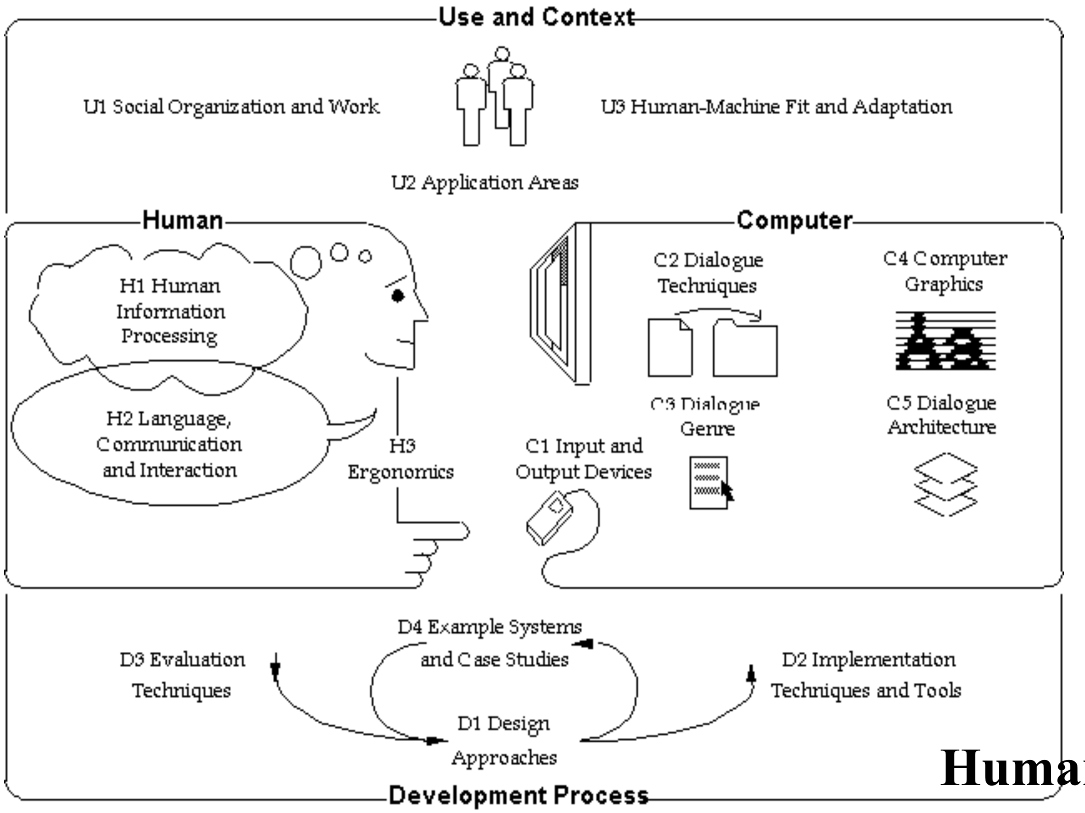
Syllabus

Part	Content	Week/Date	
Introduction	UI Revolution	1,2/Sep16，23	HCI的价值取向、历史，研究方法
	Text-Entry; NUI project assignment	3/Sep30	
Human Ability	Ergonomics	4/Oct07	HCI以人的性能为优化目标，了解人的性能
	Perception; Cognition	5/Oct14	
	HIP Model; NUI project Proposal	6,7/Oct21，28	
GUI	Fitts' Law study assignment	8/Nov04	GUI是主流界面模式，以交互效率为优化目标；**实验理解**优化模型
	GUI Design	9/Nov11	
	Fitts' Law Study/NUI project Checkpoint1	10/Nov18	
	Hick's Law, KLM	11/Nov25	
	Evaluation Methods	12/Dec02	
NUI	AR, VR; Wearable; Smart Space /NUI project Checkpoint2	13/Dec09	NUI,**实验实现**自然交互技术
	VUI; Gesture; Hands free; Eyes-free NUI project report/demo	14,15/Dec16, 23	
	~~NUI discussion; NUI project report/demo~~	~~16/Dec30~~	

What's HCI? The Definition



Interdisciplinary

Theoretical basis (Psychology，Social Science, Ergonomics, Information Theory, Computer Science, Linguistics, etc.)

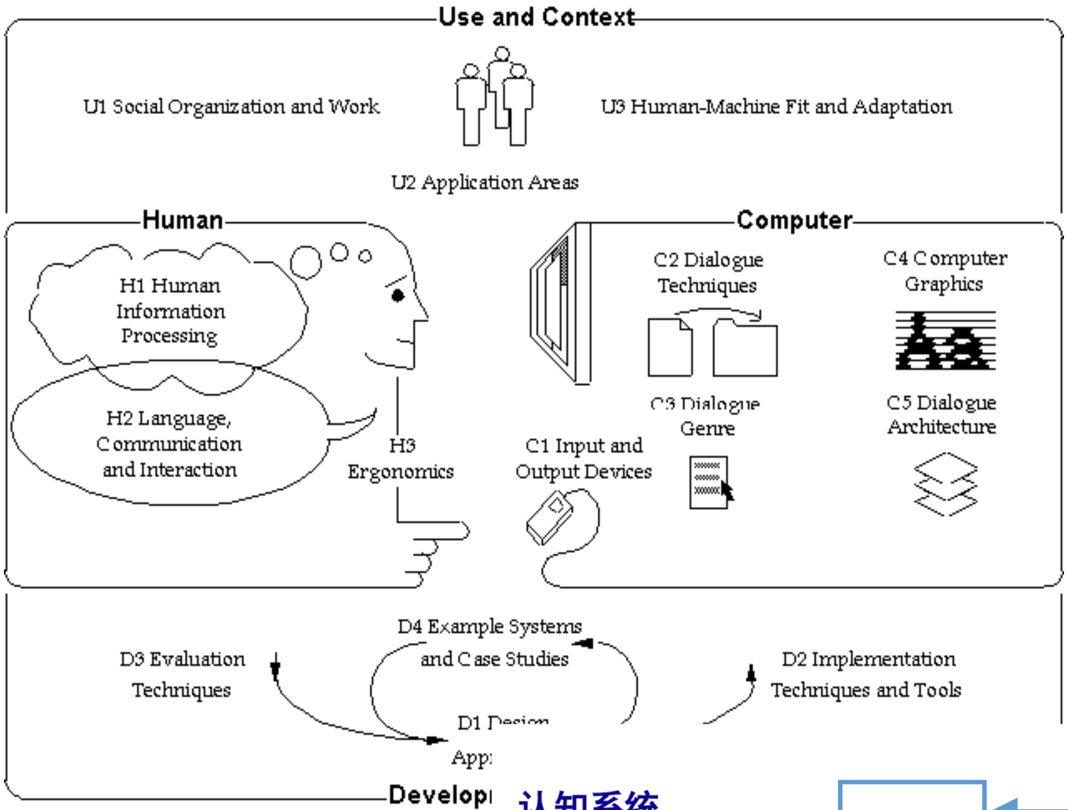
- **Interaction (Information exchange) between users and computers occurs at the User Interface (or simply interface)**

Human-computer interaction is:

"concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them"

