A Survey of Alternate Text-Entry Methods

At Eatoni, we are often asked how our lingusitically optimal text-entry methods compare to other approaches to the problem of text entry on mobile devices. Here are some brief answers.

Chording methods

Pressing several keys at once. A consistent commercial failure: too hard to learn.

Chording methods are exemplified by the Microwriter shown here, or the "Twiddler" handheld device, familiar in ubiquitous computing circles. These devices achieve economy of size and number of keys by using combinations of several keys pressed simultaneously--chords--to encode letters. In a typical system three or more keys must be pressed in some arbitrary combination in order to type each letter. Just as piano playing skill is given to but a few, so is the market for such devices limited. The device shown, The Microwriter personal digital assistant made by Agenda B, now defunct, gained a small but loyal following in England in the early 1990's. The black keys are pressed in combinations with the fingers and thumb of the right hand to form chords which code letters.

Miniature Keyboards

Since you can't touch type on them, there's little advantage to a Qwerty layout.

There have been many recent attempts to market devices of this type. They consist of a Qwerty keyboard reduced in size so that it can fit into a pocket, or almost. The keyboard may be integrated into another device, as in the Nokia 9110 Communicator, connectable to a mobile phone or other device, or completely self sufficient, such as the devices required by the Pocketmail service. These latter devices connect to a telephone via an acoustic coupler in order to send email composed on the keyboard. The problem is that when the Qwerty keyboard is small enough to fit into a pocket, it is no longer big enough to touch type on, and when it is big enough to touch type on, it is bulky. There are other related solutions, such as foldable keyboards, but all of these are too awkward to have widespread commercial appeal. One of the most popular devices (pictured) is manufactured by Research in Motion.

Handwriting Recognition

For many, handwriting recognition is a massive step backward.

The personal digital assistant (PDA), as exemplified by 3Com's Palm Pilot, will eventual disappear as a class of handheld machine, as its functionality is absorbed into a converged computing/communication device. PDAs became a fetish item thanks to the Pilot's handwriting recognition hardware and software. Qualcomm, with its pdQ phone which incorporates a full PalmOS implementation, is betting that handwriting recognition will survive convergence.

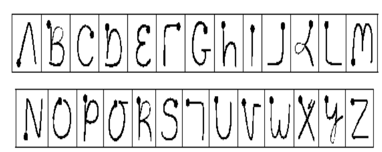
The Palm Pilot is judged by many as "pretty good". Already some 6 million people have learned how to operate the "Graffiti" text-entry method.

Writer's cramp for the new millennium.

.Graffiti is as difficult to learn as any mass market text-entry method can be. Many find that Graffiti is in fact not very easy to learn. Numerous reviews of the Pilot, while generally extolling its virtues, conclude with an expression of desire for a keyboard.

The main difficulty is that a certain level of competence must be reached, a new way of writing many letters must be memorized, before most text can be entered. In this sense, Eatoni's system is much easier to learn than Graffiti. Nothing needs to be memorized for an Eatoni-equipped device to be used for text entry. Labels on the keyboard guide the user to the correct manipulations to enter text. As the manipulations are learned, text entry becomes faster. Indeed, one can anticipate a much higher terminal text entry rate with Eatoni than with Graffiti. There is no reason that a user could not type as fast with Eatoni as with a regular Qwerty keyboard, where speeds of 40 words per minute are routinely achieved, and rates over 100 words per minute have been recorded. Writing script, however, is rarely possible at speeds greater than 20 words per minute.

