Expressive Keyboards: Enriching Gesture-Typing on Mobile Devices

Jessalyn Alvina[1,2] **Joseph Malloch**[1,2] **Wendy E. Mackay**[1,2]

[1]Inria, Université Paris-Saclay F-91405 Orsay, France

[2]LRI, Université Paris-Sud, CNRS, Université Paris-Saclay F-91405 Orsay, France

{alvina, malloch, mackay}@lri.fr

ABSTRACT
Gesture-typing is an efficient, easy-to-learn, and error-tolerant technique for entering text on software keyboards. Our goal is to "recycle" users' otherwise-unused gesture variation to create rich output under the users' control, without sacrificing accuracy. Experiment 1 reveals a high level of existing gesture variation, even for accurate text, and shows that users can consciously vary their gestures under different conditions. We designed an *Expressive Keyboard* for a smart phone which maps input gesture features identified in Experiment 1 to a continuous output parameter space, i.e. RGB color. Experiment 2 shows that users can consciously modify their gestures, while retaining accuracy, to generate specific colors as they gesture-type. Users are more successful when they focus on output characteristics (such as red) rather than input characteristics (such as curviness). We designed an app with a dynamic font engine that continuously interpolates between several typefaces, as well as controlling weight and random variation. Experiment 3 shows that, in the context of a more ecologically-valid conversation task, users enjoy generating multiple forms of rich output. We conclude with suggestions for how the *Expressive Keyboard* approach can enhance a wide variety of gesture recognition applications.

Author Keywords
Continuous Interaction; Expressive Communication; Gesture Input; Gesture Keyboard; Mobile; Text Input.

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

INTRODUCTION
People have been writing for thousands of years, using a wide variety of techniques, including cuneiform on clay tablets, carved runes, heiroglyphics, Chinese caligraphy, and illuminated manuscripts. The development of moveable-type printing presses brought a measure of standardization to text, since each letter was no longer directly produced by a person. Once digital computers arrived, this uniformity became perfect–digital symbolic values for each letter are defined according to specific schemes, e.g. ASCII and Unicode, removing the need for reinterpreting the possibly ambiguous visual appearance of inked, carved, or otherwise rendered text.

This valuable reduction in ambiguity resulted in a corresponding reduction in personalization that had been present in earlier writing systems. Letters are recorded perfectly with stylistic information stored separately, and applied to large, heavily quantized blocks of text. For example, the subtle emphasis encoded implicitly in a continuously varying pen stroke is now simply rendered as a standard *italic* typeface. Along with a severe reduction in the granularity of control, this approach also discards potentially valuable channels for implicit communication of personal style, mood or emotional state, and temporal or situational contexts. Users can of course edit the font, typeface size and color of the rendered text, but this is necessarily separate from the actual text input.

Computer keyboards are usually constructed as an array of labeled momentary switches (buttons), but interestingly, most mobile devices capture text input via "soft" keyboards displayed on high-resolution 2D touchscreens. Thus, although the output is symbolic, the input is highly oversampled in both space and time, giving us the opportunity to explore more continuous forms of control. For example, dynamic key-target resizing based on models of likely words or letter sequences increases the apparent accuracy of soft keyboards and partially resolves the "fat-finger" problem [13].

Gesture-typing [24] is a more interesting alternative that offers an efficient, easy-to-learn, and error-tolerant approach for producing typed text. Instead of tapping keys, users draw the shape of each word, beginning with the first letter and continuing through the remaining letters. Typically, a recognition engine compares each word gesture to a pre-designed "template" representing the ideal word shape. Word-gestures are not unique for each word, but can be robustly matched using a combination of kinematic models, multidimensional distance metrics, and language models to resolve ambiguities. Gestures that vary significantly may thus still register as correct.

As with other soft keyboards, the goal of gesture-typing keyboards is to produce the single, "correct" typed word intended by the user; it is either correct or incorrect, and input variation is of interest only for the purpose of designing tolerant recognition systems. Gesture variation is treated essentially as a deformation of the correct shape and discarded

Jessalyn Alvina, Joseph Malloch, Wendy E. Mackay. Expressive Keyboard: Enriching Gesture-Typing on Mobile Devices. In *Proceedings of the 29th ACM Symposium on User Interface Software and Technology (UIST 2016)*, ACM, October 2016. pp 583–593.

ISBN 978-1-4503-4189-9/16/10.
http://dx.doi.org/10.1145/2984511.2984560

as unwanted noise. To be sure, a small part of the variation is motor-system or digitizer noise, and cannot be considered meaningful. However human experience with handwriting clearly shows the potential for personal and stylistically-communicative variation of output media through performed human gestures.

What if we could leverage at least part of the natural variation in gesture-typing to increase the richness and nuance of text-based communication channels? Mobile devices already include high-resolution sensors capable of measuring the variation, and commercialized gesture typing systems are widely installed and are already designed to tolerate deformations of the "ideal" gesture template. Capturing continuous features of the variation and mapping it to properties of the rendered text could re-enable some of the benefits of handwriting, such as recognizable personal styles, implicit communication of mood, activity, or context; and explicit communication of emphasis, sarcasm, humor, and excitement.

Expressive Keyboards

We introduce *Expressive Keyboards*, an approach that takes advantage of rich variation in gesture-typed input to produce expressive output. Our goal is to increase information transfer in textual communication with an instrument that enables users to express themselves through personal style and through intentional control. This approach adds a layer of gesture analysis, separate from the recognition process, that quantifies the differences between the gesture template and the gesture actually drawn on the keyboard. These features can then be mapped to output properties and rendered as rich output. Before we can build an *Expressive Keyboard*, we must first address four research questions:

1. Does gesture-typing performance actually vary substantially across users (due to biomechanics or personality), or context (activity or environment)?
2. Can this variation be quantified as detectible features?
3. Can users *deliberately* control these additional features of their gestures *while gesture typing real text*?
4. How do users appropriate *Expressive Keyboards* in a more realistic setting?

This paper presents related work, and then attempts to answer the above research questions through a series of experiments and software prototypes. Experiment 1 is designed to verify whether gesture-typing performance varies across participants and experimental conditions. We report the results and how they led to the selection of three features that form a low-dimensional representation of gesture variation. Experiment 2 is designed to test whether or not users can *deliberately* vary both the selected features and the parameters of the rendered output text using a simplified control mapping, while simultaneously typing the required text. We report on the results and how they influenced the design of a second prototype, which maps users' gestures to a dynamic font. Experiment 3 is designed to collect ecologically valid in-the-wild data consisting of real-world conversations between pairs of friends. We report the result of this study, as well as users' perceptions of the *dynamic font*. We conclude with directions for future research, including additional mappings and applications.

