31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Character Activation Time Prediction Model for Tongue-Typing: Adaptation of Fitts's Law

Héctor A. Caltenco *Student Member, IEEE,* Eugen R. Lontis, Johannes J. Struijk, Morten E. Lund *Student Member, IEEE*, and Lotte N.S. Andreasen Struijk *Member, IEEE*

***Abstract*—This paper presents the development of a character activation time prediction model for tongue-typing. This model is based on a modification of Fitts's law that is more suitable for tip-of-tongue selectivity tasks around the palatal area. The model was trained and evaluated with data from tongue-selectivity experiments using an inductive tongue-computer interface. It takes into account the movement amplitude, target position, interactions between them, character disambiguation time and error correction time.**

I. Introduction

BOTH, speed and accuracy of target-directed movements play a big role for determining the performance of text input devices. One of the major contributions to research the effects of speed and accuracy was done by Fitts (1954) [1], who was the first to propose a formal relationship linking movement time (MT) and the index of difficulty of target selection (*ID*):

$$MT = a + b\ ID \tag{1}$$

It states that the time it takes a human to move rapidly to a target area is directly related to the index of difficulty for target selection, which is the logarithm of the relation between the movement amplitude (*D*) and the target effective width (*W*).

$$ID = \log_2\left(\frac{2D}{W}\right) \tag{2}$$

In subsequent years, the formulation was used in a systematic way for numerous studies and was found to be so general that it became known as Fitts's law. The experimental validation of Fitts's law has been totally or partially confirmed for a variety of movements, limbs and muscle groups, experimental conditions and manipulation devices, see [2] for a review.

Although Fitts's law has been generally accepted by many researchers as a good and practical working tool, many have worked on several modifications to the original equation (1) that provide a better data fitting taking into account specific experimental conditions [3-7]. It is evident that despite of its generality, Fitts's law does not hold completely for every motor task, experimental condition or muscle group. The law assumes that all targets have the same "accessibility", meaning that the target accessibility does not vary according to the position of the target.

In the case of tip of the tongue (ToT) selectivity around the palatal area, Caltenco et al. [8] state that anterior and middle palatal areas are easier to access with the tip of the tongue than posterior and lateral areas. This suggests that ToT accessibility is not the same for all targets and (2) may not hold completely and must be adapted.

ToT selectivity experiments were performed to obtain information about the difficulty of target selection according to target position and inter-target amplitude. This information is used to build a character activation time prediction model (CATPM) for tongue-typing, which is presented as the addition of ToT movement time (*MT*), character disambiguation time (*DT*), time to correct errors (*CT*), and reaction and other mental computation times (*RT*).

$$CAT = MT + DT + CT + RT \tag{3}$$

The results can be used to create an optimal character arrangement for the ambiguous tongue-computer interface. This will be useful to increase the typing throughput of a tongue-computer interface and design more efficient computer input devices for individuals with severe sensory-motor impairments.

II. Methodology

A. Experimental Setup

Two different sensor layouts were tested: L_1 and L_2. Each layout contained a printed circuit board (PCB) with 8 sensors as a tongue mouse-pad area (TMP) and a PCB with 10 sensors as a tongue key-pad area (TKP). L_1 contained the TMP in the anterior and the TKP in the posterior part of the upper palate, while L_2 the other way around (Fig. 1).

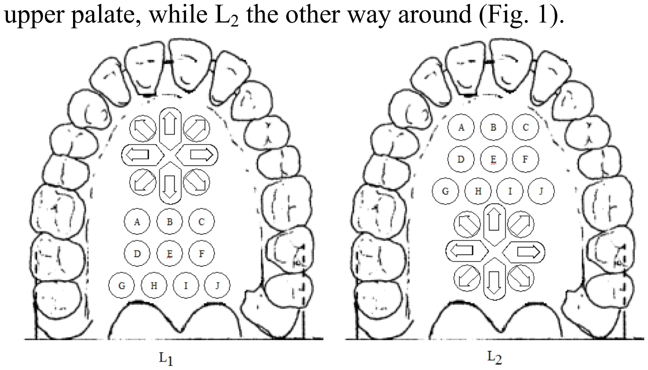

Fig. 1. Associated characters and mouse directions to each sensor for both layouts: inside view of the upper palate. Left: L1 (TMP front, TKP back), Right: L2 (TKP front, TMP back).

Manuscript received April 7, 2009. This work is supported in part by the Danish Technical Research Council, TKS A/S.