——ACM SIGCHI, 1992

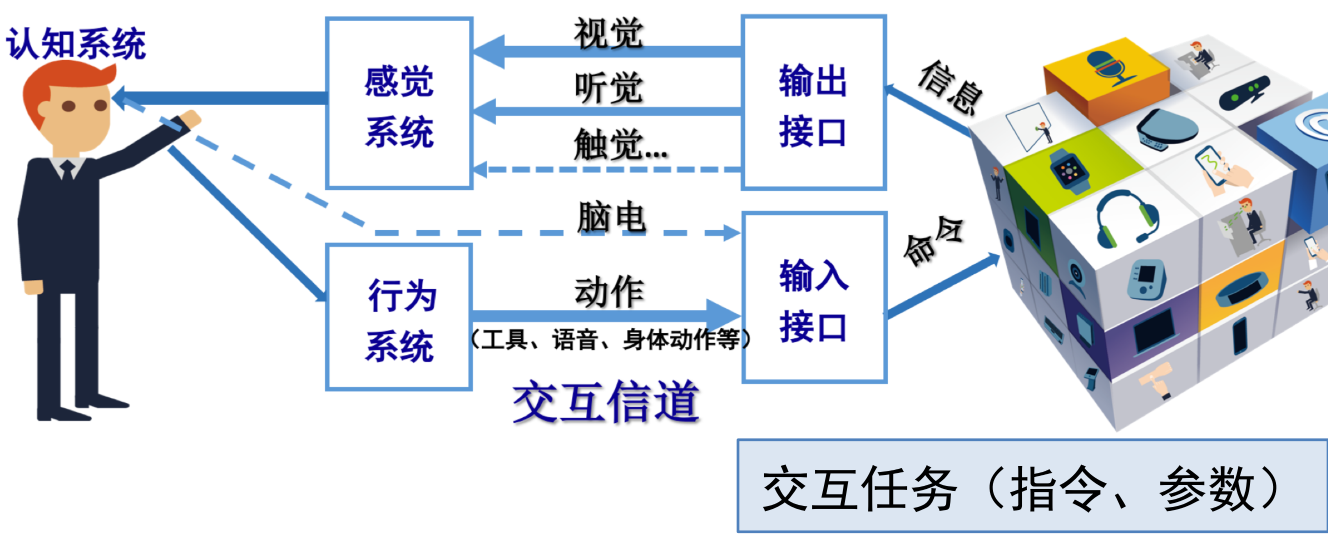
What's HCI? The Definition



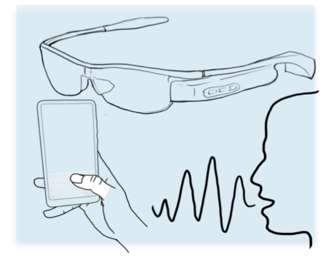
Interdisciplinary

Theoretical basis (Psychology，Social Science, Ergonomics, Information Theory, Computer Science, Linguistics, etc.)

- **Interaction (Inform**
  between users and
  at the User Interface
  interface)



交互技术VS计算模式

主机计算	个人计算	桌面计算	普适计算(移动、穿戴、嵌入)
			
批处理	命令行用户界面 CUI/Command	图形用户界面 GUI/Graphical	自然用户界面 NUI/Natural

言语 → 文本 → 键盘 → 文本 → 键盘 → 文本 → 虚拟键盘

动作 → 点击 → 鼠标 → 手势 → 动作感知

言语 → 语音 → 语音输入

无需专门输入设备

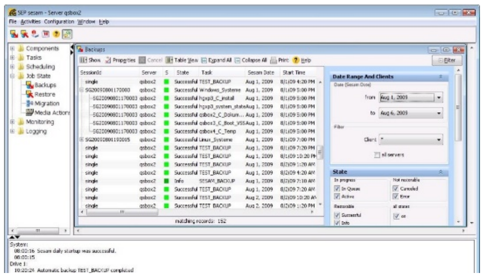
专业操作员 科学计算	→	专家用户 计算/控制	→	普通用户 桌面应用	→	大众用户 日常应用

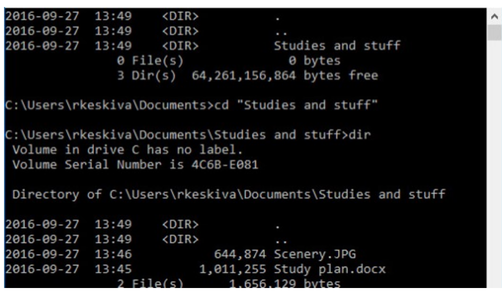
Text Entry 文本输入技术

- Background
- Design space of keyboard
- Bayesian Input prediction algorithm

Development of UI

Keyboard
/Mouse





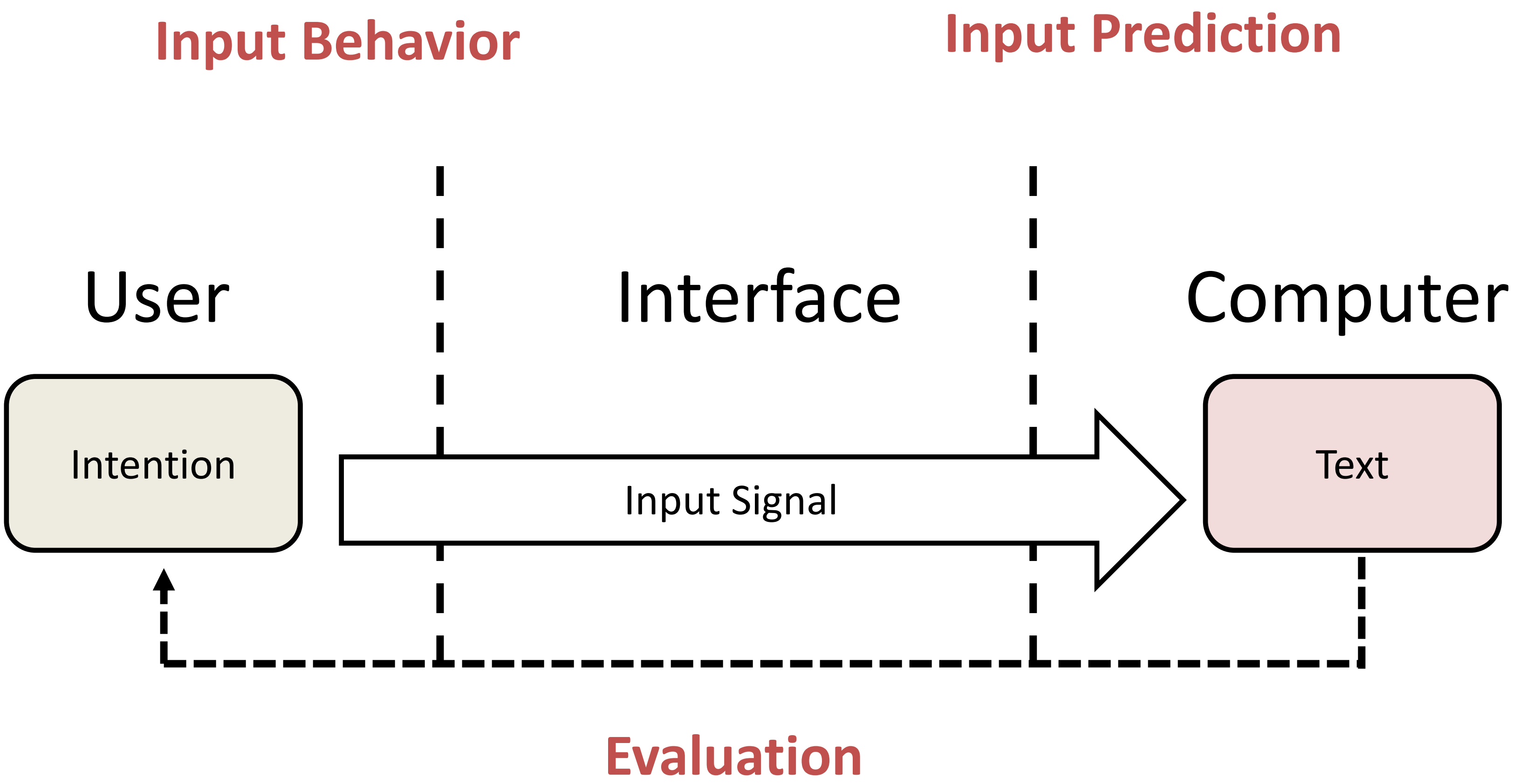
Touchscreen

Gesture

UI	Size	Tactile Feedback	Input Precision	Input Efficiency	Naturalness
CUI	Big	Yes			
GUI (KB+Mouse)	Big	Yes			
GUI (Touch)	Small	Partial			
3D UI	Big	No			

