# Keyboard Layout Design for Minority Languages - (Socio)linguistic (app)/(Im)plications

BY

HUGH PATERSON III

B.A. IN GERMAN LITERATURE, SOUTHERN ILLINOIS UNIVERSITY EDWARDSVILLE, 2007

*Table of Contents*

Ezea Full Text
Cishinini Full Text
Okphela Full Text

## *List of figures*

3 so far not listed till transfer to XLingPaper.

## *List of tables*

Several not listed here. Not listed till compiled in XLingPaper.

## *Acknowledgments*

## *List of abbreviations*

**ANSI**: American National Standards Institute - a standards organization
**ASCII**: American Standard Code for Information Interchange
**C & dR**: Chelliah and de Reuse (the authors of the *Handbook of descriptive linguistic fieldwork*)
**CMC**: Computer-Mediated Communication
**EGIDS**: Expanded Graded Intergenerational Disruption Scale
**FOSS**: Free and Open Source Software
**GIDS**: Graded Intergenerational Disruption Scale
**HFA**: Hoffnung für Alle
**ISO**: International Standards Organization - a standards organization
**JIS**: Japanese Industrial Standards - a standards organization
**KL**: Keyboard Layout
**MSKLC**: Microsoft Keyboard Layout Creator
**NLT**: New Living Translation
**OS (OSes)**: Computer Operating System (Operating Systems)
**OS X**: Operating system ten (Apple, Inc.'s operating system)
**PUA**: Private Use Area
**RVR**: La Biblia Reina-Valera
**SUM**: Sustainable Use Model
**TBU**: Tone Bearing Unit
**UI**: User Interface
**USA**: United States of America
**UX**: User Experience

## *List of languages referenced and their abbreviations*
### Languages data used from these languages
#### *Languages of the Americas*[1]

**[eng]**: ISO 639-3 code for English.
**[cso]**: ISO 639-3 code for Sochiapam Chinantec (the variety of Chinantec used in the analyzed text for this paper).
**[nav]**: ISO 639-3 code for Navajo.
**[spa]**: ISO 639-3 code for Spanish.

**[tcf]**: ISO 639-3 code for Malinaltepec Me'phaa (the variety of Me'phaa used in the analyzed text for this paper).

#### *Languages of Africa*

**[asg]**: ISO 639-3 code for Cishingini.
**[atg]**: ISO 639-3 code for Okphela.
**[bkv]**: ISO 639-3 code for Bekwarra.
**[eza]**: ISO 639-3 code for Ezaa.
**[ibo]**: ISO 639-3 code for Igbo.

#### *Languages of Europe*

**[deu]**: ISO 639-3 code for German.
**[fra]**: ISO 639-3 code for French.
**[ita]**: ISO 639-3 code for Italian.

#### *Languages of Mongolia*

**[khk]**: ISO 639-3 code for Halh Mongolian.
**[rus]**: ISO 639-3 code for Russian.

### Languages mentioned briefly

**[ara]**: ISO 639-3 code for Arabic (macrolanguage).
**[arq]**: ISO 639-3 code for Algerian Arabic.
**[axk]**: ISO 639-3 code for Yaka.
**[bla]**: ISO 639-3 code for Blackfoot.
**[chr]**: ISO 639-3 code for Cherokee.
**[crd]**: ISO 639-3 code for Coeur d'Alene.
**[ell]**: ISO 639-3 code for Greek.
**[gla]**: ISO 639-3 code for East Sutherland Gaelic.
**[hop]**: ISO 639-3 code for Hopi.
**[kee]**: ISO 639-3 code for Keres.
**[kor]**: ISO 639-3 code for Korean.
**[lkt]**: ISO 639-3 code for Lakota.
**[lut]**: ISO 639-3 code Lushootseed.
**[one]**: ISO 639-3 code for Oneida.
**[pol]**: ISO 639-3 code for Polish.

---

[1] Some languages under the "Americas" are "European" in origin, but their socio-linguistic contexts lead them to be analyzed as languages of the Americas. Languages are only menetioned once even if they are compared in mulitible regions.

**[sag]**: ISO 639-3 code for Sango.
**[shu]**: ISO 639-3 code for Chadian Arabic.
**[swa]**: ISO 639-3 code for Swahili.
**[tew]**: ISO 639-3 code for Tewa.
**[tha]**: ISO 639-3 code for Thai.
**[tzm]**: ISO 639-3 code for Amazigh.
**[urd]**: ISO 639-3 code for Urdu.
**[ven]**: ISO 639-3 code for Venda.
**[yur]**: ISO 639-3 code for Yurok.

# *Abstract*

Keyboard layout design affects language vitality. Socio-technical systems are increasingly important in today's communication ecology (Whitworth & Ahmad 2013). Language development projects and language planing programs need a way to integrate linguistic knowledge, information, and transmission practices into socio-technical systems if the languages used in these systems are going to be the mother tongue languages of minority language speakers. With the current rate of technological adaption it is more than feasible that systems will become more relevant than the traditional literacy reading primer (Blench 2012: 15). This requires addressing the design tension between requirements for minority language users and the Human Computer Interaction (HCI) requirements of computing devices. The academic linguistic community often attempts to address these tensions at the orthography "design" level (Cooper 2005: 160, Jany 2010b: 235-6). However these "solutions" often revolve around removing diacritic marks from Roman script orthographies (Boerger 2007: 134) and do not address the marking of tone in languages, such as Chinantec (Foris 2000) and some African languages (Roberts 2011), where there is a significant need to mark tone. Such solutions also do not affect key frequency issues, or diacritic marks in Ajami and Indic scripts. This project focuses on the arrangement of keys on the keyboard, or keyboard layout (KL); proposing that KL's are the cornerstone to truly adapting the digital content creation process to the needs of minority language users. In the context of minority language text input design specifications and considerations, there has been relatively little published, either for the publishing industry, linguists, or for technologists (designers and programers). The one exception is an unfinished book released in draft form by SIL's foundry NRSI (Lyons 2001). In contrast to sparse literature supporting minority language text input, QWERTY keyboard interactions, primarily dealing with English, are well studied (MacKenzie 2002, 2007, 2013, MacKenzie & Tanaka-Ishii 2007). This current study takes current practice in the HCI literature and applies it to several minority language use cases, focusing on languages which use diacritics, often as a device used to explicitly mark tone in their orthographies.