H. A. Caltenco, E. R. Lontis, J.J. Struijk M.E. Lund and L. N. S. Andreasen Struijk are with Center for Sensory Motor Interaction (SMI), Department of Health Science and Technology, Aalborg University, DK-9220 Aalborg, Denmark. (Corresponding author: H.A. Caltenco, e-mail: hca@hst.aau.dk).

Experiments were performed in 10 able-bodied individuals (5 per layout) over 3 consecutive 2 hour/day sessions. A personalized palatal plate containing the PCBs with inductive sensors was placed inside the subjects' mouth and fixed on the hard palate. A 2x2 mm. activation metal piece was glued in the subjects' ToT using biocompatible Hystoacryl®. The activation signal was sent to a computer using RS232 serial protocol.

Each session consisted on a typing task using the TKP and a pointing/tracking task using the TMP. During the typing task, the subject typed a predetermined series of character sequences. A total of 54 trials were performed per session, where the subject was asked to type the defined sequence as fast and accurate as possible for 30 seconds. The typed sequences were either repetitive (e.g. "AAA…" or "EEE…"), sequential by rows/columns (e.g. "ABCABC…" or "HEBHEB…"), or unordered (i.e. "EAICGFBHDJE…"). Results of the pointing/tracking task are reported in [9].

B. *Modeling Considerations*

A problem occurs with the original Fitts's index of difficulty (2) when the targets overlap (*D<W/2*). In this case *ID* becomes negative, which presents a serious theoretical issue. MacKenzie [6] presents a modification of Fitts's index of difficulty which deals with this problem:

$$ID = \log_2\left(\frac{D}{W} + 1\right) \tag{4}$$

Regarding the experimental data, since there are 10 character sensors, there is a total of 10x10=100 possible start-sensor to end-sensor transitions. However, only 57 transitions were identified from the typed sequences. For these transitions, the position of the end-sensor is considered as the Cartesian distance along the centrolateral and anteroposterior directions (*Ax*, *Ay*), where the origin of the coordinate system is located in the incisive papillae. While the movement amplitude (*D*) is considered as the relative scalar distance from the start-sensor to the end-sensor.

However, in order to create a model to predict *MT* for tongue-typing, it is necessary to include the effect of (*Ax, Ay)* on *MT*. Apart from this effect there could be interactions between the end-sensor position and the movement amplitude that affect *MT*. For example, *D* may have different effect for sensors located at different distance from the incisive papillae, or *Ax* may have more effect for sensors located in different position along *Ay*. Therefore, a modification of (4) for ToT-typing is presented in (5), where different weights are given to the movement amplitude (*D*), end-sensor position (*Ax*, *Ay*) and interactions between them.

The values of the distances (*D*, *Ax* and *Ay*) and its interactions, multiplied by their own weights, are used to calculate *ID*, which will be used to predict the movement time for any transition. Therefore, these values are referred as predictors throughout this paper. The predictor weights (d, a_x, a_y, i_{DX}, i_{DY}, i_{XY} and i_{DXY}) can be interpreted as the relative importance of each predictor's contribution to *ID*.

$$ID = \log_2\left(\begin{aligned}&\frac{d\ D + a_x\ Ax + a_y\ Ay}{W}\\ &+\frac{i_{DX}\ D\ Ax + i_{DY}\ D\ Ay + i_{XY}\ Ax\ Ay}{W^2}\\ &+\frac{i_{DXY}\ D\ Ax\ Ay}{W^3} + 1\end{aligned}\right) \tag{5}$$

III. Experimental Results

There is an intrinsic tradeoff between speed and accuracy in this type of target selection tasks. Typing speed is measured as the movement time between the activation of correct sensors (*MT*) and accuracy in error rate (*Err%*). Results for each performance measure are presented below.

A. *Movement Time*

The movement time for each transition is displayed in Fig. 2, where it can be observed that *MT* does not only depend on the distance between sensors but on other factors related to the position of the target inside the mouth, since L_1 generally presents higher *MT* values than L_2, which confirms that posterior areas of the palate have lower accessibility. But transitions to sensors "E-F" present higher *MT* values when from sensors "G-J" than from sensors "A-C". Also, sensor "A" is easier to access from "C" than from "E", even though the distance is shorter. This suggests that there is an important interaction between sensor position and distance.