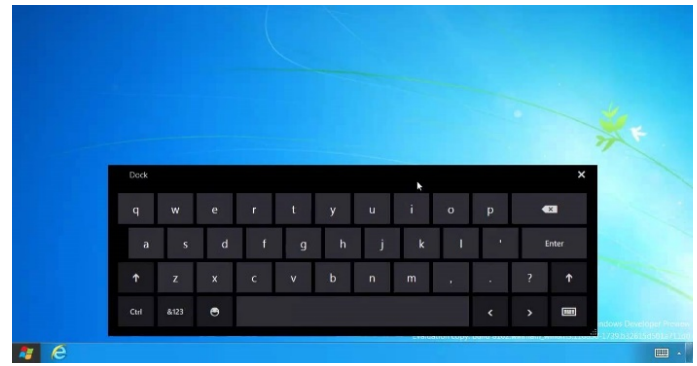
Problem： Tradeoff between input efficiency and naturalness

Text Entry Task



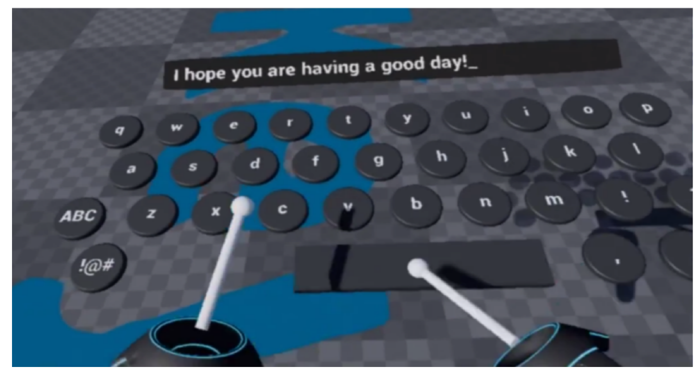
Development of Keyboard

Physical Keyboard（60-100 WPM）



Soft Keyboard (Mouse)（~10 WPM）

Touchscreen Soft Keyboard（30-40 WPM）



VR Keyboard（<10 WPM）

Key Feature： Input Prediction

Input Prediction

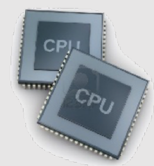
Physical Keyboard

$I_1 \quad I_2 \quad I_3 \quad I_4 \quad I_5 \quad I_6$

$W_1 \quad W_2 \quad W_3 \quad W_4 \quad W_5 \quad W_6$

$$text = signal$$

Smart Keyboard

$I \rightarrow$  $\rightarrow W_A, W_B, W_C, \cdots$

$$\underset{W}{\operatorname{argmax}} P(W|I) = \frac{\boxed{P(I|W)} \times \boxed{P(W)}}{P(I)}$$

Touch Model

Language Model

$$text = f^{-1}(signal, interface)$$

Goodman, J., Venolia, G., Steury, K., & Parker, C. (2002, January). Language modeling for soft keyboards. In *Proc. IUI'02* (pp. 194-195). ACM.

Interface: Smart Personal Devices

Phones

1.5 Billion

Tablets

0.3 Billion

Watches

0.3 Billion

Sales volume in 2015

Interface: Challenges

- Ten-finger input
- Tactical feedback
- Eyes-free Input
- Reasonable Key Size

60~100 word per minute (wpm)

vs.

- One- or Two-finger input
- Lack of tactile feedback
- Attention-engaged Input
- (Ultra-) small key size & “fat finger”

20~36 wpm for phones and tablets
Below 10 wpm for watches

Text Entry

- Background
- Design space of keyboard
- Bayesian Input prediction algorithm

Design space

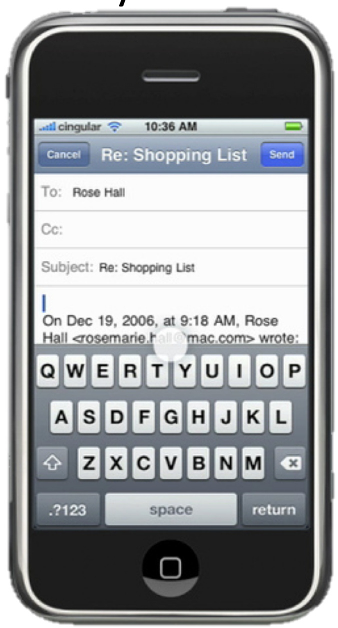
- Form Factor
- Keyboard layout
- Hand Posture
- Visual attention

Design space

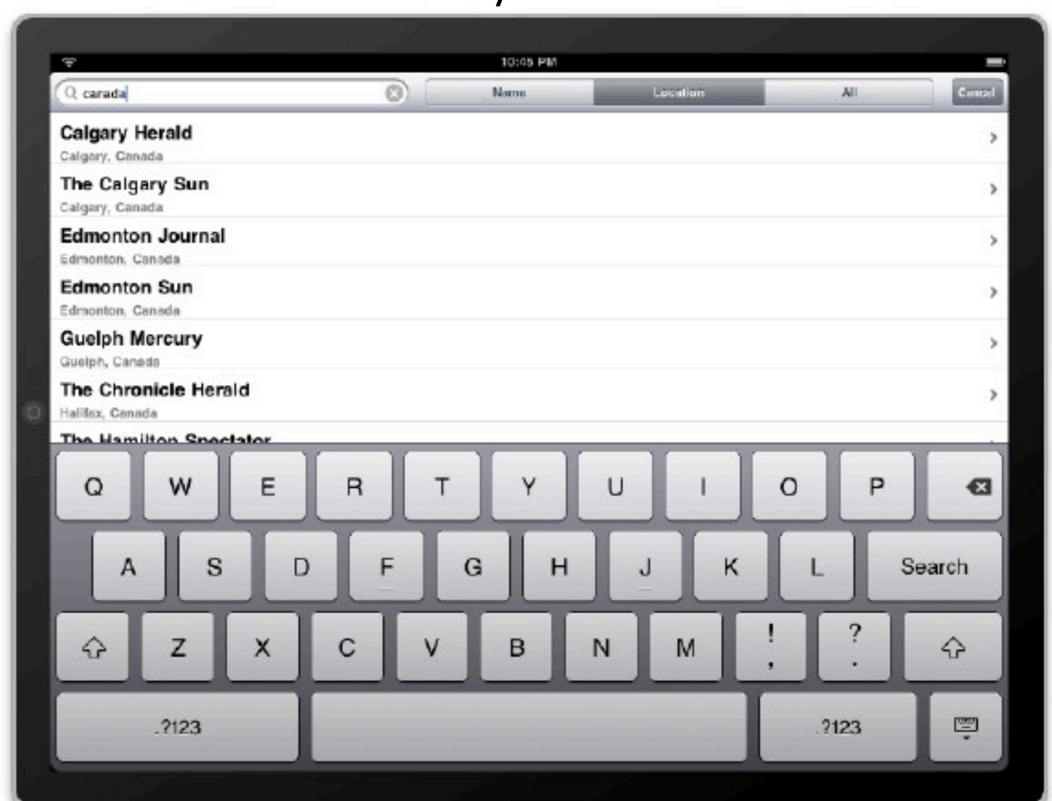
- **Form Factor**
- Keyboard Layout
- Hand Posture
- Visual Attention

1~2 inches
Key 1~2 mm

3~5 inches
Key 4~6mm



7~10 inches
Key 10mm



Design space

- Form Factor
- **Keyboard Layout**
- Hand Posture
- Visual Attention

QWERTY

Bi-Key

T9

Design space

- Form Factor
- Keyboard Layout
- **Hand Posture**
- Visual Attention

Ten finger

Index Finger

Single-handed Thumb

Two-handed Thumbs

Design space

- Form Factor
- Keyboard Layout
- Hand Posture
- **Visual Attention**

Eyes-engaged

Eyes-free

Design space

Form Factor

- Tablets 7~10”, Phones 3~5”, Watches 1~2”

Keyboard layout

- QWERTY, T9, 1Line, Bi-Key

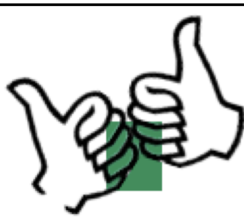
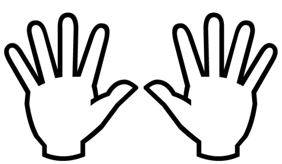
Hand Posture