Because a speaker's choice of language is based in both social and physical environments, orthography design decisions have an overall effect on the mechanics of language expression in digital forms. Emotional responses to design of orthography, of the computer operating system, and of the text input method bear upon the language user. In the discipline of language documentation and description, text input methods may initially be developed with the needs of the researcher in mind rather than the needs of a native speaker who uses the language in everyday interactions. These existing keyboard layouts that support specific languages are rarely used by the broader minority language community, and the efficacy of many keyboard layouts is limited to linguistic analysis or researcher convenience. Linguists often bring linguistic knowledge and some of their own user expectations to the keyboard design process. They may not realize that requiring a typist to negotiate a keyboard layout to access a given character can have an impact on language-use choice, orthography development, or adherence to an approved orthography. User-centric keyboard layout design for minority language community writers/typists should be an integral part of a language development project in the twenty-first century. These considerations bring us to the following question: *At what point in the design process should linguistic information be considered and applied, as opposed to other design criteria, so that maximal language usage is encouraged and made possible?* This study offers a framework for the linguist or language development worker to address crucial issues of keyboard layout design.

# 1. Writing, text-input, and typing with keyboards

The mechanical process of writing and the process of typing in digital contexts (text input) is of interest to those who study languages for at least three reasons. First, keyboarding is an expression of language and reveals some very unique ways that the human body expresses communicative thought. For instance, consider the ability to type 'LOL' without actually laughing or thinking "laugh out loud". These typing gestures can connect with our thoughts without activating the vocal or aural mechanisms which are often employed in the encoding and decoding of communicative thought. The study of the mechanics of writing is not new __(ADD early citation)__ and continues in more modern works__((van Galen 1991), (Peck, Askov & Fairchild 1980), A Decade of Research in Handwriting: Progress and Prospect)__, see__( computer v.s handwriting (Christensen 2004) teaching hand writing before typing to children (Stevenson & Just 2012))__ for discussion of contrasts between typing and handwriting; the interactive nature of computer-mediated communication (CMC) is new. Second, the language teacher is interested in language use in all mediums, including computer-mediated communication, not just oral communication. To be proficient in a language is to also know the cultural cues of when to say specific things and when to refrain. These communicative cultural norms, nonetheless carry over into communication occurring in digital mediums. The third reason that typing (text input) is of interest to those who study language is that the psychologist and linguist are interested in how the brain processes language through the process of writing, which includes typing (text input). This thesis will touch on various aspects of these three points as it proposes a framework for keyboard layout design.

Keyboard layout design is intrinsically interdisciplinary. To create a tool for language use which not only works but is embraced by a group of users requires an understanding of linguistic knowledge, script knowledge, and digital technology knowledge germane to the language entextulization challenge. For this new keyboard layout (analyzed as an object) to be embraced by a user group requires a successful application of principles from economics, anthropology, and design, especially user experience design.

The first section of this thesis takes the reader through the relevance of writing to linguistics and language development. It is often within this context that new keyboard layouts are created for monitory languages. The first section discusses entextulization and the process often followed in developing writing for the purposes of linguistic research, language documentation and language development. These settings are not without conflicting views surrounding writing as a part of language development. Just as writing is affected by various social practices and communal attitudes towards writing, so also the process of typing (text input) is affected by similar social constraints. That is, the need for writing, and therefore also the need for text input, is not felt ubiquitously.

Following the introduction to writing is a section discussing the current literature in the relevant to human computer interaction and keyboard interaction analysis. Academically, user experience design falls under the broader discipline of computer science, therefore much of the literature discussing text input (even in minority languages) does not occur in the linguistics or language documentation literature. Current literature concerning keyboard layout design, while not solely based on English language text input, is predominantly based on English language research. Furthermore, this research is rarely cited and apparently un-accessed (p.c. with various keyboard layout designers) by language development staff in the production of keyboard layouts. For these language development staff a far more pressing goal is the correct typesetting of professional documents, therefore the keyboard layout becomes a way to limit

(or quality control) data input options for text processing systems[2]. It is the goal of this thesis to integrate the HCI and the language development literatures so that the language development professional has a resource which references both literatures and provides that person with a framework upon which to design future keyboard layouts. The following diagram is a visualization of the various topics discussed in this thesis and their inter-relatedness.

