Assessing the difficulty of the text input task for minority languages

Hugh Paterson III

Hello World!
Abstract

Currently no framework exists to evaluate or rank the complexity of the text input task on a per orthography bases. We present on the challenges which must be addressed by a cross-language text input assessment framework. We discuss relevant user experience (UX) considerations for keyboard layouts and unique actions undertaken in the communicative act of ‘entextualizing’ language (typing). We follow previous work which focuses on majority language text input methods (Bellman & MacKenzie 1998; Castellucci & MacKenzie 2013; MacKenzie 1992, 2002, 2007; MacKenzie & Soukoreff 2002; Soukoreff & MacKenzie 2001, 2003a, b) and apply considerations for minority language orthographies - especially those orthographies which overtly mark tone and other distinctions via diacritics.

The ability to communicate with electronic text based devices is important in this era of globalization. Many minority language users often find it difficult to type in their languages because of the way that orthography/language specific characters are accessed through existing keyboard layouts ([Author] 2012). The keyboard layout is an essential component in text input both on mobile touch screen and traditional devices. Barriers to efficiently using text in digital mediums has a wide impact on language vitality, by affecting the way that language users perceive their language's viability in the 21st century context. The text input challenge has been often acknowledged by minority language users (Esizmetor 2009; 13; Zheltov 2005). Perceptions about the need to be able to use text based digital communication devices has sufficiently challenged language communities leading some to change their orthographies. Simons and Lewis (2010) describe the social practice of literacy (EGIDS levels four and five) as a sign of a healthy language. A text input device which does not intuitively work for language users can be seen as discriminating and be a reason for speakers to choose to not use their language in digital mediums (Trosterud 2012). We propose a language agnostic framework for text input analysis for the benefit of language development efforts and software developers alike.

References:

Text input

Basically I mean typing!

• Is not orthography
• Is complex
• Happens in a variety of environments - various combinations of devices and sociolinguistic and socio-technological settings.
Typing happens in different places

With different tools.

Multi Finger Keyboard

Single Character Key
Single Finger Keyboard  

Multi Character Keyboard
Text input

• And then we want the text to look like different things...
Sample Me'phaa Text

A nguin', tsåan' ninimba'la' juyaá Jesús, gajuma'la' rí phú gagi juwala' ídó rí nanújngala' awúun mba'a inii gajmá. Numuu nduuyaá mála' rí ídó rí nanga'la' inuu gajmá, nasngajma ne rí gakon rí jañii akian'la' juyaá Ana'ló', jamí na'ne ne rí mawajún gukuála'. Índoó má' gi'maa rí mawajún gukuála' xúgíí mbi'í, kajngó majráan'la' jamí ma'ne rí jañii akian'la', asndo rí náxá'yóo nitháan rí ja'yoo manindxa'la'. [lyii' rí ni' tháán Santiágo 1:2-4]
Sample Sochiapam Chinantec Text

Hnoh² reh², ma³hiún¹³ hnoh² honh² lî³ua³ cáun² hi³ quiunh³² náh², qui¹ la³ cun³ hi³ má²ca³lî³ ňíh¹ hnoh² jáun² hi³ tî³ jlánh¹ bíh¹ re² hi²tîn² tsú² hi³ jmu³ juenh² tsí³, má³juáh³ zia³² hi³ cá² lau²³ ca³tî²¹ hi³ taunh³² tsú² jáun² ta²¹. Hi³ jáun² né³, chá¹ hnoh² cáun² honh², hi³ jáun² hi³ tî³ tîn² hnoh² re² hi³ jmúh¹³ náh² juenh² honh², hi³ jáun² hnoh² hi³ tî³ tîn³ náh² tsá² má²hún¹ tsí³, tsá² má²ca³hiá² ca³táunh³ ca³la³ tán¹ hián² cu³tî³, la³ cun³ tsá² tiá² hi³ lî³hniauh²³ hi¹ cáun² ňí¹ con² yáh³. [Jacobo Jmu² Cáun² Sí² Hi³ Ca³tîn¹ Tsá² *Judíos, Tsá² Má²tiáunh¹ Ñí¹ Hliáun³ 1:2-4]
So our fingers dance different dances

Me'phaa
Sample Text
So our fingers dance different dances
But dancing is work... right?
But dancing is work... right?

Hit load, hand balancing, and finger balancing

Sochiapam
Chinantec
Sample Text
And we hope to do different things by using text...

Filling various social communicative functions
Assumptions about success

- allow the digital text input of an orthography
- allow typing a text without fatigue
- maximize typing speed
- reduce the number of typing errors
- allow rapid mastery of the touch typing method
Conceptualizing the problem space

• Distance - Time equation
  - The shortest distance between two points is a straight line.
Building on what?

- Model 1: Counting chorded keystrokes as a single keystroke (Constable 2001)
  - Does every reach of the fingers count or do only the produced characters count?

- Model 2: Not counting shifted characters at all (Shieh & Lin 1999)
  - Does every character count including punctuation?
Building on what?

• Model 3: Single language/orthography - ENGLISH with assumptions like $M = N$, $M > N$, but not on the 10 key layout (Yin & Su 2011)
  • Do we find ourselves in monolingual communicative contexts or in multilingual contexts?
  • Do we find ourselves in contexts where we have more characters than keys?

• Model 4: Not defining a ‘character’ Mackenzie et al. (in general) -
  • easy: ø
  • Difficult: ó, ò, ô, ø / ø, ó / ó , ô / ô
Why do we need a comparative framework?

• Can we just change the orthography to fit the keyboard?
  

• We know that typing is hard, but how hard (or what is hard about it)?
  
  Practically speaking minority language users come to us with a “problem” but are we addressing it with the appropriate solution? How much investigative effort do we put into the text input side of the “literature development” problem space?

• What do we compare?
What to compare

Six criteria

• tapping load distribution
• number of keystrokes
• hand alternation
• finger alternation
• finger posture
• hit direction (little finger to thumb)

Keyboarding Typology

• Single Character Key example - QWERTY
• Single finger keyboard
• Multiple Character Keyboard - T9 phone
• Multiple finger Keyboard
What else is helpful to compare?

- Perceptual distance
- Measuring dissonance
Dissonance

- What does the user see and have to ignore so that they can achieve the results they desire?
Perceptual distance

• The distance the user casually thinks they will travel (or the time taken) to achieve the result
References


Spanish
Sample Text
Spanish Sample Text

Spanish Full Text
English - QWERTY
Sample Text