RELATED WORK

Much of the research on digital writing uses machine learning to improve content recognition, e.g., by predicting the most likely word from the context (auto-completion) [12] or by improving spelling or grammar (auto-correction) [10]. These systems seek to predict the user's intention, at some level of probability, to produce the "correct" outcome.

In each case, the output is fixed: typing on both hard and soft keyboards produces standard output that lacks non-verbal cues [23]. Users sometimes use **bold** or *italic* typefaces, or ALL CAPS to emphasize a block of text. To convey more subtle expression, users may also insert emoticons, either by selecting them from a menu; typing a particular keyword, e.g., 'sad' to produce ☹, drawing a gesture [21], or through an emoticon recommendation system [22]. However, the act of selecting an emoticon is not integral to the production of the text and can easily distract the user from the act of writing [1, 21]. The degree of expression is also limited to the pre-defined set of emoticons.

Enhancing Text-based Communication

Some researchers have explored how to support subtle expression in text-based communication. For example, EmoteMail [1] annotates email paragraphs with the sender's composition time and facial expression. KeyStrokes [19] uses shapes and colors to visualize typing style and text content. Iwasaki et al. [15] added sensors to a physical keyboard to capture typing pressure and speed. Mobile devices offer new possibilities for generating rich text, given their touch screens and multiple sensors capable of capturing temporal, spatial and contextual features. For example, Azenkot & Zhai [3] investigated how users type on soft keyboards and found that touch offsets vary according to how they hold the device. Buschek et al. [6] combined touch offset, key-hold time, and device orientation to dynamically personalize the font.

Gesture as an Expressive Instrument

A third alternative is to use gestures. Researchers who study gesture for music or dance often take a completely different perspective, emphasizing the continuous qualities of human gestures over recognition: individual variation is valued rather than ignored or rejected. These researchers characterize gesture variation in terms of qualities of movement: spatial features [6, 7]); temporal features; continuity; power; pressure; activation; and repetitions [9].

This approach to studying and using gesture contrasts with definitions of the term from linguistics and cognitive psychology. See McNeill [18] for a more in-depth discussion of the competing conceptual understandings of the term 'gesture'.

In the artistic domain, the richness of gesture can be transformed into continuous output, e.g., [11], or to invoke a command [16] in a more integrated interaction. If the goal is to make the system 'fun' and challenging, the system should encourage curiosity [19]. Hunt et al. [14] found that continuous, multi-parametric mappings encourage people to interpret and explore gestures, although learning these mappings takes time. Human gesture variation can also be affected by movement cost [20], interaction metaphors and system behavior.

QUANTIFYING VARIATION IN GESTURE-TYPING

We conducted a within-participants experiment with three types of INSTRUCTION as the primary factor: Participants gesture type specified words “as *accurately* as possible”; “as *quickly* as possible while still being accurate”; and “as *creatively* as possible, have fun!” The *accurately* condition should provide the minimum level of variability for novice gesture-typists as they try to match the word shape as closely to the template as possible. The *quickly* condition might realistically be found in real-life gesture-typing under time constraints, and presumably results in greater variability and divergence from the template. The *creatively* condition was designed to provoke more extreme variation, and is not intended to match a real-world gesture-typing scenario.

We chose three sets of 12 words that vary systematically according to three dimensions: length (SHORT <4 characters, LONG >4 characters); angle (ZERO, ACUTE, or OBTUSE); and letter repetition (SINGLE, e.g., *lose*, or DOUBLE, e.g., *loose*). We consider angle between stroke segments because it may affect performance [20]. For example, the word *puree* is long, with a double letter ‘e’, and a zero drawing angle, i.e. a straight line on the keyboard; *taxi* is short, with a single letter and at least one obtuse angle: the chunk *axi*. Each letter appears at least once in each set.

Participants

We recruited seven men and five women, all right-handed, mean age 26. All use mobile phones daily, but none had used gesture-typing prior to this experiment.

Apparatus

We developed a custom Android application running on an LG Nexus 5 (Android 5.1) smartphone. It displays a non-interactive Wizard-of-Oz (WOZ) keyboard that matches the position and dimensions of a standard QWERTY English keyboard. We use the keyboard evaluation technique described in [4]; the WOZ keyboard collects gesture coordinates that are later fed to a word-gesture recognizer (keyboard).

Procedure

Sessions last approximately 50 minutes. Participants sit in a chair and hold the phone comfortably in their left hands, so they can perform all gestures with their right index finger. Participants are encouraged to talk aloud as they draw each word. During initial training, participants may practice until they feel comfortable using the gesture-typing technique.

Each trial in the experiment begins with an instruction displayed at the top of the screen, e.g., “Draw as accurately as possible”, with a word centered below, e.g., *queue*, and a soft keyboard at the bottom of the screen (see Fig. 1a). The trial ends when participants lift their finger, after which they answer a multiple-choice question as to their level of confidence: “Do you think you wrote *vein*?” (*Yes*, *No*, or *Not sure*). Each word is presented as a sub-block with 10 replications.

The experiment consists of 360 trials (12 words x 3 instructions x 10 replications). All participants begin with the *accurately* instruction; *quickly* and *creatively* are counterbalanced for order across participants. The 12 words are chosen from

Figure 1. Gesture variations: a) *Accurately* is straight, b) *quickly* is smooth, and c) *creatively* is inflated and highly varied.

the three word sets; counter-balanced within and across participants.

Data Collection

We record the touch coordinates in order to extract spatial and temporal characteristics of each gesture. We later simulate the gesture data on gesture-typing recognizers, KB-1 and KB-2, to derive ACCURACY, i.e. the recognizer score for the intended word (*True=1*, *False*=0). We also record the participant’s CONFIDENCERATE – an ordinal measure of the post-trial answers (*Yes=1*, *Not Sure*=0.5, *No=0*). The post-questionnaire asks participants to describe how they varied their gestures according to each instruction. We also record a kinematic log of each gesture, using screen capture, and audio record the participant’s verbal comments.

Results and Discussion

The first research question concerns the extent to which the participant’s gestures vary as they gesture-type. We first examined the subjective measures obtained through the post-questionnaire and looked at the existing variability in gesture data to identify candidates for gesture features.

We collected 4320 unique gestures. We removed 22 outliers (0.5%), defined as when 1) a participant said they made a mistake, e.g. accidentally lifting the finger before finishing the gesture; 2) they answered *no* to the post-trial question; and 3) gesture length was <100 pixels. Significance rates for confidence or recognition rate were not affected.

Figure 2. Recognized creative gestures included: a) loop and cusp for *taxi*, b) visualization of crown for *queen* and c) scribbling on keys to create the stars or a constellation for *midnight*.