The second section of this thesis __()__

The third part __()__

---

[2] The creation of keyboard layouts and text input systems is sometimes deligated to publishers (and typesetters and their foundries). These stakeholders in the publishing process are very interested in consistent encoding of texts. As an example some packages of LaTeX require the special declaration of combining glyphs to form characters and can not accept strait unicode characters (Goossens, Rahtz & Mittelbach 1997: 264-5). This more restricted approach to text input can be seen as a challenge for self publishers, who prefer a more straitforward approach to entextualization.
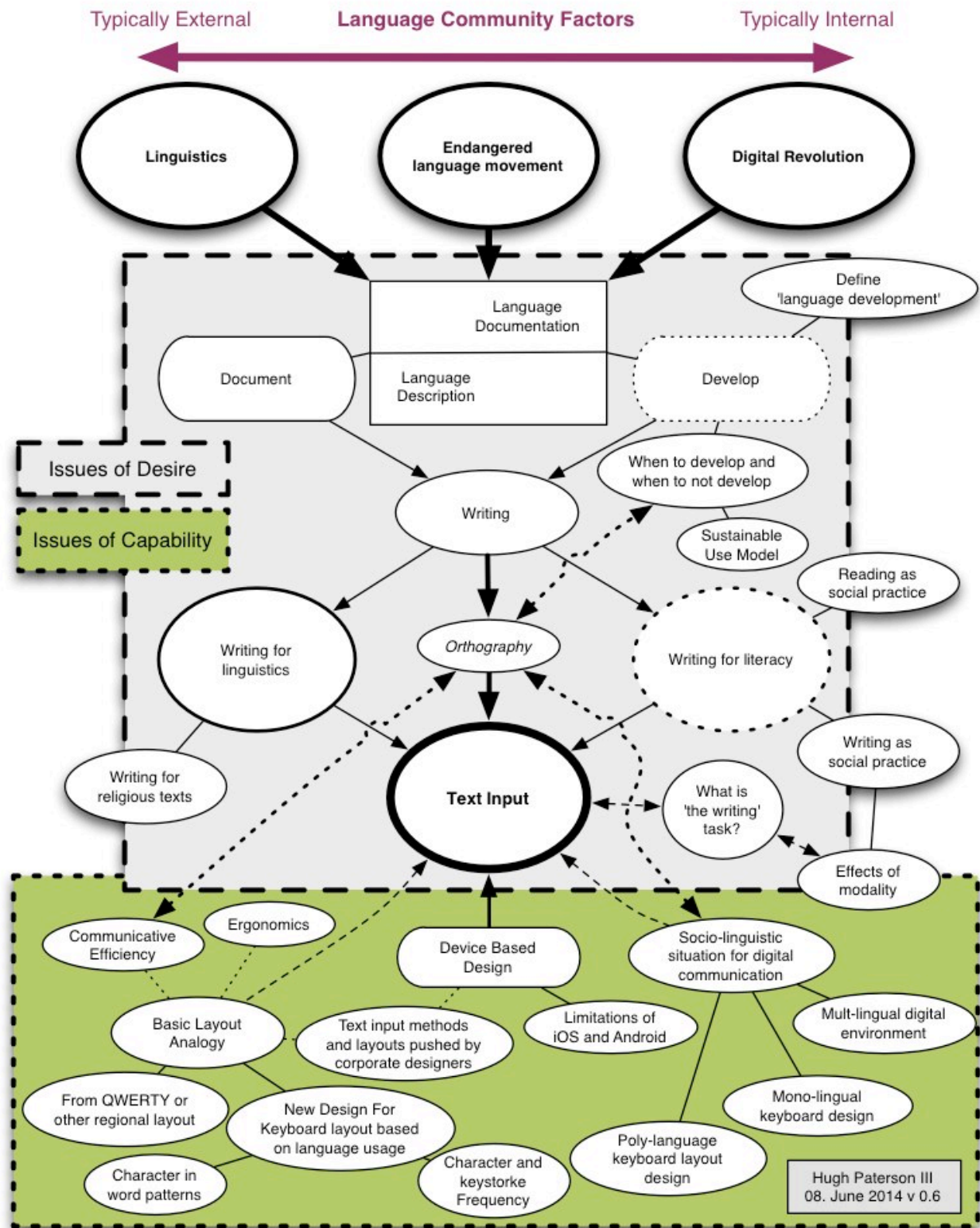
## *1.1. Language Documentation versus Language Description*

Since the early 1990's language documentation has emerged as its own discipline (Furbee 2010, Himmelmann 1998, Woodbury 2003) growing out of the field of linguistics. One of language documentation's distinctives is the collection of original language use performances (Nathan 2010) in digitally archivable formats (Bird & Simons 2003). The focus on primary data is a shift in paradigm (Thieberger & Musgrave 2007: 27-9) as much as it is in methods (Bergqvist 2012: 24). Language Documentation has focused on creating lasting and multi-purpose language artifacts, where as linguistics has traditionally chosen to focus on identifying the patterns occurring within and around language use. Along the way and facilitating the split between language documentation and linguistics, the field of linguistics has encountered two other notable movements: the endangered language movement, and the digital revolution. These movements have changed the the foundational evidence of linguistic argumentation from being an evidence based science argued from antidotal observations by linguists worried about the observers paradox (Labov 1966, 1972, 2006), and descriptions of languages based on written forms of observed of linguistic performance (for example hand transcribed Swadesh lists (Swadesh 1971: 283), to a science driven by data rich with reviewable examples of performance (Coleman 2011, Schroeter & Thieberger 2011, Thieberger 2009) gathered collaboratively by speakers and researchers (Dwyer 2006: 54-6, 2010, Kuhlmann 1992: esp. 277-278, Leonard & Haynes 2010, Penfield, et al. 2008).

## *1.2. The Digital Revolution*

The first of these two movements is the Digital Revolution. The advent of socially embraced digital communication has affected the behavior of both the observed (Kiesler, Siegel & McGuire 1984) and the observer (Crasborn 2010); the speaker and the listener (Seltzer, Prososki, Ziegler & Pollak 2012); the writer (Porter 2003) and the reader (Fortunati & Vincent 2014, Liu 2005, Mangen, Walgermo & Brønnick 2013). Digital devices are reshaping the communicative context in which 'language' exists. Handheld radios are replacing surrogate speech forms in Chinantec [cso] society (Wilfredo Flores, pc.; Mark Sicoli in segment 23:00-23:17 in D. Duncan 2013). Research in L1/L2 and L2/L1 transference, and the role of orthographies in the production of sounds suggests that devices with text based dependencies for operation stand to have the potential to expedite the reshaping of sounds in a minority language via the graphical similarity between a minority language orthography and the orthography of a language of wider communication (Detey & Nespoulous 2008, Major 2008: 69, Perre, Pattamadilok, Montant & Ziegler 2009, Simon, Chambless & Kickhöfel Alves 2010, Vendelin & Peperkamp 2006)[3]. The change of language use (including loss) is not the only impact digital devices are having on minority languages. In some contexts minority language speakers are either adapting language use habits to incorporate the use of digital devices (Lexander 2011) or adapting their language related products (orthographies) so that it can more readily be used on existing devices (Jany 2010b: 235). Digital tools not only allow for new methods of language analysis using large multimedia corpora (Crasborn, Hulsbosch, Lampen & Sloetjes 2014), but also enable people to communicate across time and space in new ways (Brinckwirth 2012, Elia 2006, Maslamani 2013). Computer and electronic device meditated communication is a reality in modern language use - both oral and written. To the 21st Century linguist this means not just studying language in

---

[3] These claims are not universally accepted. Inconclusive results are presented by Pytlyk (2007, 2011); and Pattamadilok et al. (2011: 121) while arguing for the orthographic influence on phonology point out: "… that whether orthographic knowledge affects the core mechanisms of speech processing (e.g., lexical access) or some more peripheral processes (e.g., explicit segmentation or decision/comparison) seems to depend strongly on the choice of the tasks that researchers use to probe speech processing." For the purpose of this paper, I take this to mean that there are likely a variety of factors affecting the orthography-pronunciation relationship; of which orthography in the digital device is one.

its non-digital contexts, but also in its digital contexts. It also means language users no-longer have a choice between the two modalities of oral v.s written, rather there is a complex array of options available to most people which cover a plethora of communicative devices and multi-modal/multi-medium scenarios. For example, interlocutor 'A' may get a short email message on his computer from interlocutor 'B' and reply via the 'Facebook Chat' app via his mobile device and carry on several exchanges with interlocutor 'B' before walking into interlocutor 'B's' office and continuing the conversation orally. All the while each segment of the conversation is constrained by the medium it occurs within. Computers have larger screens and keyboards, smart phones have smaller screens and smaller keyboards but also have oral to text features packaged with their OSes, and oral face-to-face communication usually carries with it a host of visual cues and 'reduced' speech forms.