- Single-handed Thumb, Index Finger, Two-handed Thumb, Ten finger

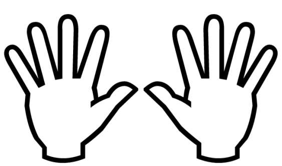
Visual attention

- Eyes-Engaged, Eyes-free

Hand Posture vs. Form Factor vs. Keyboard Layout

	7~10 inches tablet	3~5 inches phone	1~2 inches watch
	QWERTY	QWERTY, T9	QWERTY, T9, Bi-Key
		QWERTY, T9	
	QWERTY	QWERTY	
	QWERTY		

Hand Posture vs. Form Factor vs. Keyboard Layout

	7~10 inches tablet	3~5 inches phone	1~2 inches watch
			
			
	QWERTY		

Individual configurations are
- suitable for different scenarios
- of different input speed
- of different touch accuracy

Text Entry

- Background
- Design space of keyboard
- Bayesian Input prediction algorithm
- Keyboard techniques in VR

Problem Definition

- Input (Touch) Sequence: I

$$I = (x_1,y_1)\ (x_2,y_2)\ldots(x_n,y_n)$$

- Output word: W belongs to a word set $\{w_i \mid i=1\ldots N\}$

- Maximize the likelihood of W given I.

$$W = \max_{w} P(I, w)$$

Problem Definition (continue)

$$W = \max_{w} P(I, w)$$

Bayes' Rule

$$W = \max_{w} P(I|w) \cdot P(w)$$

Input/Touch Model

Language Model

Language Model

- A set of strings associated with its frequency
  - Indicate which words are authentic, and which are illegal or arbitrary string
  - Specify the likelihood of each word that user may input

Terms	Probability for Input
the	0.02
of	0.015
and	0.01
to	0.01
a	0.001
…	

$$\sum_{terms} P(term) = 1$$

Character Input vs. Word Input

- Input a character. Take 26 letters for example:

$$\text{Log}_2(26)=4.7 \text{ bits}$$

- Input a word. Take 10K vocabulary for example:

$$\text{Log}_2(10K)=13.3 \text{ bits}$$

- In average, word length is 5. The gap is

$$4.7*5- 13.3 = 10.2 \text{ bits}$$

- In other words, in the default keyboard (without any intelligence), we input 10.2 bits (77%) more for each word.

Key Grouping

- Increase key size to make touch easier
- Resolve ambiguity with language model

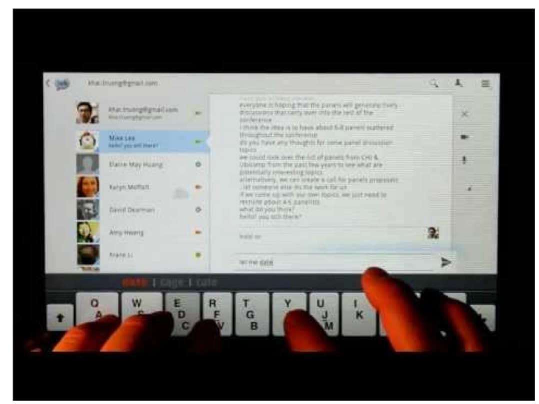
T9

Phones and watches

Samsung’s Bi-Key

Watches

1Line Keyboard



Tablets

Take T9 for example

- Word: “ON”
- Input: (6MNO)->(6MNO)
- Candidates: “NO”, “ON”

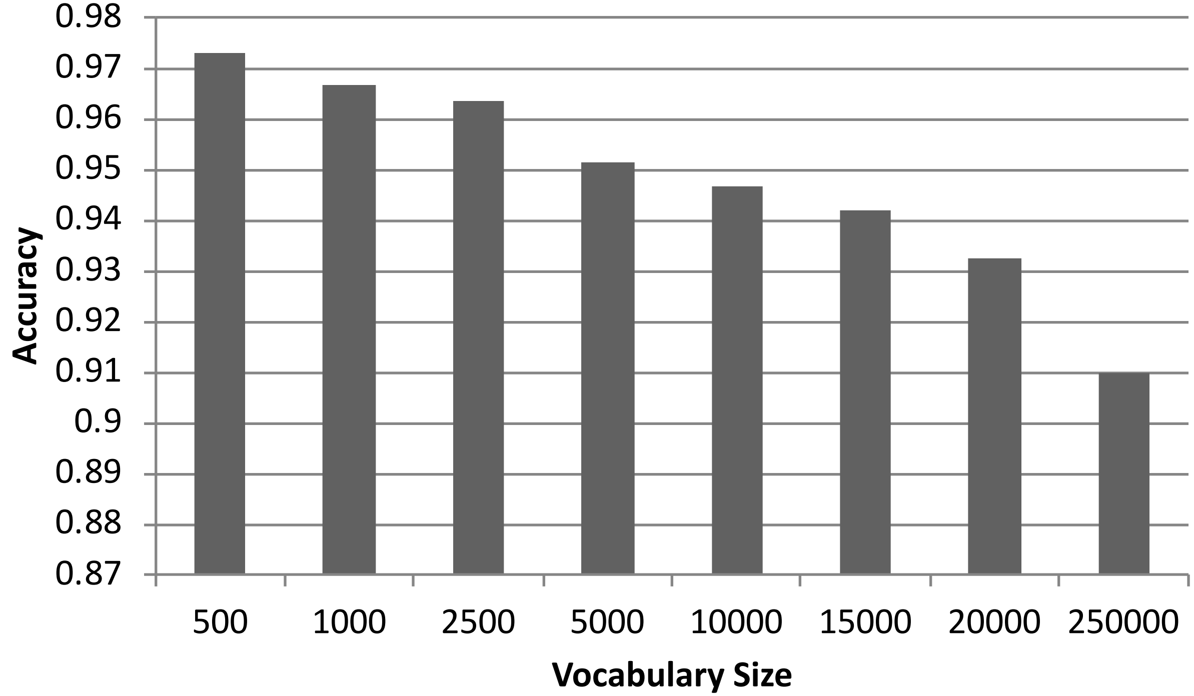
- Word: “THE”
- Input: (8TUV)->(4GHI)->(3DEF)
- Candidates: “THE”, “TIE”

- Word: “MAKE”
- Input: (6MNO)->(2ABC)->(5JKL)->(3DEF)
- Candidates: “MAKE”, “MADE”, “MALE”