Gesture Variability

Like [24], we found that participants viewed the word-gesture as crossing through “targets” i.e. each letter in a word. Participants changed the way they drew depending upon their perception of the instructions. Seventy-five percent (9/12) of the participants said they “pass through all the letters” and “stay in the [letter] box” in the *accurately* condition. This results in straight-line gestures (Figure 1a) that closely resemble the gesture template. Not surprisingly, participants drew faster when asked to draw *quickly*, which resulted in smoother, more curvy gestures (Figure 1b), although two participants mentioned that they tried to draw straight lines so it was shorter and took less time. Almost half (5/12) said they explicitly ignored precision when they drew *quickly*. Participants interpreted the *creatively* instruction as either to "draw shapes along the way from one key to another" (7/12), or to "be comfortable, likeable, or suitable" (5/12). The *creatively* instruction resulted in the highest variation across gestures, as shown in Figure 2.

Accuracy

Although recognition is not our primary goal, we are interested in how the different recognizers reacted to the variation in the users' gestures. We replay the recorded gesture data to the Android MonkeyRunner event simulation tool on a desktop computer, which communicates with the phone using the Android Debug Bridge. We use two recognizers (referred to anonymously as KB-1 and KB-2) representing the state-of-the-art for Android, with over 150 million copies installed collectively. Both keyboards have identical dimensions but different recognition algorithms: one is a commercial version of SHARK2 [24]; we do not know the other approach [1].

We send the recognizer a random word between each prompt word and erase personalization data after each participant. This ensures that the recognition results cannot be contaminated by the adaptation and personalization features in the keyboards. We expect low accuracy since both recognizers are known to use language context to resolve ambiguities between word shapes in normal use, and participants did not receive accuracy feedback during the experimental trials. The instruction *creatively* was specifically intended to provoke wide exploration and was not expected to give recognizable results – in fact, high recognition rates would indicate a failure to provoke adequate exploration.

ACCURACY for KB-2 (73.6% of the gestures recognized correctly) was significantly higher than for KB-1 (46.7%). One contributing factor is that KB-1 treats leaving the keyboard space as cancellation. However, even when these trials are removed, KB-2 achieved a significantly higher recognition rate (75% vs. 53%).

INSTRUCTION also has a significant effect on ACCURACY for both keyboards ($F_{2,22}$=140.6 and $F_{2,22}$=106.3 for KB-1 and KB-2 respectively, all p<.0001). A post-hoc analysis with Tukey HSD showed that accuracy is significantly lower for *creatively* (mean KB-1=34% and KB-2=62%), but no significant differences between *accurately* (KB-1=53% and KB-2=82%) and *quickly* (KB-1=57% and KB-2=79%). Surprisingly, for KB-1, the *quickly* instruction resulted in fewer errors than instructions executed *accurately*.

Confidence Rate

Participants expressed high overall confidence in their performance (83.6% *Yes*; 9.9% *Not Sure*; 6.4% *No*). Participants were most confident when they tried to write *accurately* (87.5% *Yes*; 7.1% *Not sure*; 5.4% *No*), and least confident when they wrote *quickly* (81.2% *Yes*; 10.2% *Not Sure*; 8.5% *No*). For *creatively* (82.1% *Yes*), participants are less sure whether or not they wrote the intended word: 12.4% *Not Sure* and 5.5% *No*. Their comments indicate that they drew more carefully in the *accurately* condition, which appears to have increased their confidence.

In summary, Experiment 1 shows that novice users modify their gestures in response to different instructions, while maintaining high confidence in their performance. Eight participants were eager to pursue more expressive gestures when writing, especially to generate emoticons or when communicating with friends and family.

GESTURE FEATURES

The next step is to determine if the variation identified in experiment one can be quantified as detectible features which can then be mapped to text output variation. We considered candidates from Experiment 1 as well as previously explored gesture characteristics from the literature [7, 9]. We evaluated them with respect to their applicability to gesture-typing: Word-gestures are restricted by keyboard layout properties, such as letter size and position, making features such as orientation, direction, or scale difficult to control while gesture-typing. We performed an ANOVA[2] to determine the effect of INSTRUCTION (user intention) on each feature.

Speed

Speed is a common measure of gesture variation, e.g., for typing activity [19, 8] and modeling the production time of a gesture [5, 7]. For gestures, speed is calculated by dividing the total length of traced distance (in pixels) by the total time (in milliseconds). We found that average speed is significantly affected by INSTRUCTION ($F_{2,22}$=216.8, p<.0001), where *quickly* is fastest (mean=1.8 px/ms), followed by *creatively* (mean=1.3) and *accurately* (mean=1.07). This suggests that the participants can vary the drawing speed according to different instructions.

Since our goal is to enable more fine-grained gesture control, we would like to increase granularity by examing each gesture over chunks of movement instead of as a whole. Given a sequence of points $P = \langle (x, y) \rangle$, we can divide P into n chunks, where P_i is the i^{th} sub-sequence. Thus, we calculate the average speed (v_i) of each chunk as follows:

$$v_i = \frac{\sum_{j=0}^{n} \sqrt{(x_j - x_{j+1})^2 + (y_j - y_{j+1})^2}}{T_i} \qquad (1)$$

where T_i is the total time for P_i. Parameterizing n increases the possibility of explorations in the feature space, which may give similar information to acceleration but with less noise. We started with $n = 2$ and compared the speed of the first and last half of each gesture to measure *speed consistency*:

$$R_{i,j} = \frac{v_j}{v_i} \qquad (2)$$

[1]We anonymize the keyboards here since the goal of our investigation probably differs from that of the keyboard designers and the samples we collected may not match typical uses of the keyboards.

[2]All analyses are performed with SAS JMP, using the REML procedure to account for repeated measures.

We found that participants were more likely to start quickly and then slow down (mean R=0.83; ratios less than 1 indicate drawing more slowly). An ANOVA showed that INSTRUCTION significantly affects the speed ratio ($F_{2,22}$=35.4, p<.0001). Drawing *creatively* results in the most constant rhythm (0.91), followed by *accurately* (0.82) and then *quickly* (0.77). Additionally, different patterns obtain when writing a long word as opposed to a short one. Participants performed faster at the end of long words with obtuse angles (1.1), such as *jewel*, performed at constant speed with acute angles (1.0), such as *joking*, and slowed down when angle=zero (0.76), such as *pure*. This suggests that participants may separate long words into separate chunks and then draw each chunk more consistently.

Inflation

Some properties of a gesture, e.g., direction and residual momentum, may also result in unintentional inflation or overshooting that goes beyond the limits of the keyboard itself. More interesting are deliberate gestures drawn outside the keyboard, as in Figure 1c. Since the gestures are constrained to pass approximately through each letter key in a word, we refer to this variation as *inflation* rather than *size*, but it can nevertheless be quantified using the ratio between the minimum bounding boxes of the performed and template gestures. We calculated the ratio between the participants' gestures and the gesture template's bounding box (Rb):

$$Rb_{word} = \frac{B_{gesture}}{B_{template}} \tag{3}$$

We found a significant effect of INSTRUCTION on inflation ($F_{2,22}$=185.9, p<.0001). A post-hoc analysis with Tukey HSD showed that words drawn *creatively* are significantly more likely to be inflated; with no significant difference between *accurately* and *quickly*, which corresponds to qualitative observations. Some participants intentionally drew outside the keyboard as they moved from one key to another, or drew very large gestures, beyond the normal bounding box.