## 1.3.  The Endangered Language Movement

The second of these two movements is the endangered language movement. The endangered language movement can be broken down into two main tenants: document and develop. Krauss is credited with sounding the cry which started this movement (Hale, et al. 1992:9). Sounding a cry that linguists have not only a responsibility to study and document these disappearing languages but also to assist their speakers in the task of developing their languages. He says:

> **We should not only be documenting these languages, but also working educationally, culturally, and politically to increase their chances of survival. This means working with members of the relevant communities to help produce pedagogical materials and literature and to promote language development in the necessary domains, including television.**

### Defining Language Development

Language development is not a new concept; being defined as early as 1968 by Ferguson (1968). However, the distinction between language documentation and language development is pressed by Simons (2011), nineteen years after Krauss[4]. Simons defines language development as:

> **... activities undertaken for the purpose of developing new functions for its language or for restoring lost functions.**

There are two pertinent remarks when considering Simons' definition. The first relates to the first movement mentioned previously - The Digital Revolution. That is, for many languages 'new domains' would include making the language viable in digital contexts, be it written, or oral, or oral with visual support (like YouTube and more generally all kinds of video).

### Distinguishing Language Development from Language Documentation

The second remark is that the long time delay in formally defining 'language development' does not mean that development and development-like activities were not undertaken prior to Krauss' call to action or in the interim between Krauss and Simons' formalized definition. Blench (2012: 13) generalizes the language development pattern in a Nigerian context for the past century saying:

> **A language was first analysed linguistically, a draft orthography was developed, primers to teach the language were printed, and as literacy initiatives were undertaken, Bible translations were very often begun. Wherever literacy took off, in major languages such as Hausa and Yoruba, this would 'leak' into the secular sphere. Books, newspapers and advertising would pick up on the possibility of targeting specific ethnic audiences.**

---

[4] Although Ferguson (1968) also does layout much of the same definition for language development.

The journal *Language Documentation & Conservation* is rife with more recent examples, perspectives, and use cases where linguists have engaged with communities to help "develop" their languages (Amery 2009, Otsuka & Wong 2007, Yamada 2007). Often these use-cases use the term 'language revitalization' to describe their language development type activities. Many ethnolinguistic communities have undertaken language 'revitalization projects' to fortify social and pragmatic positions of heritage languages (for examples see: Reyhner & Lockard 2009). However, the lack of a clear distinction between 'language documentation'[5] and 'language development' for so long a period of time leads to two observations: (1) that in general there has not been a clear distinction in the literature between language development activities and language documentation activities, by those engaged in either or both activities; and (2) that in general these activities do not usually occur individually. That is, generally the activities of language development are encountered in documentation projects as persons affiliated with the academy pursue and engage minority language users. Those activities which make it to the literature, do so because it is persons affiliated with the academy (Mosel 2006: 68) who are looking to fulfill the requirements of the academy. Requirements often include publishing (Nature 2013, Priem, Taraborelli, Groth & Neylon 2010) and a demonstrable impact (Taylor 2011) which itself is part of a larger departmental research profile (for an example see materials by: Provost of the University of Wisconsin 2014).

## *The Pressure to Develop Writing*

When language development (and by extrapolation also language documentation) is approached through the tradition of linguistics, one of the most grappled with issues when working with unwritten languages, is the role of writing in these "undeveloped" languages. The pressure to write a language primarily comes from two directions: (1) from the linguist and (2) from the community[6].

The pressure from the linguist to develop writing in the language is also two-fold. The first of these, is for the creation of the language description materials as part of language documentation. While not all forms of language documentation require a written method of transcription and translation (Reiman 2010), all known forms of advanced degrees in linguistics and related studies (like language documentation; for examples see Dobrin, Austin & Nathan 2007: Appendix 1) require some form of written format. This format usually also requires languages of study to be written either in a linguistic transcription system or a 'working' orthography (for examples see: Mahmudweyssi, Bailey, Paul & Haig 2012 and, Shokri, Jahani & Barani 2013). The second of these pressures on the linguist is often the assumption that a linguist's chief offering to a community of speakers of an unwritten language is that the linguist can provide access to, or help a community of speakers of an unwritten language through the process of orthography development, thereby 'developing' their language. One evidence of this pressure is the rising importance and frequency of addressing issues in orthography development within the language documentation literature (Cahill & Rice 2014, Grenoble & Whaley 2006, Guérin 2008, Lüpke 2011, Mithun 1992, Seifart 2006). This

---

[5] The lack of clear definition also extends into the scope of the term 'language development' and the relationship of this term with terms in the literature like: 'language revitalization', 'language maintenance', 'language revival', 'language renewal', 'language reclamation' or 'reversing language shift/loss'. This is tangental to the current discussion and will be set aside for now. For further commentary on how others see these activities interrelating see Fishman (1964, 1991, 2013), Lewis & Simon (2010) Dwyer (2012), Chelliah & de Reuse (2011: 19-20) and documentation by UNESCO (2003).

[6] The categories 'linguist' and 'community' are not always mutually exclusive (see Ajo, Guérin, Hattori & Robinson 2010, Maxwell 2010, Rice 2011). Therefore perhaps a more accurate distinction is an emic versus an etic (Pike, Headland & Harris 1990) set of pressures. Although one might argue that 'writing' has already been invented and therefore will always move from outside of an ethnolinguistic community to inside the community. As is noted by Kluge (2007) language planners and governments fit somewhere in the equation of writing and literature production. Whether these entities fall on the emic side of the equation or on the etic side varies from case to case.

commentary on orthography development for unwritten languages, though often published in 'language documentation' venues and expressed by linguists, is not always representative of external interests. It is inclusive of the various experiences as linguists have engaged with communities of speakers who have themselves been responding to pressures of globalization and the endangered language movement. These are often communities which been looking to solidify their position in society and codify their speech variety as supporting evidence of their existence.  A second evidence of the pressure from linguist to view writing as a development for a language comes from how writing is presented in the *Graded Intergenerational Disruption Scale* (GIDS) by Fishman (1991) and later the *Expanded Graded Intergenerational Disruption Scale* (EGIDS) by Lewis & Simons (2010). In these scales writing is presumed to start to exist somewhere between levels five and four. GIDS and EGIDS both use higher numbers to represent lower levels of the developmental spectrum of a language (no government support, no trans-generational transmission, and orality is presumed to be the the base modality) and lower numbers to represent language based activities representative of more developed languages (government support, trans-generational transmission, and a functioning written tradition). In contrast to written language activities, level six in both scales focuses on degrees of oral transmission.