Simulated Accuracy

Simulated Accuracy

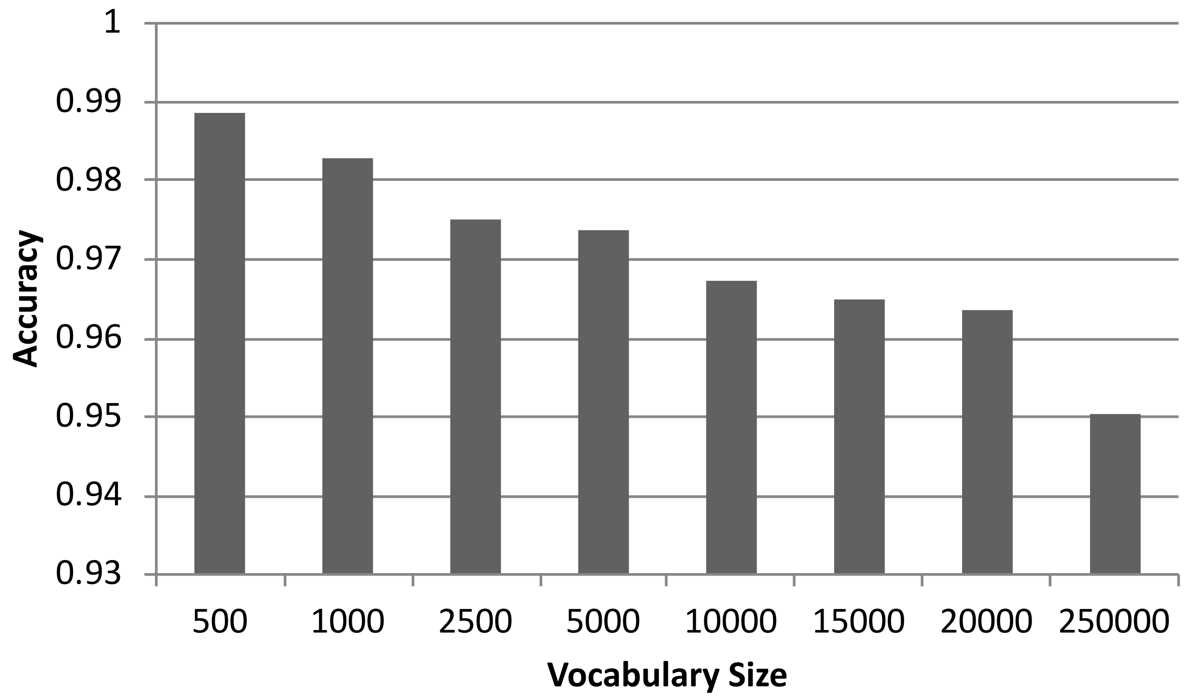
8 Groups



Probability of target word in the 3-best list

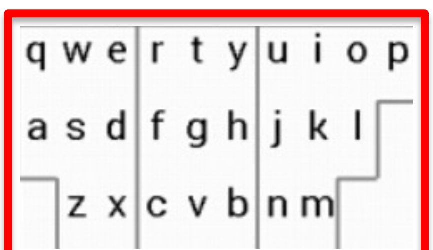
Simulated Accuracy

13 Groups

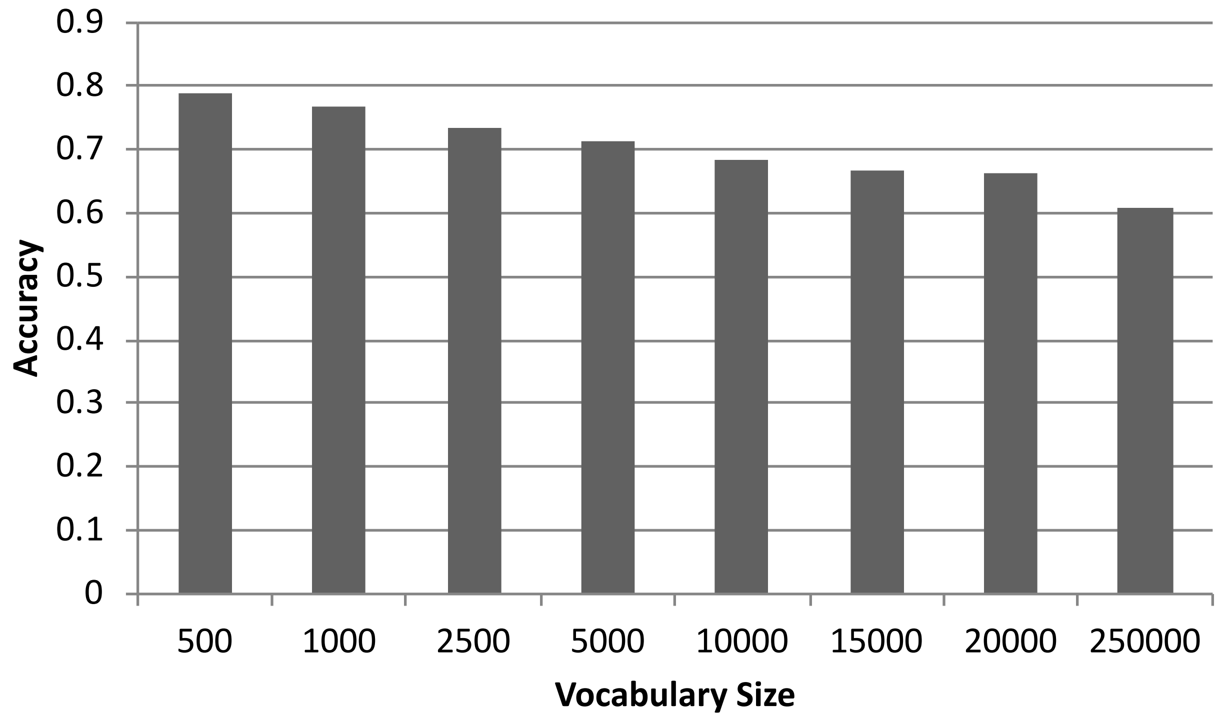


Probability of target word in the 3-best list

Simulated Accuracy



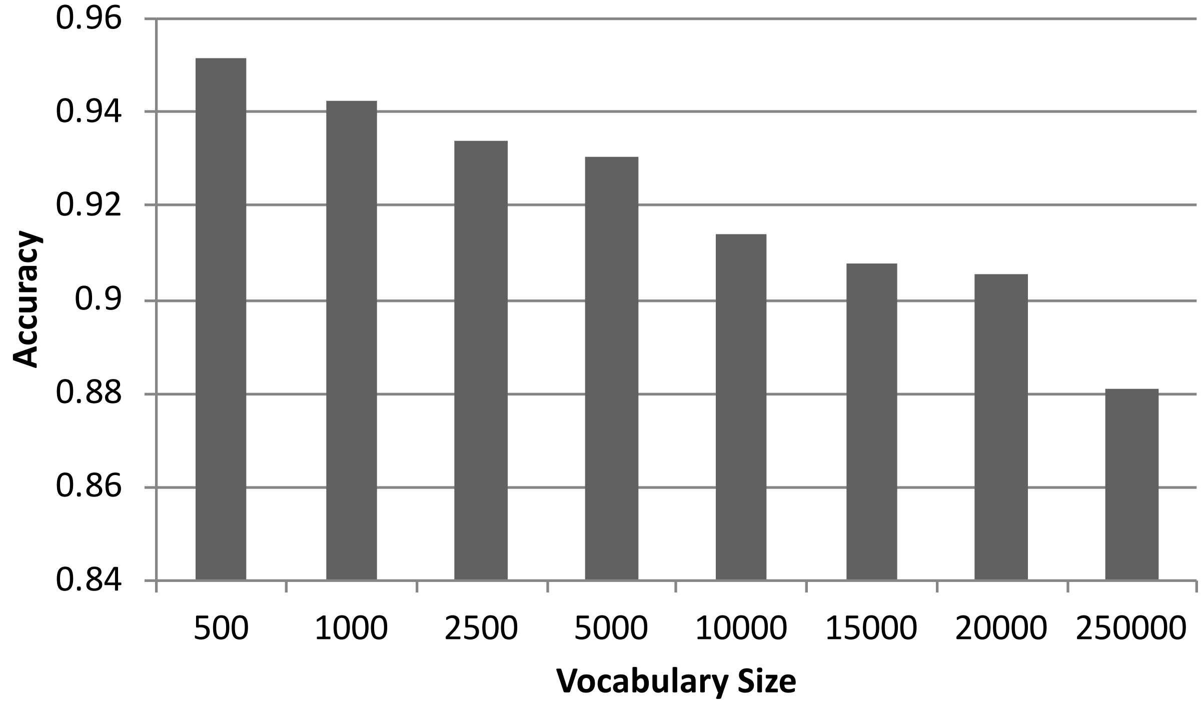
3 Groups



Probability of target word in the 3-best list

Simulated Accuracy

8 Groups



Probability of target word in the 3-best list

Touch Model

$$W = \max_{w} P(I, w)$$

$$\Downarrow$$

$$W = \max_{w} P(I|w) \cdot P(w)$$

Input/Touch Model **Language Model**

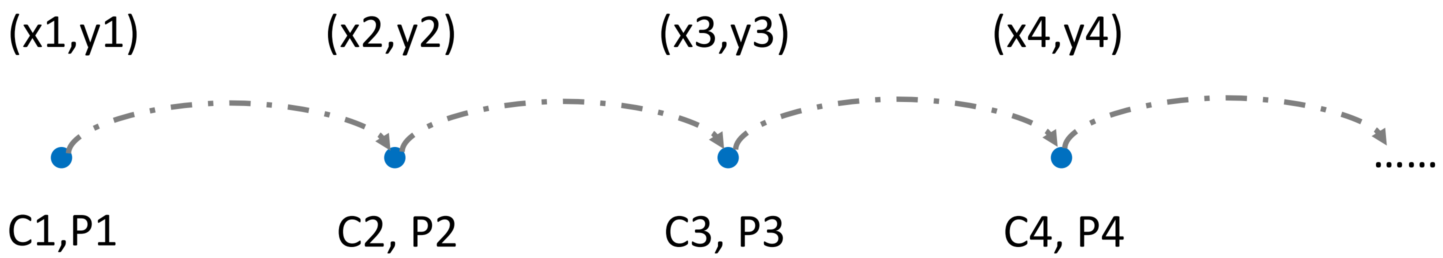
Touch Model: Independent vs. Dependent

- The sequence of characters in target word is

  C1, C2, C3…Cn

- The sequence of touch points is

  P1, P2, P3, …Pn



Independent Touch Model

- Maximize the likelihood of W given I.