Curviness

While a word-gesture template consists of lines and corners, in practice a gesture may also include curves [7]. Curviness is usually defined as the sum of the length of each segment whose radius of curvature is less than a certain threshold. In our data, calculation of a normalized curviness metric was complicated by the need to compare word templates with varying numbers of corners. We consider the absolute instantaneous angle among three points using tangents. Given a sequence of points $P = \langle p_i \rangle_{i...N}$, where $N = sizeP$, θ is the angle between vector $\vec{u} = \overrightarrow{p_i p_{i+1}}$ and vector $\vec{v} = \overrightarrow{p_{i+1} p_{i+2}}$ where $p_i, p_{i+1}, p_{i+2} \in P$, calculated as follows:

$$\theta = |atan2(|\vec{u} \times \vec{v}|, \vec{u} \cdot \vec{v})| \tag{4}$$

To emphasize the relative variations over the gesture, we measure the curviness (in degrees) by the standard deviation of all angles. The standard deviation is close to zero for straight lines but higher for curvy lines. For gestures consisting of several segments, e.g., *taxi*, even if the lines are drawn straight, the standard deviation is still affected by the corners.

Experiment 2: Block 1			
CONTROL	**INSTRUCTION**		
	Consistent	**Different**	**Varied**
Output	Make each phrase the same black color.	Make each phrase a different shade of green.	Make each phrase include at least two colors.
Input	Scribble on each letter.	Draw at different speeds.	Draw outside the keyboard.

Table 1. Examples of instructions that vary according to CONTROL (*Output* or *Input*) and INSTRUCTION (*Consistent*, *Different*, or *Varied*).

While different words may have different numbers of corners with different angles, we can still reliably distinguish a curvy gesture by setting a threshold value. If all lines are drawn relatively straight, the standard deviation is around 12; cusps (acute corners), scribbles, and mixed straight-line/curves are $>$12; while curves and loops (obtuse corners) are < 12.

An ANOVA showed that participants increase curviness when writing *quickly*, which is significantly different ($F_{2,22}$=60.5, p<.0001) from *accurately* or *creatively*. These results correspond to the qualitative observations.

In summary, participants tend to draw more curvy gestures when writing *quickly* and straighter gestures when writing *accurately*. This is not surprising when speed is the goal, given that the human motor system maximizes smoothness to reduce movement cost [20].

EXPERIMENT 2: CONTROLLING INPUT AND OUTPUT

Experiment 2 investigates whether users can explicitly control both their movement (gesture input) and the final result (gesture output). We chose RGB color as the output parameter space, since it is continuous, easy to quantify, and has relatively unambiguous semantics – most people agree on the meaning of the descriptor "red" as opposed to e.g. "messy".

Block 1 is a [2x3] within-participants design with two primary factors: CONTROL, with two levels (*input*; *output*) and INSTRUCTION, with three levels (*consistent*; *different*; *varied*). Participants are asked to gesture type phrases either by controlling a particular characterisic of their *input*, – "Go outside of the keyboard"; or by controlling a characteristic of their *output*, – "Try to make each phrase the same color of green". For each type of control, participants are asked to draw phrases that are *consistent*, *different*, or *varied*, as shown in Table 1). Phrases were chosen randomly from MacKenzie and Soukoreff's three- or four-word phrase sets [17].

We are also interested if users can control their gestures based on their relationship to the recipient or their current emotional state. Block 2 is a one factor within-participants design with two levels: *message recipient* and *sender emotion*. The task is open-ended: participants can choose how to interpret these instructions and make their own mappings between the instruction and the resulting variation in the output. Participants gesture-type three phrases, to three different recipients or to express three different emotions, as illustrated in Table 2. For example: "Draw the phrase for your partner" or "Express how you feel: happy".

Experiment 2: Block 2	
Recipient	Draw the phrase for your partner
	Draw the phrase for your best friend
Emotion	Express how you feel: happy
	Express how you feel: angry

Table 2. Experiment 2, Block 2: Sample instructions based on message recipient or sender emotion, replicated three times.

Participants

We recruited five right-handed men and seven women (mean age 27). All use mobile phones daily. Four use gesture-typing daily, the others are non-users. No participants had participated in Experiment 1.

Apparatus

We used the same LG Nexus 5 (Android 5.1) smartphone as in Experiment 1. We chose KB-2 for the gesture-typing recognizer since it allows drawing outside the keyboard area. Touch events are captured by Android Debug Bridge running on a desktop computer connected to the device via a USB cable. The touch data is sent back to the smartphone and post-processed by the application to determine gesture features.

We implemented a prototype *Expressive Keyboard* that maps gesture variation to the full range of RGB colors. We mapped inflation ratio to *red*, curviness to *green*, and speed consistency ratio to a *gradient of blue*. Gestures that use a constant speed when following the gesture template – a straight line from middle point to middle point of each letter in the word – map to the color black. For example, as the user slows down, the gesture turns from blue at the beginning of the phrase to another color at the end of the phrase. This mapping makes it technically feasible for users to generate all possible RGB color combinations.

Procedure

Sessions last from 30-60 minutes. Participants sit in a chair and hold the phone comfortably in their left hand, so they can perform all gestures with their right index finger. Participants are encouraged to talk aloud as they draw each word. They are asked to practice until they feel comfortable with the recognizer and can reliably produce different colors.

Each trial displays an instruction at the top of the screen, e.g. "Try to make each phrase a different shade of green"; with a three- or four-word phase centered below. Participants gesture-type three phrases in succession, on three separate lines, according to the condition (see Fig. 5.) For example, the *Varied-Output* condition "Make each phase include at least two colors; use as many colors as you can" is accomplished by speeding up or slowing down to create varied colors within each successive phrase.

In both blocks, participants may write the phrases as often as they like, before pressing 'next' to submit the current result and move to the next trial. After each condition, participants used a five-point Likert-style scale to rate how their output compared to their expectations.

The complete experiment consists of 30 trials: Block 1 includes 18 trials (2 CONTROLS x 3 INSTRUCTIONS x 3 replications); and Block 2 includes 12 trials (2 INSTRUCTIONS x 6 replications). Trials are counter-balanced within each block and across participants using a Latin Square.

At the end of the experiment, participants are asked to explain how the system generated colors in block 1; how they generated variations when asked to write to a particular recipient; and how they expressed specific emotions in block 2. We intentionally did not inform participants how the mappings work, so that we could determine whether or not they controlled their variation "unconciously".