Lewis and Simons (2010: 105) summarize Fishmann's GIDS levels four and five (1991: 95-103) as the following:

**Level 4 Literacy in the language is transmitted through education**

**Level 5 The language is used orally by all generations and is effectively used in written form throughout the community.**

Lewis and Simon, in their extended definitions (2010: 110) present levels four[7] and five as the following:

**Level 4 Educational Literacy in the language is being transmitted through a system of public education.**

**Level 5 Written The language is used orally by all generations and is effectively used in written form in parts of the community.**

In Lewis and Simons' discussion (2010: 111) describing each of the levels which deal explicitly with writing they say:

**EGIDS Level 3 (Trade) – This level encompasses languages that may not have official recognition but are "vehicular" in that they are used as a second language by members of multiple first-language communities and serve important functions for business and intergroup communication. They are learned outside of the home either formally or informally and often have a standardized (though perhaps not officially sanctioned) written form.**

**EGIDS Level 4 (Educational) – This level includes languages that are used either as media of instruction or as subjects of instruction in a system of institutionally supported, widely-accessible education. It may be the first language of literacy for speakers of minority languages with eventual acquisition of and transition to one of the languages at a higher level on the EGIDS for more extensive written use. This is the stage that is often described as "mother tongue literacy" or "first language literacy". Institutional support for literacy acquisition may be primarily situated in the local community and be provided by more-or-less formally constituted local institutions that are sustainable. Lee and Mclaughlin (2001) make the distinction at this level between institutions which are primarily under local control (Level 4a) and those which are under the control of outsiders (Level 4b). That distinction may well be useful in many contexts. Here we focus primarily on the existence of institutional support for education in the minority**

---

[7] Fishman (1991) has an extended discussion pertaining to the gradation of literacy within a society in chapter twelve. This accounts for the difference in terminology between Lewis & Simons and Fishman.

**language in contrast to introduced literacy without such institutional support (EGIDS Level 5).**

**EGIDS Level 5 (Written) – This is the level at which literacy is incipient, more often-than-not informally transmitted and with only weak or transient institutional support. Although the introduction of literacy can serve powerfully to improve the prestige of a minority language and may increase its prospects for survival in many cases, the stronger institutional support for literacy acquisition and maintenance found at the levels above is required for ongoing transmission of local-language literacy from one generation to the next.**

This is to say that the increased role of writing (and reading) within a society is, according to GIDS and EGIDS a mark of more advanced progression along a language's path of development.

## Objections to Language Development

It follows then that there can be two objections to the development of a writing system for a language: (1) there can be an objection to the development of a language in any form, and (2) there can be an objection to the specific step in development which calls for the creation of an orthography for a given language. Neither of these objections particularly infer that the person also objects to documenting and describing the language or for its words to be entexualized[8]. That is, linguists may desire to entexualize a language for the purposes of writing a phonetic transcription or for the purpose of presenting the data as support for some theoretical claim and yet hold either of these two objections. For instance Ladefoged (1992) points out that in some cases the development of a language may go against the national policy for the use of the national language. Additionally, it is noted that for some ethnolinguistic communities the speaker population does not warrant the expense of language development.

Arienne Dwyer (2010: 209) also warns against assuming that writing is always the appropriate option saying:

**Sometimes, ethical collaborations may seem to go against the interests of linguistic science and the academy. For instance, a speaker community may not want its language committed to written form or want to maintain a particular orthography out of convention, despite redundancies or missing contrasts.**

Chelliah and de Reuse (C & dR) (2011: 18) echo the sentiments that not all language communities may be favorable towards defining an orthography or entexualizing their language saying:

**The goal of religious organizations such as the Summer Institute of Linguistics (SIL, nowadays called SIL International) and its missionary arm, the Wycliffe Bible Translators (WBT), is ultimately Bible translation. However, these organizations also encourage literacy among indigenous people who do not have a written language (Gudschinsky 1957, Pittman 1948). The reason for this is obviously that if the Bible is translated into an indigenous language, the indigenous people themselves have to be able to read it. Furthermore, literacy is conceived of as a valuable educational goal for the integration of indigenous**

---

[8] By this term I mean the use of markings to encode language in a 'text based' format, be that format handwriting, or typing. Sometimes this process is also referred to as *language codification* (for example see: Bielenberg 1999). Unfortunately, the term *codification* is ambiguous in that language can be codified in several ways - visually (such as with sign languages and co-speech gesture), orally (with speech) or textually (via the use of a system of writing). *Language Codification* can also be used to discus the formalizing of a one or more lects of vernacular through standardization for educational purposes (for example see: Birken-Silverman 1997). The word 'text' also suffers from an unfortunate ambiguity in that it can refer to both a written or typed string of language (with a visual modality), or it can also refer to a segment of a corpus (irrespective of modality). These segments of corpora can then also be oral or in video formats, and not have any written or typed component meaning that a 'text' is not always 'entexualized', but a text will always be 'encoded' though the 'encoding' may take the form of an audio encoding or graphical/visual encoding such as by system of writing. For further review of the term and uses of 'entexualization' see Poveda, Cano & Palomares-Valera (2005: 92).