$$W = \max_{w} P(I, w)$$

$$W = \max_{w} P(I|w)P(w)$$

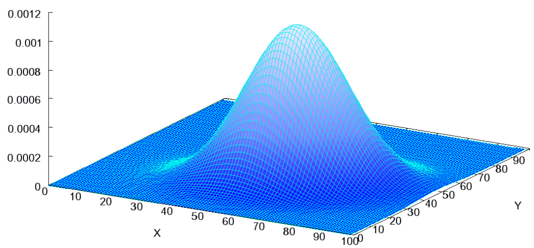
$$\Downarrow$$

$$W = \max_{w} P(w) \prod_{i=1}^{n} P(x_i, y_i | c_i)$$

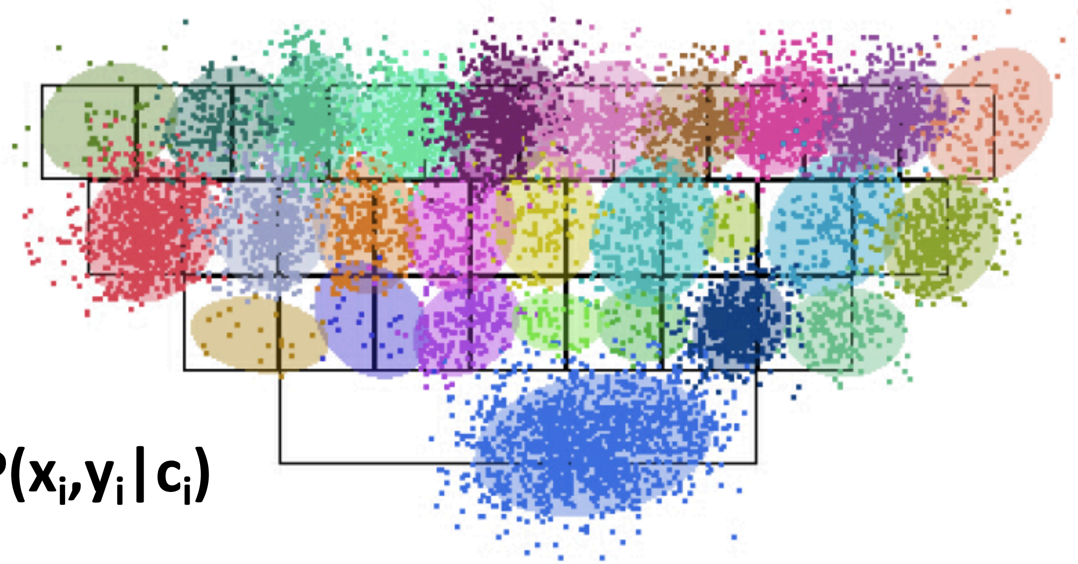
Independent Touch Model

- Hypothesized Distribution (e.g. Normal Distribution)

Average Error Rate for individual key is above 50%



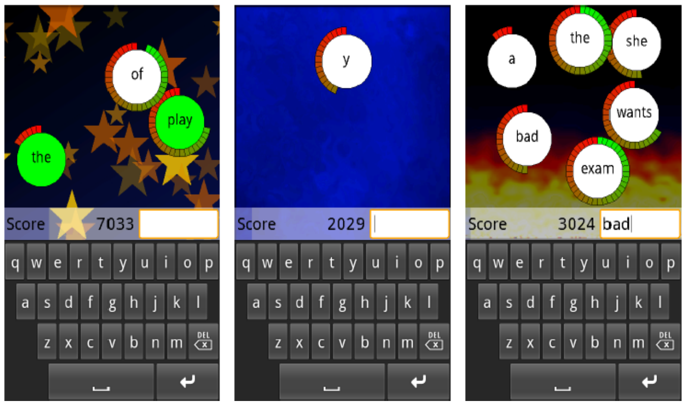
$$P(x_i, y_i | c_i)$$



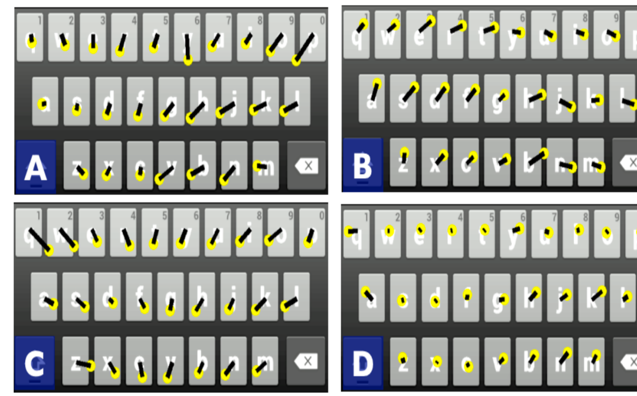
Touch distribution on smart watch (key size = 2mm)

Research on Touch Model

Typing patterns

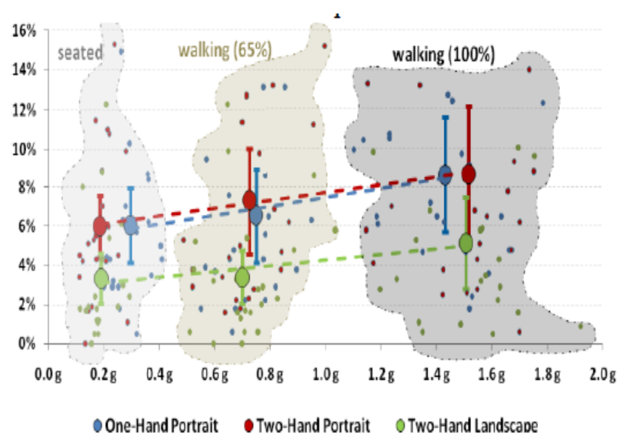


CHI'12: Touch in Game



CHI'13: ContextType

Effect of Mobility

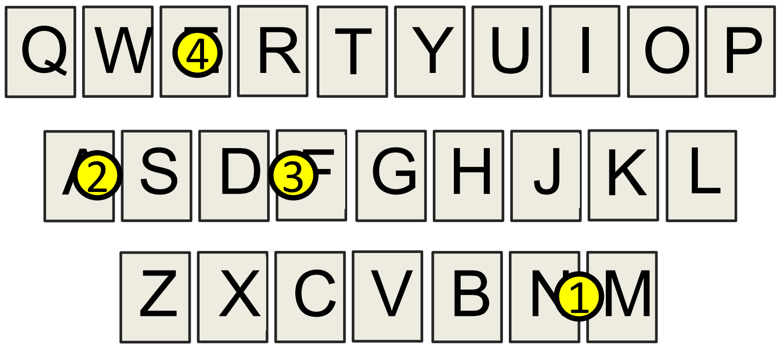


CHI'12: Thumbs Typing

CHI'12 WalkType

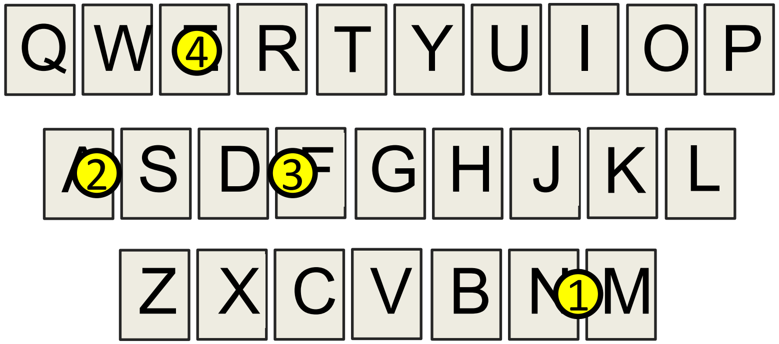
A Concrete Example

- Suppose a user needs to input “made”
- The produced four touch points are
  - P1 (350, 175)
  - P2 (69, 104)
  - P3 (181, 103)
  - P4 (123, 30)
- The uncorrected output is “nafe”



A Concrete Example

- Suppose a user needs to input “made”
- The produced four touch points are
  - P1 (350, 175)
  - P2 (69, 104)
  - P3 (181, 103)
  - P4 (123, 30)
- The uncorrected output is “nafe”



The keyboard algorithm works like this:

For each word w in vocabulary, we compute $P(w)\prod_{i=1}^{n} P(x_i, y_i | c_i)$

and then recommend the output according to the results.