Data Collection

In addition to touch events, we record values for CORRECTRATE, WORDACCURACY and FEATUREACCURACY. WORDACCURACY is when the gesture produced the correct word; and FEATUREACCURACY is when the gesture produced both the correct word and correct output properties. We also measure inflation, curviness, and speed consistency ratio. To reduce noise caused by dependencies in word characteristics, and to increase variation in general, we average each measure progressively throughout each phrase. CORRECTRATE (0-2) measures the participant's success in each condition. We defined a threshold value for each condition based on a pilot test and results from Experiment 1. For example, in the *varied-output* condition, a successful trial has a high blue value ($> 100/255$), and at least one other RGB color component in RGB that differs from the other phrases (difference $\geq 20/255$). Fulfilling both conditions results in CORRECTRATE=2, whereas fulfilling only one results in CORRECTRATE=1. We count number of errors based on how many times the participant erased a word before submitting a results. We record the screen and audio throughout to capture verbal comments.

Results

Participants were generally able to control the variation in their gestures (overall CORRECTRATE is 1.3 out of 2.0), but were unable to fully meet the goals of each condition. An ANOVA showed that Both CONTROL and INSTRUCTION significantly affect CORRECTRATE ($F_{1,11}$=18.5 and $F_{2,22}$=28.7 respectively, all p<.0001). Participants achieve significantly higher success rates when controlling output (1.4) than when controlling input (1.1) (Figure 3).

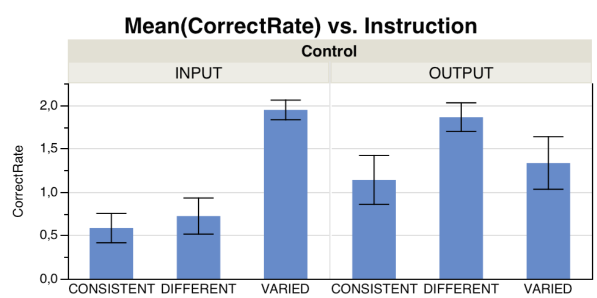

Figure 3. Participants are significantly more likely to control gestures based on output than on input.

A post-hoc test with Tukey HSD also showed a significant interaction between CONTROL and INSTRUCTION ($F_{1,11}$=37.5, p< .0001). Participants were most successful in the *Varied-Input* and *Different-Output* conditions (1.9 and 1.86 respectively) which are both significantly different from the others. The least successful conditions were the *Consistent-Input* and *Different-Input* (0.6 and 0.7 respectively).

Variation in Gesture Features

A post-hoc analysis with Tukey HSD revealed a significant interaction between CONTROL and INSTRUCTION for all gesture features. Participants were clearly able to control the size of the bounding box, as in the *Varied-Input* and *Varied-Output* conditions, but otherwise chose not to. Participants drew curvier gestures when in the *Constant-Input* and *Different-Output* conditions and significantly more straight gestures in *Constant-Output*. However, their natural inclination is to draw curvy gestures with non-constant speed. Explicitly controlling speed consistency appears to be more difficult.

The post-questionnaire on expectation-match level reveals that overall, 70% of the tasks matched their expectations (41% strongly satisfied, 18% neutral, and 12% dissatisfied). Based on the measurement criteria, 76.4% of the tasks successfully fulfilled at least one condition (50% fulfilled both).

Accuracy

The overall WORDACCURACY is 78.8%, which suggests that the participants are able to control both their input and output while retaining reasonable accuracy as compared to baselines from Experiment 1. Overall FEATUREACCURACY is lower at 45.1% (57.2% of correct words); participants sometimes re-wrote correctly recognized words to modify their output properties. The participants erased more incorrect words during the first trial replication (mean 2.4 vs. 1 word for the last trial replication). They also erased more correct words to modify the output properties in the first trial replication (mean 4 vs. 2). Both CONTROL and INSTRUCTION significantly affect the number of errors. Participants in conditions that focused on output made significantly more errors than in conditions focused on input (7.2 > 2.0, $F_{1,11}$=27, p<.0001). Participants also erased significantly more often when trying to be consistent than when trying to be different (6.5 > 2.9, $F_{2,22}$=4.3, p<.0001).

Conveying Emotion & Writing for Different Recipients

Participants varied their gestures when expressing certain emotions or when writing to different recipients. Participants used different strategies: six deliberately varied their gesture input; five varied their gesture output; and only one participant varied both.

When the hypothetical recipient was their boss or their parent, five participants reported they wanted to make the text color darker, and thus wrote more slowly and accurately. In contrast, when writing to a close friend or child relative, seven participants said they drew more slowly, with curvier gestures and detours, resulting in brighter colors (Fig. 5).

Three participants chose pink or red to write to their partners; P9 drew a heart shape that left the keyboard area. Participants expressed negative emotions using slower, straighter gestures, resulting in darker colors. Four participants associated 'angry' with greater speed and most expressed being busy by drawing faster, curvier gestures. Only one participant stated that they did not change their style of gesture-typing.

Discussion

Participants were able to control aspects of their gestures to produce their intended outcomes, although perhaps not as consistently as they would like. Participants found it easier to control the output of their gestures, i.e. to explicitly control the color of a phrase, than to control the characteristics of their input, i.e. to control the curviness, size and speed of their gestures. Surprisingly, only 25% (3/12) were able to correctly guess the mapping between color and their gestures.

Some participants took advantage of the variations in the color feedback to reflect upon and modify their performance, which was not possible for the *input* conditions. This suggests that participants faced a trade-off during the first time using *Expressive Keyboard*: they expended more effort rewriting the phrase until they got the desirable colors, but in the end were more successful in fulfilling the instructions. Clearly, the participants need more time to practice, since most participants were using gesture-typing for the first time. While both types of accuracy must be improved, we believe this is a promising start: novice users are unlikely to have a clear understanding of how gesture recognition algorithms work, but this should not prevent them from generating rich output by varying their gestures.

Participants demonstrated that they are capable of drawing certain types of gestures, e.g. extremely curvy or large gestures, even if they do not choose to do so when typing without rich output. By contrast, participants have difficulty maintaining a constant drawing speed. Gestures typically start quickly and then slow down, which with our mapping produces a color gradient in the *Varied-Output* condition. This should be a particularly simple gesture to control, but interestingly, only three found it easy and none could articulate how it works.