**peoples into the larger society. The relationship between literacy, literacy development, and fieldwork is somewhat controversial, since some indigenous communities might want to keep their language oral and are therefore opposed to literacy.**

While both texts argue for respecting the wishes and desires of the ethnolinguistic community over the wishes and desires of linguistic researchers, both fail to follow up with any use cases where a community has collectively opted to *not* have its language written down either by the creation of an orthography or by linguistic transcription. It is possible that implicitly there are just so few cases where communities do not want to write their language that no use cases were cited. Although it is almost certainly true that there have been some languages over the course of time which have reached a state of dormancy on EGIDS and also have no written record either by a linguist from outside of the community or by a native speaker; the lack of a written record of the language does not mean that the community consciously made a choice to *not* write their language down. In contexts like West Africa where language is often pragmatically viewed, language as part of a socio-economic ecosystem is subject to only those requirements which the knowledge transmission economy places on it (Mufwene 2002). In these cases the choice to not engage in creating a written record may be, and most likely was prior to the 1950s[9], an unconscious choice dictated by a pragmatic awareness of a socio-economic ecology. It is conceivable that many of these languages' speakers knew that writing existed yet had no knowledge of the skill of writing, or that they had no means of writing, or yet still no cultural value which told them that their language was inherently valuable to record in a written format. Therefore, it should not be assumed that because there has been no written record in these languages that no written record is or ever will be desired on the part of the speakers of a given language.

There are both self-imposed and externally imposed reasons communities and language users choose to not engage in the practice of writing. These can be divided into three types of situations where communities may intentionally *not* engage in the practice of writing.

---

[9] This is not to say that no-one ever wrote some of these languages down. The manuscripts at archives in Timbuktu indicate that some West African languages have been entextualized (Hassane 2008: 116), though the extent of literacy in these scripts and languages is not currently known. The frequency and prevalence of Ajami script to record minority languages is disputed, but seems to be wider than Western scholars previously thought (Mumin 2014: 43).

|  | Engaged in Writing Practice | Not Engaged in Writing Practice |
| --- | --- | --- |
| Self-Volitional Actions | I want to Read and Write! | Engaging in writing is part viewed as a part of inter-generational transmission |
|  |  | Preservation of 'sacred space' requires disengagement with writing |
| Imposed Actions | You will learn to Read and Write! (Western education systems) | Prohibition on learning to write based on social-class |

The first is where there is a conscious choice to not engage in any activity which may enable intergenerational transmission of the vernacular.  Dorian (1987) testifies about the speakers of East Sutherland Gaelic [gla][10], characterizing them as wanting to refrain from actively participating in intergenerational transmission of the vernacular in any medium: oral, written, video, or audio recording. It is significant to say, that this language attitude is a very different attitude than an specific objection to writing. It should also be noted that the community was not against Dorian transcribing their speech, only that they were not going to engage in the activity.

The second type of situation is where there is a conscious objection to the vernacular appearing in a written medium - (particularly where the language use situation crosses into a language domain considered as sacred[11]). The lone case(s) in the literature which seems to substantiate the claims of

---

[10] Three letter items within square brackets are ISO 639-3 codes (International Organization for Standardization 2007). These codes correspond to a variety of language names used for languages and provide a level of clarity when referencing languages. All code values used in this document are current up to the 2014 release of code table changes as released by the ISO 639-3 registrar.

[11] I make no scological claim on the definition of the term as it has been used by various researchers. I am merely pointing out here that not all language domains are treated with the same way within a society.

Dwyer and C & dR comes from the southwest USA. In some Native American[12] speaking language communities, like those reported in McCarty [kee] (2003), Pecos & Blum-Martinez [kee] (2001), Webster [kee], [tew], [hop], [nav] (2006), and Redish & Lewis [tew] (2009) recorded language expression, including written materials, are disfavored with an overall preference by community elders and language leaders for transmission of the languages in oral forms. However, this view on written expression of the language is not held by all community members especially non-fluent younger members of the Keres community (Romero n.d.).

The language situation in the Keresan speaking language communities might be considered stable diglossia as defined in FAMED condition 'D' by Lewis (2010a: 10-1).

> **...functions assigned to the language (i.e the uses which make the language useful) must be compartmentalized, (i.e., clearly identified and identifiable) to the extent necessary so that the use of the language for those functions becomes the norm based on a community-wide consensus. Failure to use the language for those functions comes to be seen as aberrant and draws notice. Where lapses in such patterns of use occur, they are remarked on, pointed out, corrected, and to some degree censured.**

Definitions of diglossia often stemming from observed socio-linguistics situations in European contexts has been defined __((Ferguson 1959))_Look at pages in Ken's book_/__((Fishman 1967))__ where either of two (or more) languages can be used in the same social contexts. Clearly, this is in contrast to how Lewis uses the term. The language situation described by Lewis is sometimes referred to as stable multilingualism __((Clifton 1994))__. In the Keresan contexts there are social constraints in which some members of the community want to prevent one of the two language options from filing specific social communication functions. Thereby creating a unique space (social context) for each language within the total linguistic repertoire of the community. When considered through the definition of stable multilingualism, the Keresan community members do not appear to be opposed to the technology of writing, or its use within other domains of their linguistic repertoire (covered by other languages), but rather are opposed to the perceived influences of writing and their potential impact on culturally significant social contexts. Therefore the issue does not seem to be one merely of objection to the technology as may be simply inferred in more cursory discussions.

In contrast to the efforts of some North American indigenous groups to preserve the culturally sensitive spaces within their social contexts by opposing the technology of writing, other North American indigenous groups have found that the written form of their languages is and has been very important to maintaining the vitality of their language. Several groups are very appreciative that their language has been recorded in the past by linguists: for example Blackfoot [bla] (Kipp 2007: 38-9), Coeur d'Alene [crd] (Vincent, et al. 2013), and Yurok [yur] (A. Garrett 2011) all have heritage language learning programs which have made extensive use of linguistic materials which were once written and archived by linguists. Still other indigenous language communities from North America like the Cherokee [chr] (W. Walker 1984, White 1962), Oneida [one] (DeJong 1998), Lakota [lkt] (Powers 2009), and Lushootseed [lut] (Vincent 2009: 346), all have some sort of language development program which includes a written language component.

While it may be possible to view the language-use/written language situation in Keresan languages as a novelty, it should be pointed out that many languages have domains in which social norms pertaining to writing are different from social norms pertaining to oral language use. There is an emerging awareness among speakers of global majority languages that there are some situations in which

---

[12] The language communities which are reported to have the strongest objections to writing are the Keres speaking communities in the southwest United States. Webster (2006) compares the practices and attitudes of writing in Hopi [hop], Tewa [tew], Keres [kee] and Navajo [nav] communities which all share some level of written langauge restrictions. By using the term "Keresan" I am not implying that these langauges form a genteic affiliation, rather I am using it to refer to the archtype.

language use should not take a written form. These situations often include domains of work, personal opinion and social media (as a medium). For example, Lee Landor, a deputy press secretary to Manhattan Borough President Scott Stringer lost her job following remarks she made on facebook (Maiman 2009), while Paris Brown a Police commissioner in the UK lost her job due to comments made on twitter (Dodd 2013). It is perhaps useful to consider these emerging written discourse domains as sacred[13]. Though, when discussing sacred language use, it is equally important to note that not all domains involving sacred (or religious) areas of language use have Keresan like prohibitions against writing. Various religions take approaches to entexualize their sacred texts. Major world religions like Islam, Hinduism and Christianity all use entexualized versions of their sacred texts, but often prayers or other sacred and personalized speech acts are not entexualized or recorded.

