keys labeled "?123" that toggle between the alphabetic keys and an alternate keyboard layer with punctuation marks and numbers. The alternate layer was modeled based on existing touchscreen keyboards; that is, the layout was largely based on an Android keyboard, with the exception that the comma and period keys were replaced by an "ABC" toggle on the right side of the keyboard. Comma and period were not evaluated in our study and commonly appear on the first layer of keyboards large enough for 10-finger typing (*e.g.*, Apple iPad).

The *gesture keyboard* is an implementation of the 10-finger keyboard described by Findlater *et al.* [3] (Figure 1). It is a standard Qwerty keyboard that allows users to draw punctuation symbols atop the keyboard using the process shown in Figure 1. Single-touch gestures are defined for each punctuation mark, while multi-touch options are also provided for $, " and #. For example, to enter the pound sign (#), a user can draw either four separate strokes, or use two fingers to draw a pair of lines at time (Figure 1c). The gesture set in the original work [3] required the user to draw a circles rather than just dots when entering a question mark and a percentage sign; based on piloting we extended the set to support dots as well.

Informed by the character analysis of the Twitter and Google corpora, ten punctuation marks were included for the study: ' - $ # " ? ! @ & %. The period and comma were not included because they commonly appear on the primary keyboard layer with the letter keys and thus should not incur a mode-switching cost. We also implemented but did not formally evaluate a *combined keyboard* that allows users to enter punctuation marks with gestures or by switching to the alternate layer.

Method

Ten participants, aged 21–30 (M = 24.8, SD = 2.9), were recruited through on-campus mailing lists. Two participants were female and two were left-handed (8 were right-handed). All participants had experience with touchscreen and four had used a gesture input keyboard (*e.g.*, Swype). Participants were compensated $15 for their time.

Participants completed two tasks with each keyboard: phrases and random words. For the phrases task, we adapted the MacKenzie and Soukoreff phrase set [10] by adding exactly two punctuation marks to 100 phrases, one near the middle of the phrase and one at the end (*e.g.*, physics & chemistry are hard!). The pairs of punctuation marks were chosen with a uniform distribution from the ten included in the study, except for the single and double quotations, which were always paired with themselves (*e.g.*, video camera with a good "zoom"). We inserted the punctuation marks manually and made them semantically meaningful when possible. The randomly generated words task, though not as realistic, allowed us to collect more punctuation input without inundating the user with all possible character bigrams. To do so, we created 5-character words from a uniform probability of each character being a letter or symbol (*e.g.*, %va$k).

The study employed a 2x2 within-subjects factorial design with factors of: *Keyboard* (alternate *vs.* gestural) and *Task* (phrases *vs.* randomly generated words). The orders of presentation for both Task and Keyboard were fully counterbalanced. For each keyboard, participants completed 20 training trials, which were repetitions of each of the 10 symbols, followed by 40 trials each for the phrases and randomly generated words tasks (*i.e.*, 40 phrases and 40 words). Between each keyboard condition, participants rated the keyboard on several scales and answered open-ended questions about their experience. Finally, at the end of the session, users briefly used the combined keyboard, typing up to 10 phrases. Study sessions lasted 60–90 minutes.

In total, participants entered 28,486 characters. To analyze overall typing speed, we ran a two-way repeated measures ANOVA (Keyboard x Task). For corrected and uncorrected error rates, which violate the normality assumption of parametric tests, we used Wilcoxon signed ranks tests to compare the two keyboards within each task. For each of the switching costs we analyzed (*e.g.*, from letter to punctuation mark), paired two-tailed t-tests were used to compare speed across the two keyboards.

Results

Overall speed and error rates for the keyboards are shown in Table 2. No statistically significant main or interaction effects of *Keyboard* were found for these measures, strengthening the previous preliminary findings [3]. However, only examining overall typing performance may mask differences specifically when it comes to punctuation input. To evaluate the cognitive cost of entering

	Phrase Task				5-character Word Task			
	Gesture		Alternate		Gesture		Alternate	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
WPM	23.0	3.4	23.9	4.6	9.7	2.3	10.9	1.9
Uncorrected Error (%)	0.4	0.2	0.8	1.5	0.4	0.4	0.5	0.7
Corrected Error (%)	10.5	4.2	9.1	4.0	6.6	2.4	2.8	2.7

Table 2. Performance results from gesture and alternate keyboards in the experiment.

punctuation marks, we thus considered four types of mode switches: punctuation to letter, letter to punctuation, punctuation to space, and space to punctuation. We considered letters and spaces separately since they should incur different costs with the alternate keyboard condition, where space appeared on both layers.

The average speed for switching into punctuation mode was similar for both conditions whether from a letter (alternate = 1862ms, gesture = 1789ms) or from a space (alternate = 1691ms, gesture = 1871ms). However, the gesture keyboard provided a significant advantage when switching from a punctuation mark to a letter, likely because the letter keys always remain visible and thus require less visual search: 913ms versus 1632ms on average ($t_9 = 6.15$, $p < .001$).

In comparison, the alternate keyboard was significantly more efficient than the gesture keyboard when entering a space after a punctuation mark: 556ms *vs.* 819ms ($t_9 = 3.62$, $p = .006$). Since the space bar is always displayed on both keyboards, the relatively poor performance of the gesture keyboard in this case suggests that there is a cognitive cost in transitioning from drawing a gesture to tapping again on the keyboard.