Ease-of-use Considerations for Graffiti vs. WordWise

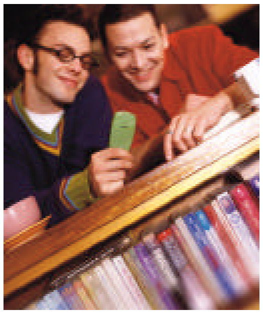


Ease-of-use has many aspects, familiarity, physical comfort, difficulty of learning, and so on. One of the most important from the standpoint of early adoption of a new technology is difficulty of learning. Graffiti is simple enough that millions have learned it, yet it is complicated enough that millions more have resisted learning it. That there is a market at all for keyboards into which a palm pilot can be placed indicates that Graffiti is not the ultimate answer for handheld text entry. In the following tables, we compare the difficulty of learning of Graffiti and WordWise by listing some of the facts that a learner must assimilate to learn each system. We are assuming here that the consumer already knows how to write using a pen, and type using a Qwerty keyboard. For both Graffiti and WordWise, there is a base case to be learned which covers most of the letters. For Graffiti, the base case is: print capital letters, make sure they are written large, filling most of the text-entry portion of the touch screen. For WordWise, the base case is: press the key labeled with the letter. From there, certain exceptions must be learned for certain letters. The more exceptions, the more difficult the system is to learn. And, more importantly, the more different the exceptions are from each other, the more difficult the system is to learn. Consider the difficulty of learning a foreign language for which the grammar contains many exceptions, as compare to leaning one which is regular. In the following tables, the letters with exceptions are listed, along with a description of the exception. In the case of Graffiti, the table is far from exhaustive, in the case of WordWise, it is exhaustive.

WordWise Base case: type the letter	
c	**hold down the shift**
e	**""**
h	**""**
l	**""**
n	**""**
s	**""**
t	**""**
y	**""**

Graffiti Base case: print capital letter; write big!	
a	**leave out horizontal line**
e	**don’t square off**
f	**square off, and leave out the bar**
g	**don’t forget the little bar to the right**
h	**lower case**
i	**no dot**
k	**special**
m	**starts at the bottom**
q	**special**
t	**special, confuse with f**
v	**special, confuse with j**
y	**is cursive, don’t raise up until the end, else get v**

Voice To Text Systems



The last alternate technology to be considered is voice to text, that is, systems which can translate spoken language into written text.

For now, and for some time to come, it's not an issue.

Voice-to-text systems has been "around the corner" for some time, in the same way that artificial intelligence was around the corner for decades. While they are far from perfect, voice-to-text systems do exist today for server-class machines. Moore's law states that computing power doubles every 18 months. Even supposing that Moore's law will hold indefinitely, it will be many years until a mobile phone has the computing power of today's PC server. Consider, for instance, memory. A typical mobile phone has 2 megabytes of memory, while a typical server might have 128 megabytes of memory. Moore's law predicts that mobile phones will have the memory of today's server in 9 years.

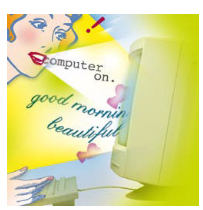
Eatoni believes that even when voice-to-text systems arrive on mobile phones, many or most consumers will still need rapid **manual** text entry systems. We take an even stronger stance: Even perfect voice recognition would be an inferior to WordWise as a solution to the mobile text entry problem. Some of our reasons for this belief are as follows:

Typing provides privacy, silence, and a parallel channel.

Silence is Golden. We can already see the rising tide of a mobile phone backlash. The backlash against mobile phones will follow the course of the anti-smoking campaign: the public will not kindly suffer noxious second-hand noise. Voice-to-text promises to aggravate the problem, and textual conversations are the solution.

Fish-bowls are for fish. By the same token, even if others are willing to overhear your mobile phone conversation, you may not wish them to know the details of your personal or business life. Again, voice to text is part of the problem, not part of the solution. In short, speaking on a mobile phone makes private space public, while textual conversation reclaims privacy in any public space.

Come again? Between our ears lies an extremely sophisticated voice-recognition system. Still, we are often led to ask, "what did you say?" or "can you spell that please?" Voice, like dictionary-based predictive text systems, is not very good at dealing with noise, with uncertainty, with new words, or names of any kind beyond the most familiar.



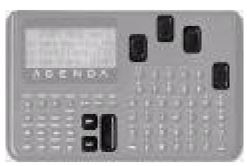
The second law of communication: The number of available modalities increases with time. Typed text will provide a parallel channel to voice recognition systems when they arrive. Consider giving voice commands to your handheld device while you are using it to have a conversation. How do you inform the machine that you are talking to it rather than your human interlocutor, or vice versa? In general, how do you keep several information streams open and separate using only voice recognition?.