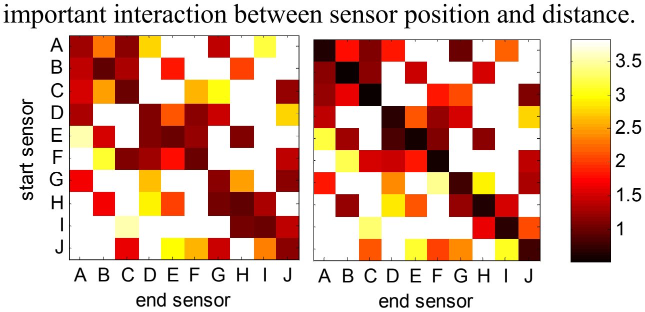

Fig. 2. Movement time matrix (in seconds) for the possible sensor transitions; blank squares indicate that the transition was not recorded. Left: L_1, Right: L_2.

B. *Typing Accuracy*

Even though, the user was asked to type as fast and accurate as possible, there were several erroneous activations during the typing tasks. This is not surprising since the tongue is not naturally used for these types of task, and there is no visual feedback that closes the motor control loop. Therefore it is easy to undershoot or overshoot a target, especially for transitions with large amplitudes.

A confusability map was obtained to describe the frequency of erroneous activations between "wanted" and "typed" characters (see Fig. 3). The highest error rates resulted from selecting a sensor one position anterior to the target sensor, especially when aiming for the posterior row (e.g. typing "F" while aiming for "J"). This can be attributed to the fact that the anterior part of the palate has better accessibility than the posterior part.

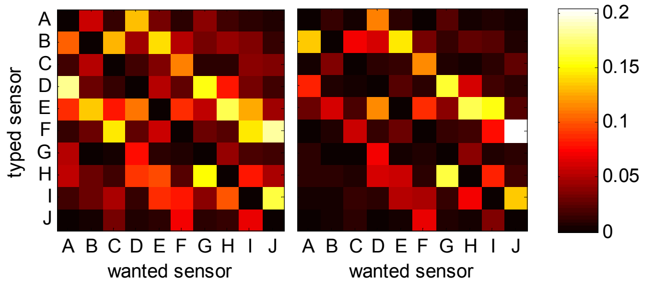

Fig. 3. Confusion matrix for rate of erroneous activations between wanted and typed targets. Left: L_1, Right: L_2.

IV. Character Activation Time Prediction Models

A. Predictors and Interactions Considered

From the k=1,…,7 predictors, some predictors will be included and some will be excluded from the model. This inclusion/exclusion of each predictor (k) is represented with a logical value (L_k). Therefore (5) can be represented for each training data point (t) by:

$$ID^t = \log_2\left(\sum_k w_k P_k^t L_k + 1\right) \tag{6}$$

Where:

$$P^t = \left\{\frac{D^t}{W}, \frac{Ax^t}{W}, \frac{Ay^t}{W}, \frac{D^t Ax^t}{W^2}, \frac{D^t Ay^t}{W^2}, \frac{Ax^t Ay^t}{W^2}, \frac{D^t Ax^t Ay^t}{W^3}\right\}$$

$$w = \{d, a_X, a_Y, i_{DX}, i_{DY}, i_{XY}, i_{DXY}\}$$

$$L = \{L_k \in \{0,1\}\}\ \forall k = 1,2,...,7$$

There could be different combination of predictors (and interactions) that could be used to build CATPM. However, not all the possible combinations make sense and not all the predictors have a sufficiently strong correlation to MT. Predictors Ax, Ay and $Ax*Ay$ resulted to have low correlation with MT (r < 0,20). This suggests that they alone have no effect on movement time, but they could be considered as an interaction with the distance.

On the other hand, when the correlation between predictors is high (r > 0,85), including both predictors in the same model will provide unnecessarily redundant information. Therefore two highly correlated predictors will not be included in the same prediction model, such is the case of D and ($D*Ax$ or $D*Ay$), Ax and $Ax*Ay$, and $D*Ax*Ay$ and ($D*Ax$ or $D*Ay$). In order not to deviate very much from the original Fitts's Law, predictor D is always included either alone or as a part of an interaction.

TABLE I
POSSIBLE CATPMS THAT CAN BE BUILT DEPENDING ON THE INCLUSION/EXLCLUSION OF PREDICTORS (VECTOR L)

	L_1	L_2	L_3	L_4	L_5	L_6	L_7
$CATPM_1$	1	0	0	0	0	0	0
$CATPM_2$	1	1	1	0	0	0	0
$CATPM_3$	1	0	0	0	0	1	0
$CATPM_4$	1	0	0	0	0	0	1
$CATPM_5$	0	1	1	1	1	0	0
$CATPM_6$	0	1	1	0	0	0	1
$CATPM_7$	0	0	0	1	1	0	0
$CATPM_8$	0	0	0	1	1	1	0
$CATPM_9$	0	0	0	0	0	1	0
$CATPM_{10}$	0	0	0	0	0	0	1

Based on the considerations above, different values for L are chosen in Table I. With which different IDs (and therefore different CATPMs) can be built assuming a linear (1) or other type of relation between ID and MT, and finding predictor weights (w_k) depending on the predictor values used for the training set (P_k^t) and the inclusion criteria (L_k).