The third kind of situation where a community may choose to not engage in writing is where the practice of not writing or reading is affected by conditions imposed from outside of a given social community. Generally these restrictions affect a certain class of members of the community, for instance: The gender gap report by the World Economic Forum (2013) presents a global picture which can be interpreted such that women in countries which have an Islamic majority culture generally have a lower education, including the ability to read and write. However the gender divide is not true for all countries with an Islamic majority culture.[14] It should be noted that the Qur'ān does call for the equal moral status before Allah of men and women, and some Islamic traditions apply this to include educational opportunities. However, the application of equal treatment through gender roles in Islamic societies varies in its application and can be subject to local interpretation in support for local sociological ideals or values. For example, Boko Haram in Northeastern Nigeria, has a strong position on the kinds of educational opportunities afforded to people, preferring a strict rejection of Western education and an adherence to Islamic education (al-Ashanti 2013: 17, A. Walker 2012: 7)[15]. This position is arguably insupportable as being 'a proper teaching of Islam' based on Islamic sacred texts, but none the less is a common cultural position across Islamic societies (Esposito 1995). A second example of class based limits on the use of entexualized language, which also was justified under economic and sociological ideals can be found in past state law of North Carolina, United States of America (General Assembly of the State of North Carolina 1831). In this case law was made by which African-American slaves were legally prevented from engaging in the practices of learning to read and write so that they would not be encouraged to escape from their owners, causing economic loss to the slave owners.

## *Counter objections and agreement on limited development*

As discussed previously, Ladefoged (1992) presented two objections to universal language development. Implying that language development might be undertaken in some contexts, but should not be under taken in all contexts. Ladefoged's second objection, not previously mentioned, was that linguists should not try to work against national language policies. For reasons already discussed, language development is relevant to the discussion of writing (and typing/text input) in minority languages. However, Ladefoged's second objection deserves some further discussion for two reasons. The first reason is because Ladefoged's hypothesis that the work of a linguist is apolitical is rejected by many in the academic community of linguists. Dorian (1993: 575) in her rebuttal to Ladefoged claims that there is nothing apolitical about a linguist's work. Like Ladefoged, she does not espouse open rebellion to governments but maintains a position that a linguist's work is either seen as encouraging tribalism or

---

[13] Should I discuss how I am using the term 'sacraed', because this is a hotly debated term in Anthroplogy?

[14] Should I include: (Mrkić, Johnson & Rose 2010) and or CNS News coverage (Goodenough 2010)? Respetivly this is another geneder report and an interpretation of the first set of statistics. It is also likely that national boundires are simply not fine grained enough to accruatly describe the cultural boundires nessisary to show significant distincitons.

[15] Al-Ashanti's work is a translation of Murdadā (2012).

nationalism, and depending on the social context this encouragement can be perceived as either a threat or a service to the ruling government. Bernard (1996) takes this observation one step further to include the in-group politics of speech communities, indicating that for any community who has members which want to relinquish their language in favor of some other more prestigious language, there are others in the community who are eager to see their language developed. The fractured nature of speech communities with respect to the desires of language development point to an important issue - "who is the 'language community'?"

Defining the language community is a concept many linguists have struggled to define. For instance, Dorian (1982) challenges us to rethink the 'the speech community' to include previously excluded margins (semi-fluent speakers). Jany discusses how one community was defined in the process of creating an orthography (2010a: 4). Patrick (1999, 2008) provides us with a general review of various perspectives and definitions presented in the linguistics literature. Defining the 'language community' or the 'speech community' remains a challenge in language documentation and in language revitalization efforts. Even archives which attempt to implement community appropriate restrictions and access services struggle with definitions of "community" (Chang 2010: 51, Dobrin, Austin & Nathan 2007: 62-3). Within the scope of definitions, a more inclusive and sociological definition of "speech community" is suggested by Lewis (2010a: 2) as he argues for the sustainable use model (SUM) of language development. In his definition he incorporates the concepts of a multilingualism and social network connections, however he falls short of describing which attributes, or ties in the social network should be counted or weighted more heavily.

> **In a world where contact is the norm, minority language communities are increasingly multilingual. Language development programs therefore must take into account the entire linguistic repertoire of a community, and "community" must be understood in terms of social networks (networks of contact and communication) rather than in terms of a single language treated in isolation. The longstanding primary focus of language assessment on the identification of discrete languages (linguistic similarity, intelligibility, etc.) is less crucial for decision making than is the nature of the relationships of the members of a community to each other and to others around them.**

Defining the "language community" is an issue not just in orthography development, as Jany demonstrates, but also an issue in typing and keyboard layout design. As Lewis points out, individuals at various places in the social network are going to have different language use needs. This translates to also having various uses for specified language input desires. In contrast to embracing or negotiating the notion of 'speech community', keyboard layout designers/creators can and should embrace the notion of a user group or a user base for their product. Much as speech communities use language to fill various functions, so keyboard layout users utilize keyboards layouts to meet various communicative needs. To address the needs of minority language users of keyboard layouts I will use concepts from user experience design which focus on product use rather than notions of "community" as are often sought by linguists. Admittedly there is some overlap in the membership of the two groups: 'user base' and 'speech community'.

The second reason Ladefoged's hypothesis deserves further consideration is because it espouses a synchronic point of view. That is, it does not appear to take into account that national policies and positions regarding language use can and do change over time. It follows then that the perception of threats and benefits of developed and developing languages also changes. In contrast to the change in perception of threats, the imminence of globalization only grows, meaning that the economic and language ecology pressures on minority languages which pressure them increase on the EGIDS scale do not simply disappear with changes in national policy towards languages. In fact, Mufwene (2002: 1) suggests that comprehensive local globalization (by which he means the connection to global economies) has a lager impact and responsibility in language vitality than colonization, to which nationalism and associated language policies are often a response.

