African Languages: Assessing the text input difficulty

Communication with electronic text based devices is prolific in this era of globalization. In many socio-cultural contexts the ability to input digital text undergirds sustainable social practice of literacy. 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), adversely affecting language development efforts. The challenge then is to provide a digital text input method (1) that works well for the orthography of any given language and (2) that users also find intuitive.

Paterson (2015) presents a framework to evaluate or rank the complexity of the text input task on a per orthography bases. We apply this framework and present the results from five Nigerian languages: Ezza [eza], Bekwarra [bkv], Cishingini [asg], Okphela [atg], and Igbo [ibo]. We discuss relevant user experience (UX) considerations for keyboard layouts and unique actions undertaken in the communicative act of 'encoding' 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.

Many minority language users often find it difficult to type in their languages because of the way that orthography specific and language specific characters are accessed through existing keyboard layouts (Paterson 2014). 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. Minority language users acknowledge the difficulty of text input in their languages (Esizmetor 2009: 13, Zheltov 2005).

Perceptions about the need for text based digital communication devices has sufficiently challenged language communities leading some to change their orthographies (South Pacific - Boerger 2007: 133; Central Asia - Cooper 2005: 149, 160; Americas - Jany 2010:235-6). However, it is unclear if the issue is the text input process or some other interaction between users and an orthography. The African context is not immune to these challenges (Bailey 2007). Africa is probably more sensitive to text input difficulties than other geographical regions like the South Pacific because many Roman Script African orthographies overtly mark tone and other distinctions via diacritics.

References:

Bailey, Dwayne. 2007. Creating a single South African keyboard layout to promote language. *Lexikos* 17.1: 212-25. http://www.ajol.info/index.php/lex/article/view/51533

Bellman, Tom & I. Scott MacKenzie. 1998. A Probabilistic Character Layout Strategy for Mobile Text Entry. *Proceedings of Graphics Interface '98*, 168-76. Toronto: Canadian Information Processing Society.

Boerger, Brenda H. 2007. Natqgu Literacy: Capturing Three Domains for Written Language Use. *Language Documentation & Conservation* 1.2: 126–53.

Castellucci, Steven J. & I. Scott MacKenzie. 2013. Gathering Text Entry Metrics on Android Devices. *Proceedings of the International Conference on Multimedia and Human-Computer Interaction - MHCI 2013*, 120.1-.8. Ottawa, Canada: International ASET, Inc.

Cooper, Gregory. 2005. *Issues in the Development of a Writing System for the Kalasha Language*. Ph.D dissertation, Macquarie University.

Esizmetor, David Oshorenoya. 2009. What Orthography for Naijá? Paper presented at Conference on Nigerian Pidgin, University of Ibadan, Nigeria.

Jany, Carmen. 2010. Orthography Design for Chuxnabán Mixe. *Language Documentation & Conservation* 4.1: 231-53.

Lewis, M. Paul & Gary F. Simons. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55.2: 103–20.

MacKenzie, I. Scott. 1992. Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction 7*, 91-139.

MacKenzie, I. Scott. 2002. Introduction to this special issue on text entry for mobile computing. *Human-Computer Interaction* 17.2-3: 141-5.

MacKenzie, I. Scott. 2007. Evaluation of text entry techniques. In I. Scott MacKenzie & Kumiko Tanaka-Ishii (eds.), *Text entry systems: Mobility, accessibility, universality*, 75-101. San Francisco, CA: Morgan Kaufmann.

MacKenzie, I. Scott & R. William Soukoreff. 2002. A Character-level Error Analysis Technique for Evaluating Text Entry Methods. *A character-level error analysis technique for evaluating text entry methods. Proceedings of the Second Nordic Conference on Human-Computer Interaction -- NordiCHI 2002*, 241-4. New York: ACM.

Paterson, Hugh J., III. 2015. Assessing the difficulty of the text input task for minority languages. Paper presented at the *4th International Conference on Language Documentation & Conservation,* Ala Moana Hotel in Honolulu, Hawai‘i, (February 26-March 1, 2015).

Paterson, Hugh J., III. 2014. Keyboard layouts: Lessons from the Me'phaa and Sochiapam Chinantec designs. In Mari C. Jones (ed.), *Endangered Languages and New Technologies*, 49-66. Cambridge, UK: Cambridge University Press.

Soukoreff, R. William & I. Scott MacKenzie. 2001. Measuring errors in text entry tasks: An application of the Levenshtein string distance statistic. *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems - CHI 2001*, 319-20. New York: ACM.

Soukoreff, R. William & I. Scott MacKenzie. 2003a. Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. Paper presented at *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, Florida, USA.

Soukoreff, R. William & I. Scott MacKenzie. 2003b. Input-based Language Modeling in the Design of High Performance Text Input Techniques. *Proceedings of Graphics Interface 2003* (CIPS, Canadian Human-Computer Communication Society), 89-96. Halifax, Nova Scotia: A K Peters.

Trosterud, Trond. 2012. A restricted freedom of choice: Linguistic diversity in the digital landscape. *Nordlyd (Tromsø University Working Papers on Language and Linguistics)* 39.2: 89-104.

Zheltov, Pavel V. 2005. Minority languages and computerization. The situation in the Russian Federation. *OGMIOS* 3.3: 8-11.