B. Direct Multiple Linear Regression

Several researches have proposed alternative quadratic or power relations between MT and ID [7, 10], instead of a linear relation proposed by Fitts (1). These quadratic and power laws are based on well documented kinematic theory and provide better data fit for their target selection tasks. Instead, let us assume an exponential (base2) relation (7) between ID and MT.

$$MT = a + b\,2^{ID} \tag{7}$$

The obtained curve of this assumption resembles the power law proposed by Kvalseth [7]. It does not only provide better or equal data fitting than the quadratic, power or linear relations on our training set, but has the advantage of simplifying the construction of the CATPM as a multiple linear regression problem:

$$MT^t = b\sum_k w_k P_k^t L_k + (a+b) \tag{8}$$

It is worthy to mention that (a+b) is a derived from non-informational aspects of sensor selection, independent from the difficulty of target selection, but dependent on other factors like layout and feedback complexity. These non-informational aspects could be interpreted as the combination of the user and system's reaction time (RT) to select a target with the ToT.

C. Character Activation Time Estimation

The TKP contains fewer sensors that there are letters in the English alphabet. Therefore more than one character should be mapped into each sensor and a character disambiguation method must be used. If an optimal character arrangement should be done (not necessarily in alphabetic order), different optimization criteria are considered: MT, DT, CT, and RT.

These optimization criteria depend on different factors, like predictors (P), desired character disambiguation order (KS), dwell time to type a character (dt), error likelihood of a target sensor (E_j), average time to correct an error (ct), and other factors that affect reaction time like: layout complexity, feedback, user frustration, user motivation, etc.

Equation (9) describes a linear relation between CAT and each predictor (P_k) for any start-sensor "i" to any end-sensor "j". But it introduces the rest of the optimization criteria for character arrangement based on (3).

$$CAT(i,j,KS) = MT(i,j) + DT(KS) + CT(j) + RT$$

$$CAT(i,j,KS) = b\sum_k w_k P_k^{i\to j} L_k + dt\ KS + ct\ E_j + (a+b) \tag{9}$$

$$CAT(i,j,KS) = \sum_k \tilde{w}_k P_k^{i\to j} L_k + dt\ KS + ct\ E_j + \tilde{w}_0$$

D. Regression Results

Using multiple linear regression, the predictor weights were obtained for each model. The model with highest adjusted coefficient of determination (adjR^2) was $CATPM_4$ for both layouts. This model was also the one that presented lesser mean squared error (see Table II). The model is presented in (10) and (11) for L_1 and L_2, respectively. The performance of the models is evaluated with each data pair of the testing dataset (v). A comparison between the predicted and the test set values (MT^v) is presented in Fig. 4.

TABLE II
ADJUSTED COEFFICIENT OF DETERMINATION AND MEAN SQUARE ERROR FOR THE TRAINING AND TESTING DATA SETS AND FOR EACH LAYOUT

	adjR^2		MSE train		MSE test	
	L_1	L_2	L_1	L_2	L_1	L_2
$CATPM_1$	0,181	0,335	1,025	0,478	0,874	0,336
$CATPM_2$	0,183	0,336	1,021	0,478	0,866	0,334
$CATPM_3$	0,183	0,336	1,022	0,478	0,866	0,334
$CATPM_4$	0,192*	0,350*	1,010*	0,467*	0,838*	0,328*
$CATPM_5$	0,182	0,275	1,023	0,521	0,872	0,364
$CATPM_6$	0,104	0,169	1,120	0,597	1,063	0,438
$CATPM_7$	0,175	0,263	1,032	0,530	0,884	0,377
$CATPM_8$	0,175	0,270	1,031	0,525	0,883	0,369
$CATPM_9$	0,000	0,001	1,250	0,718	1,312	0,569
$CATPM_{10}$	0,088	0,156	1,140	0,607	1,102	0,456

* Indicates maximum R^2 or minimum MSE

$$CAT_{L1} = 0,795\frac{D}{W} - 0,042\frac{D\ Ax\ Ay}{W^3} + dt\ KS + ct\ E + 0,811 \quad (10)$$

$$CAT_{L2} = 0,855\frac{D}{W} - 0,079\frac{D\ Ax\ Ay}{W^3} + dt\ KS + ct\ E + 0,461 \quad (11)$$