A Concrete Example

$$W = \max_{w} P(w) \prod_{i=1}^{n} P(x_i, y_i | c_i)$$

For “name”:

P(P1,P2,P3,P4|“name”)=P(P1|”n”) P(P2|”a”) P(P3|”m”) P(P4|”e”) $=0.15*0.20*0.0001*0.25=7.5*10^{-7}$

P(“name”)=0.0017

P(P1,P2,P3,P4, “name”)=$0.0017*7.5*10^{-7}=1.275*10^{-9}$

For “made”:

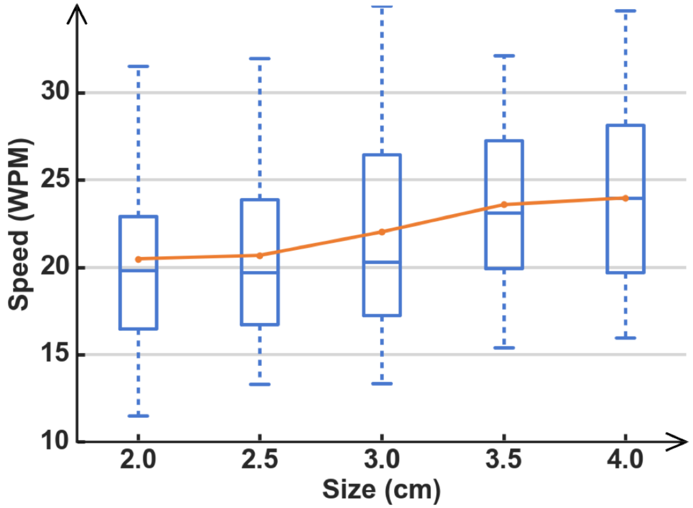
P(P1,P2,P3,P4|“made”)=P(P1|”m”) P(P2|”a”) P(P3|”d”) P(P4|”e”) $=0.08*0.20*0.02*0.25=8*10^{-5}$