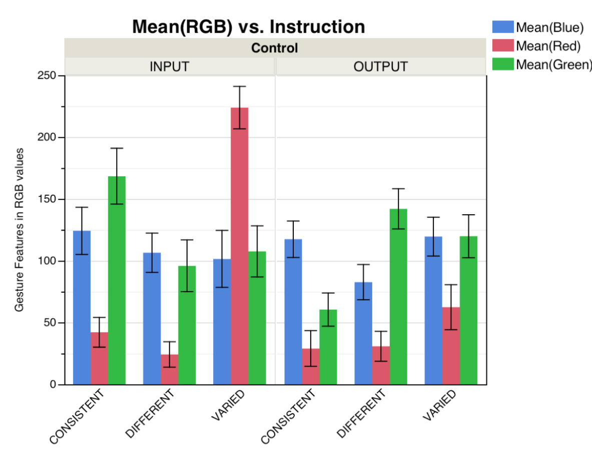

Figure 4. Participants can intentionally control gesture size when asked (BoundingBoxRatio), but do not vary it otherwise.

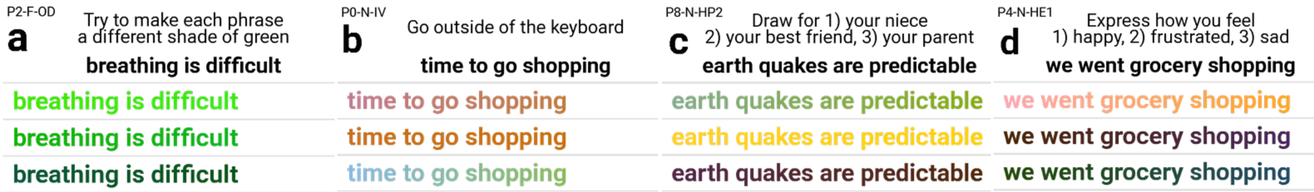

Figure 5. Successful control of color through gesture: a) [*instruction: Different-Output*] bright-green indicates curviness, dark green indicates straight lines; b) [*instruction: Varied-Input*] inflating the gesture increased red values; c) [*instruction: recipients*] P8 changed the color deliberately for different recipients; d) [*instruction: express emotion*] P4 made curvier gestures with more detours for 'happy', and slow and less curvy gestures for 'sad'.

We observed large individual differences across participants with respect to the above features. Perhaps not surprisingly, this suggests that each individual is likely to appropriate expressive gestures in their own way, and generate distinctive, personal gesture styles, just as they do with their handwriting.

EXPERIMENT 3: ECOLOGICAL VALIDITY

We designed a third experiment to explore how *Expressive Keyboards* are used in a more ecologically-valid setting, and when mapped to more complex features than color. The experiment is a [2x2] repeated measures with two factors: KEYBOARDTYPE {*baseline keyboard*, *Expressive Keyboard*} and TEXTTYPE {*user-generated*, *prescribed*}. The *baseline keyboard* is a standard gesture-typing keyboard. The *Expressive Keyboard* generates a *dynamic font* which shape and colors changes depending on the gesture features.

Implementation

We implemented an *Expressive Keyboard* that maps gesture features to a user-defineable dynamic font. The dynamic font is created through a simple application we developed that lets users draw each letter by defining control points. Users can create several typefaces and dynamically interpolate between them to generate new intermediate fonts continuously. The interpolation between n typefaces changes the position of component control points based on a weighting function:

$$typeface_{interpolated} = \frac{\sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{typeface_j weight_j}}{\sum_{\mathrm{j}=1}^{\mathrm{n}} \mathrm{weight_j}} \quad (5)$$

where each typeface is a vector of control points for each letter.

For Experiment 3, we predefined a font with n=2; one typeface is more skewed (italic) than the other. We used the same gesture features as for Experiment 2, with the speed consistency ratio mapped to the weight ratio for font interpolation; inflation ratio mapped to stroke thickness (bold); and curviness mapped to the magnitude of random offsets applied to each control point. We also used the color mapping used in Experiment 2.

Participants

We recruited six pairs of friends, seven men and five women (age range 19-40, mean 25.4); all use mobile phones daily. Half of them use gesture-typing daily, the others are non-users. No participants had participated in the two previous experiments.

Apparatus

For the first session, we customized an Android chat application[3] to capture the gesture data as well as the typed words. For the second session, we developed a custom Android application that presents either the *baseline keyboard* or an *Expressive Keyboard* that renders the dynamic font. In addition, we developed a simple font engine that lets users define static typefaces by connecting control points to form each letter; this software was used to design the font sets used in the experiment but not in the experiment itself. For recognition, we used KB-2 on the same LG Nexus 5 (Android 5.1) smartphone as in the previous experiments.

Procedure

Participants sit comfortably in a chair while gesture-typing. The experiment is divided into three sessions. In the first session, we set up live conversations between pairs of participants to collect gesture-typed texts in a natural setting. Each pair of participants chat for 15 minutes without any restriction on how to gesture-type. For the second session, we select five sentences from the chat as the *user-generated* text.

The second session consists of three blocks of five trials each. This is an individual task where the participant has to write five *user-generated* and five *prescribed* sentences (from news, blogs, etc). Participants are instructed to gesture-type as if writing to their peer and to assume the peer will see the same output. The session always starts with the *prescribed* text with *baseline keyboard*, followed by an introduction to the *Expressive Keyboard* and the mapping used. Participants are encouraged to practice to understand how the system works; no participant practiced longer than five minutes. For the next two blocks participants are asked to write both the *prescribed* and *user-generated* texts (counter-balanced across participants). We do not specifically tell them how to use *Expressive Keyboard* and let them use it as they like. Throughout the session, we ask the participants to describe aloud what they want to do and what they are thinking.

The third session is a quiz (three blocks of three trials). We ask them to gesture-type "hello" three times with specific output goals: 1) bold and red, 2) italic and containing blue, 3) green. Finally, we interview them regarding their preference and hold a mini brainstorming session on how they might define their own features and mapping if they could. An experiment session last for 40 minutes.

[3] AndroidHive: http://www.androidhive.info/

KeyboardType	TextType	
	User-generated	*Prescribed*
WordAccuracy		
Baseline Keyboard	95%	75.4%
Expressive Keyboard	68.5%	79.5%
FeatureAccuracy		
Expressive Keyboard	55.9%	61.7%

Table 3. Accuracy in Experiment 3. Participants changed the way they gesture-type with different keyboards, and are more likely to explore with *Expressive Keyboard*. Although using *Expressive Keyboard* does not necessarily decrease word-accuracy, participants erased correct words to modify the output properties leading to low values for feature-accuracy.

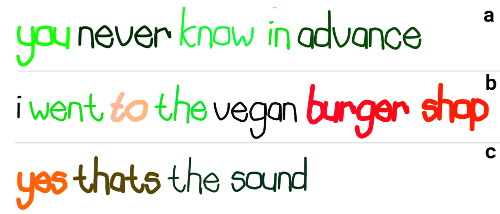

Figure 6. Using Expressive Keyboard in Experiment 3: a) P2 naturally made curvy gesture (green) but made straight gestures (dark green) to emphasize some words, b) P6 deliberately made two words ("burger shop") the same shape and color, c) P3 emphasized the first word, but then deliberately made his gesture more precise (dark color) instead of curvy.