Do you really want to have to dictate all your correspondence? Not everyone is given the skill of dictation, and even those skilled in dictation may prefer the process of writing by typing. Writing a text is typically not the process of registering words in a linear stream, rather it is a highly interactive process between the writer and the written, in which the writer constantly revises the written in non-linear way. Physical typing supports this process much better than speech recognition.

Don't believe the hype. It's much faster to type. An edited report from Dr. Robert Baily, a human factors researcher, summarizing recent work on automatic speech recognition:

> Automatic speech recognition technology has been under development for over 25 years, but has not yet received widespread use. One of the main reasons that speech recognition has not gained greater acceptance is that speech recognition errors are fundamentally different than keying errors. Most keying errors can be tracked back to users, while most speech errors are tracked back to mis-recognition of the speech by the computer. In the latter case, user input simply does not match computer output.[Eatoni note: This is why lookup errors are very bad, they are computer error, not human error.]
>
> Even though people can dictate faster than they can type, actual throughput is usually much slower with automatic speech recognition systems than with keying. A major problem is that error correction takes much longer with speech. The most commonly used correction methods used with speech input are:
>
> (a) deleting and repeating the last phrase,
> (b) deleting and repeating a specific word,
> (c) deleting and selecting a correct word from a list of alternative words,
> (d) typing the correction.
>
> Past studies have suggested that switching modality [using typing and speech together] could speed up interactive correction of recognition errors. Suhm, Myers and Waibel (1999) at Carnegie Mellon University found that switching between modalities eliminated repeated recognition errors. They found that if users simply repeated their speech to correct errors, correction accuracy was much lower than if users switched to a different modality (keyboard and mouse).
>
> [...]
>
> Throughput is the number of correct words produced per minute.
> The key variables are:
> (a) the accuracy of the speech recognition system,
> (b) the speaking rate of the user, and
> (c) the time required to correct errors.
>
> [In a study,] **the fastest users** spoke at an average of 107 uncorrected words per minute, **which resulted in about 25 corrected words per minute**. The "keyboard-mouse" group completed almost three times more words per minute than did the "voice-only" group.
>
> Lewis (1999) at IBM evaluated the performance of participants using a speech recognition dictation system. The participants received training in one of two correction strategies, either "voice-only" or "voice, keyboard and mouse." In both cases, users spoke at about 105 uncorrected words per minute. The multimodal (voice, keyboard, mouse) corrections were made three times faster than "voice-only" corrections, and generated 63% more throughput. Participants observed that they were usually aware of when a typing error occurred, but were much less confident of being aware of when a speech error occurred. Users must either constantly glance at the display for errors, or rely heavily on proofreading after the speaking has ended.

Summary of Alternate Approaches

Mechanism	Examples	Advantages	Disadvantages
Qwerty keyboard integrated in a single handheld device.	Nokia 9110. 	Could be touch typed if the keys were large enough. The Qwerty keyboard is well-known. Smooth learning curve: a typist with no training can perform useful work, speed increases with use.	The Qwerty keyboard has too many keys to fit on a handheld device without excessive bulk. The keys are too small for touch typing.
“Outboard” Qwerty keyboard which plugs into a mobile phone, or makes a connection to a fixed phone.	The Pocket Board 	Somewhat larger keyboard, so easier to type.	Requires the purchase, and more importantly, the transport of two devices, one of them too large to fit into a pocket. Keys are typically still too small for touch typing.
Chording Keyboards in which their are fewer keys than letters. Letters are selected by operating several of the keys simultaneously.	Microwriter. 	Unambiguous. Fits on a handheld device.	Abrupt learning curve in the sense that a user must receive a significant amount of training before performing useful work. Requires special hardware and software.
Handwriting Recognition.	Graffiti, Jot. 	Fits on a handheld device.	Abrupt learning curve. Maximal speed is limited, since handwriting is slow. Requires special hardware and software. Requires use of a stylus. High error rate.
Voice to Text.	None commercially available for a handheld device.	Needs no keyboard.	No privacy. Dictation rather than writing. Even if perfected, would need constant error correction. Slow.

For Further information concerning Eatoni Ergonomics, Inc., see: www.eatoni.com.