Overall, subjective feedback was mixed, with some comments reflecting the performance analysis of mode switching costs. Six participants preferred the gesture keyboard, while four preferred the alternate keyboard. One participant who preferred the gesture keyboard said that it "…is easier to use as you reduce the amount of click [tapping] to switch; I enjoyed hand written commands" and two others said they felt the gesture keyboard made typing "more enjoyable" and "like a game." Those who preferred the alternate keyboard liked that it was similar to what they were used to on other devices. One participant said that they found the alternate keyboard less confusing and commented on the cost of switching from punctuation marks to letters with the gesture keyboard: "Sometimes for gesture, I would try to draw in a letter."

At the end of the study session, participants briefly used the combined keyboard. Preliminary feedback on this version was positive, suggesting that participants recognized tradeoffs with the keyboards. Given the three options (alternate, gesture, or combined), six participants said they would prefer to use the combined keyboard, three preferred gesture only, and one preferred the alternate keyboard only.

DISCUSSION AND CONCLUSION

As touchscreen devices and new communication mediums like Twitter, Facebook and SMS increasingly influence how we communicate, we need to consider the implications for text input research. Here, we have focused on punctuation input, which has received limited attention in past research. Our findings motivate the need for increased emphasis on punctuation when designing and evaluating novel text input techniques. The Twitter corpora we analyzed included almost twice as much punctuation as the more traditional English in the Google N-gram corpus. The @ and # symbols were unsurprisingly common in the Twitter data since they play a specific role in tweets. However, even apart from @ and #, punctuation in general was more common in the Twitter data than the Google corpus.

Punctuation input on touchscreens is particularly problematic due to the cost of switching from punctuation to letter input. The comparison of two existing techniques for entering punctuation symbols—an alternate keyboard and a gesture-based keyboard—revealed significant costs of switching back to letters or spaces after entering a punctuation symbol. We used a text transcription task in this study, but it is possible that the penalty would be even greater in a text composition task (*e.g.*, writing an email or tweet) because the mode switch could take the user out of the flow of writing. Of course, more work is needed to test this hypothesis.

In terms of future work specific to the keyboard techniques evaluated here, it would be interesting to study use of the combined keyboard over a longer period of time to see how users adopt one or both of the input techniques and to evaluate how the cost of mode switching changes with more experience. We would also like to implement the gestural input technique on a capacitive tablet large enough to support 10-finger typing; while the slow typing speeds observed in this study are partly due to the difficulty of typing punctuation compared to regular text, the Microsoft Surface's vision system also often skipped keystrokes when the user typed quickly.

Many questions also remain about how keyboard design itself may impact language use. We observed some differences between the mobile and desktop tweets, but, for example, we were not able to assess whether typing on different implementations of touchscreen keyboards—say, a standard iPhone keyboard versus Swype—impacts the content of the text and frequency of symbol usage. Additionally, the keyboard was implemented on a touchscreen tabletop. While this is practical for a ten-finger keyboard, they are currently not as widely used as touchscreen tablets and handheld devices.

ACKNOWLEDGMENTS

We thank Rajan Zachariah for insight and feedback on this research. Funding was provided by Google.

REFERENCES

1. Anthony, L., and Wobbrock, J. O. A lightweight multistroke recognizer for user interface prototypes. Proc. Graphics Interface (GI'10) (2010), 245–252.
2. Bi, X., Chelba, C., Ouyang, T., Partridge, K., and Zhai, S. Bimanual gesture keyboard. Proc. UIST'12 (2012), 137–146.
3. Findlater, L., Lee, B., and Wobbrock, J. O. Beyond QWERTY: augmenting touch-screen keyboards with multi-touch gestures for non-alphanumeric input. Proc. CHI'12 (2012), 2679–2682.
4. Goldberg, D., and Richardson, C. Touching-typing with a stylus. Proc. INTERCHI'13 (1993), 80–87.
5. Goodman, J., Venolia, G., Steury, K., and Parker, C. Language modeling for soft keyboards. Proc. IUI'02 (2002), 194–195.
6. Hart, S., and Staveland, L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Human Mental Workload 1 (1988), 139–183.
7. Kristensson, P. O., and Zhai, S. Shark2: a large vocabulary shorthand writing system for pen-based computers. Proc. UIST'04 (2004), 43–52.
8. Li, F. C. Y., Guy, R. T., Yatani, K., and Truong, K. N. The 1line keyboard: a qwerty layout in a single line. Proc. UIST'11 (2011), 461–470.
9. Lin, Y., Michel, J. B., Lieberman, E., Orwant, J., Brockman, W., and Petrov, S. Syntactic annotations for the Google books ngram corpus. System Demonstrations at ACL'12 (2012), 169–174.
10. MacKenzie, I. S., and Soukeroff, R. W. Phrase sets for evaluating text entry techniques. Extended Abstracts of CHI'03 (2003), 754–755.
11. MacKenzie, I.S. and Zhang, S.X. The design and evaluation of a high-performance soft keyboard. Proc. CHI'99 (1999), 25–31.
12. Mankoff, J. and Abowd, G.D. Cirrin: a word-level unistroke keyboard for pen input. Technical Note at UIST'98 (1998), 213–214.
13. Zhai, S., Hunter, M., and Smith, B. A. Performance optimization of virtual keyboards. Human-Computer Interaction 17 (2002), 89–129.