Data Collection

We calculate the three gesture features as well as WordAccuracy and FeatureAccuracy as in Experiment 2. We log the timestamp and 2D coordinate of each touch event, and record the screen and audio to capture verbal comments.

Results and Discussion

Out of 2312 performed gestures, we removed 183 that were not gesture-typed, most of which were single-letter words. For the quiz, we collected 108 gestures and removed 2 outliers where the participants had lifted their finger at the start of the gesture.

Standard Gesture-Keyboard vs. Expressive Keyboard

We ran an ANOVA test to compare performance of KeyboardType and TextType. WordAccuracy is significantly affected by KeyboardType ($F_{1,11}$=30.7, p<0.0001). There is a significant interaction between KeyboardType and TextType ($F_{1,11}$=15, p=.0001). However, further analysis with Tukey HSD showed that the only significant difference is found when the participants wrote *user-generated* text using *baseline keyboard* compared to *Expressive Keyboard*. There was no significant difference between *baseline* and *Expressive Keyboard* when writing a *prescribed* text. This suggests that the use of the *Expressive Keyboard* does not necessarily decrease WordAccuracy; instead the context when chatting may help the participants better match the design parameters of the recognizer.

However, with *Expressive Keyboard*, there were cases in which the participants erased a correctly-recognized word to modify the output properties: 18.9% for *user-generated* and 23.3% for *prescribed*. The overall FeatureAccuracy, which represents accuracy of *Expressive Keyboard* is shown in Table 3. The time spent to draw a gesture also significantly increased when using *Expressive Keyboard* ($F_{1,11}$=64.9, p<.0001), mean 2 seconds per word, while they spent the least time when chatting (0.7spw).

There is a significant effect of KeyboardType on inflation and curviness ($F_{1,11}$=30.6, p<.0001 and $F_{1,11}$=5.6, p<.0177 respectively). The participants significantly inflated their gestures when using *Expressive Keyboard* (mean 1.6) as compared to *baseline keyboard* (mean 1.2). There is no significant difference with regards of the speed consistency.

From the post-questionnaire, we learned that the participants took advantage of the fact that they could change the output properties. All participants stated that they changed the way they gesture-typed with *Expressive Keyboard*. Most used it to highlight a specific word or phrase in the sentence, rather than trying to control the appearance of every word. Two of them mentioned they changed the properties to match their intonation when reading the sentence. Three of them stated that they expressed their feelings or mood when writing, e.g. "when it's something happy I tried to write faster [so] that text becomes green and blue" (P6).

An interesting appropriation of *Expressive Keyboard* is to reflect on their own gesture-typing habits. P1, P10, and P11 realized that they tend to make curvy gestures, and they deliberately let the output change according to their natural input style. On the contrary, P3 and P8 made a special effort to make the output text as similar as possible (e.g. all black). This suggests that continuous changes to output properties can provide important feedback for the participants and may change their behaviour when gesture-typing – not only to customize the output, but also to try to gesture-type more precisely.

Control of Gesture Features

In the quiz section, the average WordAccuracy was 86%. Participants were most accurate when controlling speed consistency (33 out of 36 trials, or 92%) and curviness (32/36 or 89%), but less accurate when controlling inflation (26/34 or 76%). Further analysis revealed that the recognition error was caused by 1) too much deformation (6 out of 15 errors), 2) faulty start or end position (6/15), and 3) removing the finger too early (3/15). This suggests that while many factors affect recognition rate, certain types of intentional variation may increase error to a small extent.

The instruction had a significant effect on all the gesture features (all p<.0001). A post-hoc test with Tukey HSD showed a significant difference between the instruction "italic" and the others. In Fig. 7, we can see that in general the participants speed-up (mean rate=1.5) with higher variability when asked to make the output italic; however both the value and the variability dropped for instructions which required them to keep the speed constant (0.86 and 0.84 for "bold" and "green" respectively).

Similar significant differences also appear for inflation rate. Pairwise t-tests with Tukey HSD showed that there is a sig-

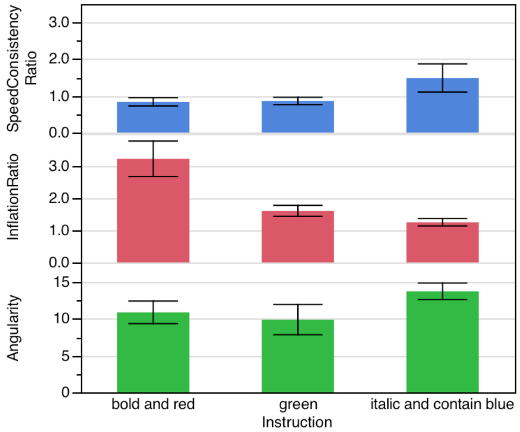

Figure 7. Participants successfully varied or kept constant the inflation and speed consistency when instructed. They naturally made curvy gestures, but could make them more curvy when instructed.

nificant difference between the instruction "bold" and the others. However, for "green", the inflation rate is a bit higher (mean=1.6). Based on our observation, this is because the participants overshot when trying to make a curvy gesture.

Meanwhile, the gestures are quite curvy for all instructions: with means 10.9, 9.9, and 13.8 for "bold", "green", and "italic", respectively (a value of 12 indicates minimum curviness). Gestures are most curvy when instructed implicitly, however, it is not significantly different from "bold". This suggests that inflating the gesture also increased curviness. Overall, participants naturally made curvy gestures, but were able to increase curviness when necessary.

GENERAL DISCUSSION

When provided the tools, participants took advantage of the variation in their gestures to enhance communication. Experiment 3 also confirms that participants were able to selectively control speed and inflation, and naturally made curvy gestures. Nine participants felt that using *Expressive Keyboard* was more enjoyable than using a standard gesture-typing keyboard (1 neutral, 2 disagree). This means that despite the fact that it is more time-consuming and less error-prone, the participants were more willing to engage in gesture-typing itself.

Since all participants were necessarily novice users, a more longitudinal study of *Expressive Keyboards* is needed to determine whether accuracy improves with experience. Additional factors may also affect user performance when using Expressive Keyboard, for example, in Experiment 3 six participants mentioned that they sometimes changed their hand position to increase the precision of their finger in gesturing, e.g. from using thumb to index finger.

While our investigation mainly focused on intentional variation, we believe unintentional variation is as interesting and can enrich inter-personal communication. Simply mapping this variation to small differences in the rendered output would generate a degree of expression that could be recognizable as personal styles. A user may not specifically control their gesture when writing on a bumpy bus, yet the output can reflect the writing process – something that handwriting can capture easily. Conscious control of this variation would not be required to implicitly communicate style, personality, context, or mood. We are not interested in 'identifying' expression or emotion based on gestures, since we believe richer interpretations will be result from a system in which users can develop their own communication contexts and related meanings.