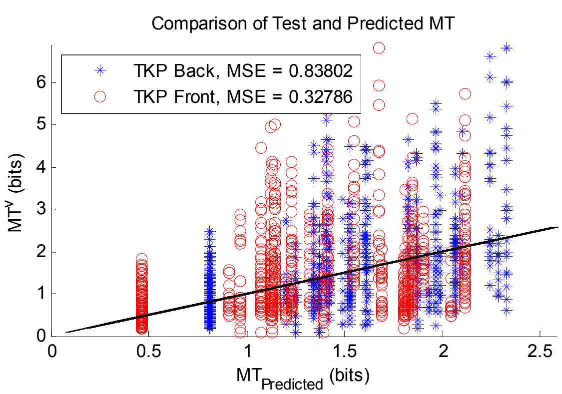

Fig. 4. Comparison of recorded (test dataset) and predicted movement time for each layout. The line represents the equation $MTv = MT_{Predicted}$

V. Discussion and Conclusion

This paper presents the development of a character activation time prediction model, which is based on a modification of Fitts's law that is more suitable for tongue-typing. During the typing tasks, the subjects were asked to type as fast and accurate as possible. However, accuracy resulted to be low. The main reasons for this is that sensors are embedded on a flat PCB and do not provide any "tactile" or "visual" feedback of weather the sensor that is currently being selected is the correct one or not.

The regression was done assuming that in our dataset, there was a single character associated to each sensor (KS = 1) and there was no dwell time (dt = 0). Also the subjects did not correct any error during the typing trials (ct=0). However if we want to predict the performance of "real typing", dt and ct should be different from zero, and KS will be greater than one. Further experiments are being realized with different dwell times and with visual or tactile feedback. It would also be interesting to include the effects of dwell time and activation threshold in the error likelihood (E_j).

The experiments took place over 3 consecutive 2 hours/day sessions. This may not be enough to precisely asses the character activation performance. However, since the performance between session 1 and 2 was significantly different ($p<0.01$) but between session 2 and 3 was not ($p>0.05$), we can assume that results of the 2nd and 3rd session are reliable for building the CATPM.

The results of this model can be used to optimally arrange characters for the ambiguous tongue-keypad, and to increase the typing throughput by using the different optimization criteria. This will be useful in order to make it a more efficient computer input device for disabled individuals.

Acknowledgment

The authors thank Javier San Agustin and John P. Hansen, from IT University, Copenhagen for the valuable discussions. Also, special thanks to Henrik V. Christensen for technical help and Bo Bentsen for Clinical assessment during the realization of the experiments.

References

[1] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement." *J Exp Psychol Gen,* vol. 121, pp. 262-269, 1954.

[2] R. Plamondon and A. M. Alimi, "Speed/accuracy trade-offs in target-directed movements," *Behav. Brain Sci. ; Behav. Brain Sci.,* vol. 20, pp. 279-349, 1997.

[3] R. J. Jagacinski, D. W. Repperger, S. L. Ward and M. S. Moran, "A test of Fitts' law with moving targets," *Hum. Factors; Hum. Factors,* vol. 22, pp. 225-233, 1980.

[4] E. R. Hoffmann, "Capture of moving targets: A modification of Fitts' Law," *Ergonomics; Ergonomics,* vol. 34, pp. 211-220, 1991.

[5] E. R. Hoffmann, "Fitts' Law with transmission delay," *Ergonomics; Ergonomics,* vol. 35, pp. 37-48, 1992.

[6] S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Hum. -Comput. Interact. ; Hum Comput Interact,* vol. 7, pp. 91-139, 1992.

[7] T. O. Kvalseth, "An alternative to Fitts' law (1980)," *Bulletin of the Psychonomic Society,* vol. 16, pp. 371-373, 1980.

[8] H. A. Caltenco, S. Boudreau, R. Lontis, B. Bentsen and L. N. S. Andreasen-Struijk, "Learning to type with the tip of the tongue: A performance study for a tongue-computer interface," in *2nd Annual IEEE Student Paper Conference,* 2008, pp. 5.

[9] E. R. Lontis, H. A. Caltenco, B. Bentsen, H. V. Christensen, M. E. Lund and L. N. S. Andreasen Struijk, "Inductive pointing device for tongue control system for computers and assistive devices," in *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 2009,

[10] R. Plamondon, "A kinematic theory of rapid human movements - Part II. Movement time and control," *Biol. Cybern. ; Biol. Cybern.,* vol. 72, pp. 309-320, 1995.