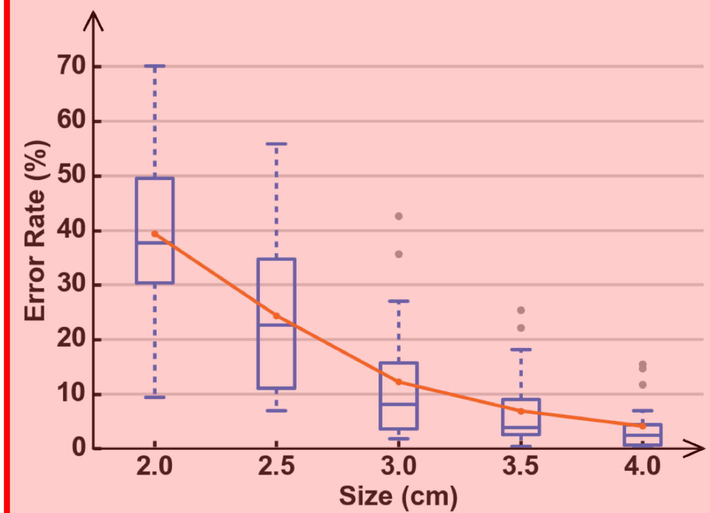
P(“made”)=0.00022

P(P1,P2,P3,P4, “made”)=$0.00022*8*10^{-5}=1.76*10^{-8}$

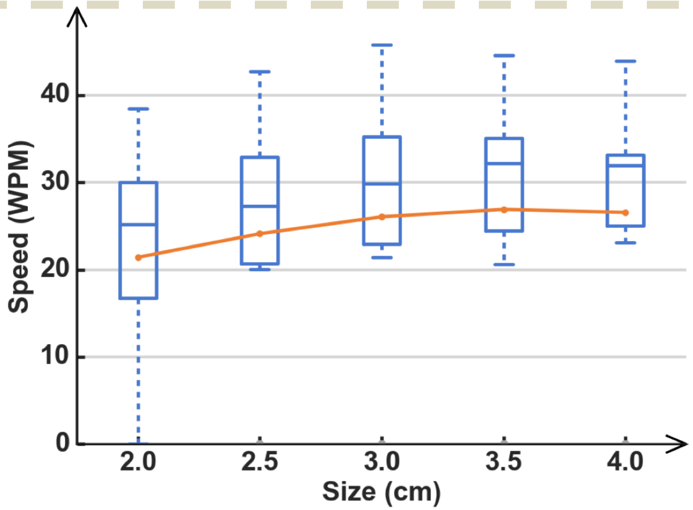
On Smartwatch Keyboards

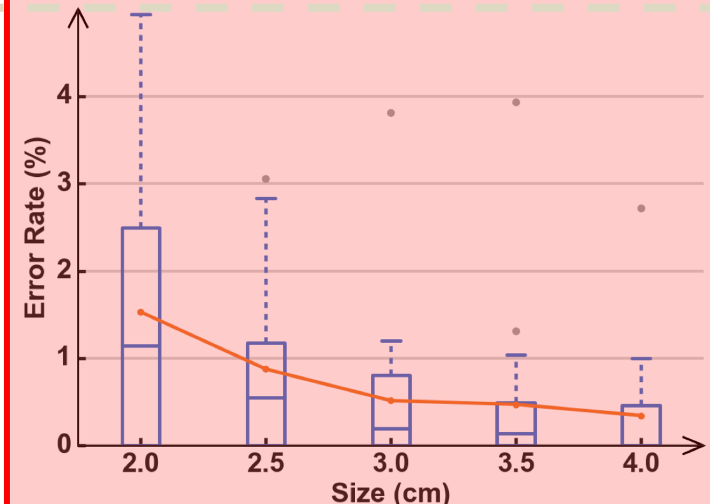
No Prediction





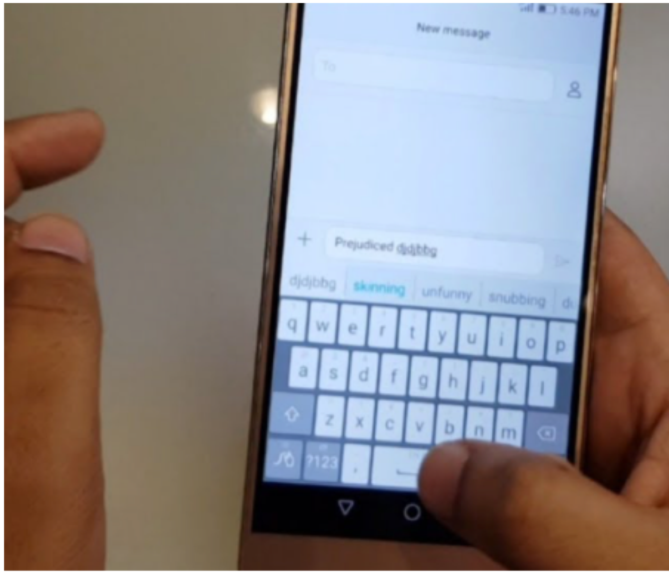
With Prediction

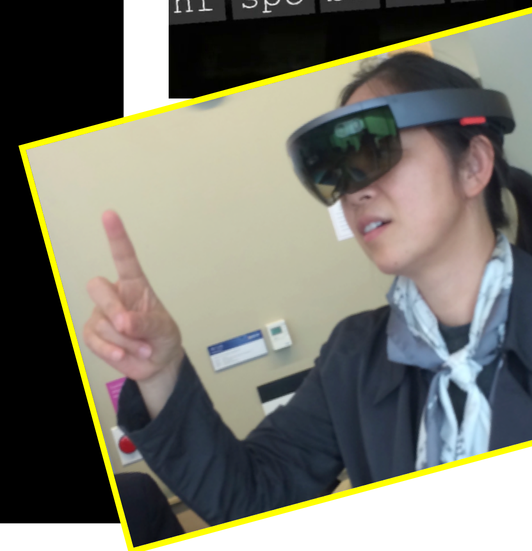




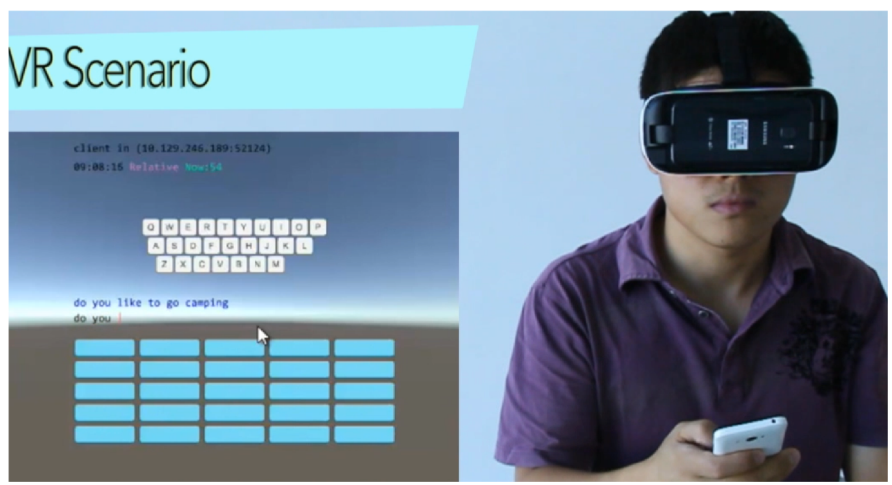

RESET
NEXT
Q W E R T Y U I O P
A S D F G H J K L
Z X C V B N M

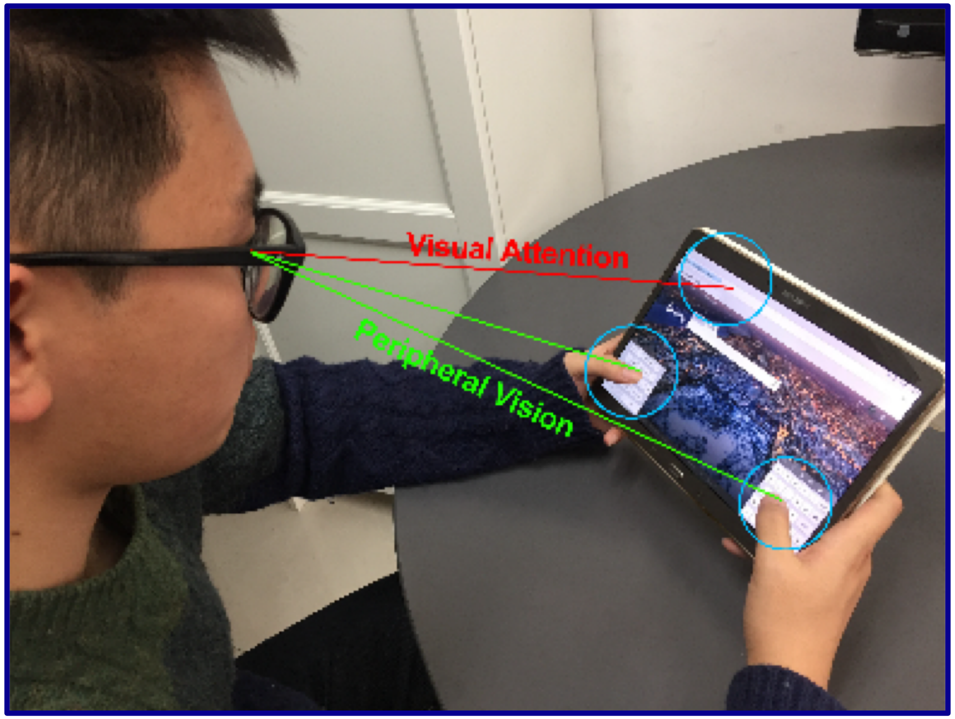

短信/彩信
我 你 不 好 在 嗯
Q W E R T Y U I O P
123
智慧輸入法

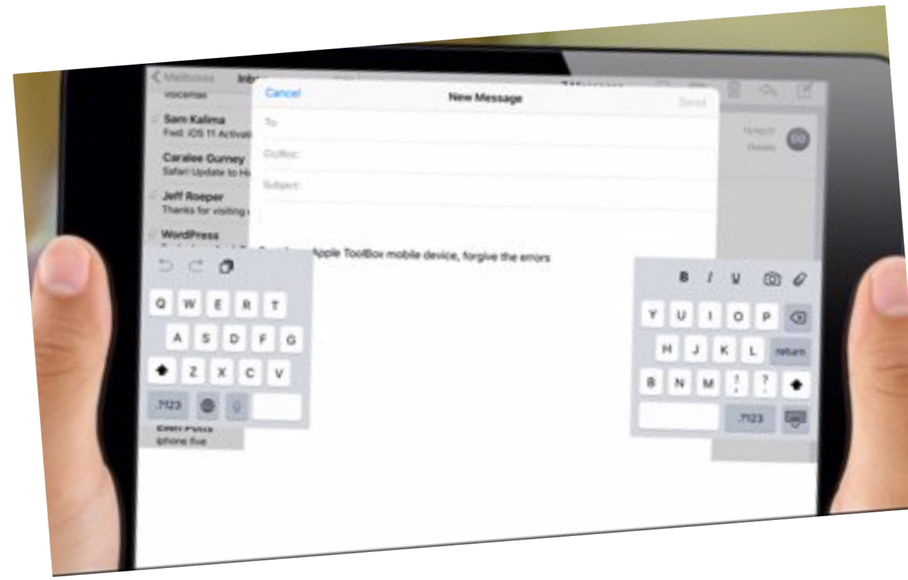

New message
Prejudiced djdjbbg
djdjbbg skinning unfunny snubbing

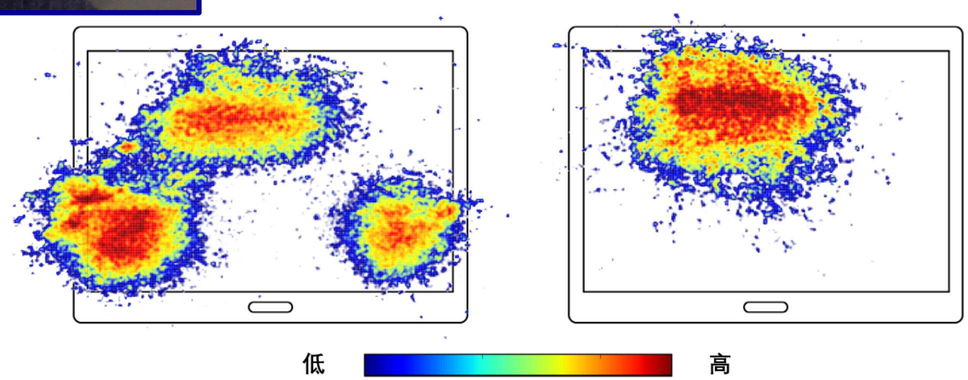


Eyes-free Typing in VR/TV




Visual Attention
Peripheral Vision


Cancel
New Message


低
高