In real usage, users should be able to vary the sensitivity of the output variation or turn it off completely [2], especially in cases where the user prefers to generate more formal output. Users should also be able to (re)design their own feature detectors, text-rendering properties, and mappings linking the two.

Finally, we are not seeking a slavish recreation of traditional handwriting, nor a replacement for other communication channels such as emoticons. Instead, we aim for new forms of nuanced textual communication supported by and interacting with new technologies: touch screens, probabilistic language models, vector graphic fonts, etc. The precise ways in which these new opportunities would be used for communicating style or emotional expression likely cannot be *designed*, but must emerge through appropriation of the system by communities of users.

CONCLUSION AND FUTURE WORK

This paper proposes a novel approach for gesture-typing that maps gesture variation to one or more continuous properties of the rendered text, producing rich output that is both accurate and under user control. Using *Expressive Keyboards*, users can simultaneously control both text content and style – captured as changes in shape, weight, color or other output parameters. Through a series of experiments, we established that users do indeed vary their gesture-typing in quantifiable ways, that they can control such variations intentionally, and that they find a system that uses the variation to enrich the rendered text interesting and enjoyable.

This work validates some of the concepts behind leveraging variation in gesture-typing behavior, opening up a wide range of research directions. In addition to improvements in the variety and quality of calculated gesture features and our dynamic font engine, we are investigating how *Expressive Keyboards* might be used to improve gesture-typing accuracy through progressive feedback.

We also plan to expand the different forms of output that can be mapped to user gesture, including parametric emoji, handwriting, and nuanced speech synthesis. We are particularly interested in studying the use of *Expressive Keyboards* in ecological settings to see how it is appropriated for providing new types of digital communication.

ACKNOWLEDGMENTS

This research was supported by ERC grant number: 321135, *CREATIV: Creating Co-Adaptive Human-Computer Partnerships*. We would like to thank Shumin Zhai for his helpful comments on earlier drafts, as well as members of ExSitu for their valuable suggestions.

REFERENCES

1. Ängeslevä, J., Reynolds, C., and O?Modhrain, S. Emotemail. In *ACM SIGGRAPH 2004 Posters*, R. Barzel, Ed., ACM Press (2004).
2. Arif, A. S., Kim, S., Stuerzlinger, W., Lee, G., and Mazalek, A. Evaluation of a smart-restorable backspace technique to facilitate text entry error correction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, ACM (New York, NY, USA, 2016), 5151–5162.
3. Azenkot, S., and Zhai, S. Touch behavior with different postures on soft smartphone keyboards. In *Proc. ACM MobileHCI 2012*, ACM Press (2012), 251–260.
4. Bi, X., Azenkot, S., Partridge, K., and Zhai, S. Octopus: evaluating touchscreen keyboard correction and recognition algorithms via "remulation". In *Proc. ACM CHI 2013*, ACM Press (2013), 543–552.
5. Burgbacher, U., and Hinrichs, K. Modelling human performance of stroke-based text entry. In *Proc. ACM CHI 2015*, ACM Press (2015), 137–143.
6. Buschek, D., De Luca, A., and Alt, F. There is more to typing than speed: Expressive mobile touch keyboards via dynamic font personalisation. In *Proc. ACM MobileHCI '15*, ACM Press (2015), 125–130.
7. Cao, X., and Zhai, S. Modelling human performance of pen stroke gestures. In *Proc. ACM CHI 2007*, ACM Press (2007), 1495–1504.
8. Caramiaux, B., Bevilacqua, F., and Tanaka, A. Beyond recognition: using gesture variation for continuous interaction. In *ACM CHI EA 2013*, ACM Press (2013), 2109–2118.
9. Caramiaux, B., Donnarumma, M., and Tanaka, A. Understanding gesture expressivity through muscle sensing. *ACM Trans. Comput.-Hum. Interact. 21*, 6 (Jan. 2015), 31:1–31:26.
10. Dunlop, M., and Levine, J. Multidimensional pareto optimization of touchscreen keyboards for speed, familiarity and improved spell checking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, ACM (New York, NY, USA, 2012), 2669–2678.
11. Godøy, R. I., and Leman, M., Eds. *Musical Gestures: Sound, Movement, and Meaning*. Routledge, 2009.
12. Goodman, J., Venolia, G., Steury, K., and Parker, C. Language modeling for soft keyboards. In *Eighteenth National Conference on Artificial Intelligence*, American Association for Artificial Intelligence (Menlo Park, CA, USA, 2002), 419–424.
13. Gunawardana, A., Paek, T., and Meek, C. Usability guided key-target resizing for soft keyboards. In *Proceedings of the 15th International Conference on Intelligent User Interfaces*, IUI '10, ACM (New York, NY, USA, 2010), 111–118.
14. Hunt, A., Wanderley, M. M., and Paradis, M. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research 32*, 4 (December 2003), 429–440.
15. Iwasaki, K., Miyaki, T., and Rekimoto, J. Expressive typing: A new way to sense typing pressure and its applications. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, ACM (New York, NY, USA, 2009), 4369–4374.
16. Kristensson, P. O., and Zhai, S. Command strokes with and without preview: Using pen gestures on keyboard for command selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, ACM (New York, NY, USA, 2007), 1137–1146.
17. MacKenzie, I. S., and Soukoreff, R. W. Phrase sets for evaluating text entry techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, ACM (New York, NY, USA, 2003), 754–755.
18. McNeill, D. *Gesture and Thought*. University of Chicago Press, 2005.
19. Neumann, P., Tat, A., Zuk, T., and Carpendale, S. Keystrokes: Personalizing typed text with visualization. In *Proceedings of the 9th Joint Eurographics / IEEE VGTC Conference on Visualization*, EUROVIS'07, Eurographics Association (Aire-la-Ville, Switzerland, Switzerland, 2007), 43–50.
20. Pastel, R. Measuring the difficulty of steering through corners. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, ACM (New York, NY, USA, 2006), 1087–1096.
21. Pirzadeh, A., Wu, H.-W., Bharali, R., Kim, B. M., Wada, T., and Pfaff, M. S. Designing multi-touch gestures to support emotional expression in im. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '14, ACM (New York, NY, USA, 2014), 2515–2520.
22. Tu, H., Ren, X., and Zhai, S. Differences and similarities between finger and pen stroke gestures on stationary and mobile devices. *ACM Trans. Comput.-Hum. Interact. 22*, 5 (Aug. 2015), 22:1–22:39.
23. Walther, J. B., Loh, T., and Granka, L. Let me count the ways: The interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of Language and Social Psychology 24*, 1 (2015), 36–65.
24. Zhai, S., and Kristensson, P. O. The word-gesture keyboard: Reimagining keyboard interaction. *Commun. ACM 55*, 9 (Sept. 2012), 91–101.