If we momentarily set aside the issue of comprehensive local globalization and treat it as a covert or indirect consequence of social organization. There are still many overt or intentional aspects of social

organization which affect with language development and minority language use (including writing). One of the more tangible issues is the issue of the legal position and policy positions many nations have regarding the use of minority languages in various contexts. The position of policies affecting the development and social status of minority languages varies from country to country. Sometimes this position is used to help fortify notions of political identity at the international level. This is the case that Ladefoged refers to with Swahili [swa] in Tanzania__(Cite Malawi book chapter which was scanned)__. Some countries (e.g. Malaysia and Philippines) overtly limit the use of minority languages in schools requiring the language of instruction to be the national language or the official language. Other situations (e.g. Russia) are more covert where, economical and social development policies affect minority language speaking populations and influence their language use (Zamyatin 2012). Yet other governments take a position to encourage minority language development activities in the educational, societal, and cultural spheres of language use__(Hawai'ian)__.

In a similar vein to the Tanzanian example cited by Ladefoged, the language situation in Malaysia and the Philippines was one of … to instill ideas of nationalism.

However, Philippine change of policy - as a counter to ladefoged

(Llamzon 1977) (Yabes 1977) (Gonzalez 1999)

Malaysia
2011. National language planning & language shifts in Malaysian minority communities: speaking in many tongues. Amsterdam; Manchester: Amsterdam University Press Manchester University Press [distributor].

In all of these situations There is still a looming question which Ladefoged ends with, which is who am I to… judge one way or another. However, What is the role of the linguist then in advocacy or activism? or is this perhaps the difference between linguists and language development workers?
While ladefoged's objections universal language development are and have been countered in the literature and to some degree at the International level of the UN, there still remains an unmentioned factor in the limiting of language development. This factor is mentioned in the SUM literature is cost.

Language Development Experts agree - Cite SUM literature
(Lewis 2010b)
(Lewis & Simons Forthcoming)
(Sackett & Humnick 2013)
(Quakenbush & Simons 2012)
(Lewis & Simons 2010)

Mufwene needs cited here some time too…
(Csató & Nathan 2003) Need for quote about

(Eisenlohr 2004)

## Pressure to develop writing from within the community

As previously stated, the pressure to write a language primarily comes from two directions: (1) from the linguist and (2) from the community. The previous section addressed the perspectives of linguists and development professionals. However, it is important to also acknowledge those pressures that come from

within the community. That is, some language communities do want to develop their languages. They want to embrace writing technologies and the global data and information exchange ecologies. Some communities do actively seek out ways to write their language.

Case of cherokee writing, but more to the point today is perhaps the use of minority languages in text messages via cell phones.

__(Cite cases of Russian requests for language development through keyboards for the language)__

(Guérin 2008: 63) Says:

**On the other hand, due to the misconception that written languages are "superior" to spoken ones, providing a writing system for a language may bring pride to the language community (see also Terrill 2002:214–15, Tsunoda 2005:189). Training speakers as writers may also help to widen the range of uses of the language. New genres can develop, from children's literature to diaries, to more practical uses like writing reports or notes. Creating orthographic conventions for an endangered language may also help safeguard that language. Vernacular education is often seen as a key activity in sustaining languages against the pressure of dominant languages (Brenzinger et al. 2003:12, Crowley 2000a:79, Crowley 2000b:383–384), and although "the existence of linguistic descriptions is neither a necessary nor a sufficient condition for language maintenance" (Mühlhäusler 2000:321), it is an essential condition for language revitalization. Once language shift is near completion, revitalization is possible only if the speakers have access to descriptive and written materials in the language (Sasse 1992), since they may shelter the only surviving language resources (Hinton and Hale 2001:241).**

**However, according to Mühlhäusler (1990:190, 199, 203), not only is literacy in the metropolitan languages detrimental to a vernacular language, but "the most general long term effect of literacy in the vernacular has been language decline and death," because, he argues, the goal of vernacular literacy is transitional—to prepare learners to read in a nonlocal language (but see Crowley 2000b, which disagrees).**

**Harrison records that some last speakers of languages are "resigned to fate," or think of language shift as "progress," but last speakers may also regret the loss of their language (2007:9):**

**Svetlana D., one of the last speakers of Tofa, told me in 2001: 'The other day my daughter asked me, 'Mom, why didn't you teach us Tofa?' ... I do not know why. Such a beautiful, difficult language! Now it is all forgotten.'**

(Unseth 2011) west african scripts...

## *1.4.  Writing in Society*

Broadly speaking there are two factors in the role of writing in society: desire and capability.  The first - *Desire* - can be expounded as the sociological practice of writing - the when, the where, and the how writing is done in a given medium and genre, by which I mean the the extent to which language is entexualized in various social communicative interactions: be they SMS messages, E-Mail, blogs, ledgers, shopping lists, books, newspapers, hand written love-notes, or advertisements, etc.  The second factor - *Capability* - can be expounded as the technology used in the writing process, by which I mean the orthography and the tools used to implement the orthography: be they a reed and clay, a quill and ink, a pencil, a manual typewriter keyboard, a laptop keyboard, or a touch screen mobile device.

### *Literacy and Writing as social practice*

Within the larger discussion of literacy, David Barton (1994) discusses the various levels of social use of literacy. The ability (and need) of members of a society to decode entexutalized information varies. This the variation in the perceived need to "be literate"[16] in a language can cross or follow social, economic, cultural or language divisions. Literacy as social practice is where the members of a society embrace the skills of literacy to the point where it becomes a positive social value and common practice to be literate. Such a society and language situation could reasonably be classified as EGIDS level 5. However. there is assumption built into EGIDS assessments which are important to the assessment and adoption of writing in a social setting which uses a particular (set of) languages. This assumption is that if people are reading their language that something is being written in the language. A social value of literacy (as narrowly defined) does not include a social value of writing. EGIDS does not directly address or provide a tool for assessing the social value of writing in a particular language situation.

Writing as a social practice is less often discussed than reading as a social practice, and is relatively undiscussed in the minority language context. However, one would suppose that if reading in a minority language were to become a social practice that writing should also be a social practice. Bernard (1996: 5) provides some interesting counter examples to this assumption showing that at least in some cultures/ language contexts that these values do not necessarily happen together.  <span style="color:red">some examples of written but not literate, and literate but not socially used.</span> That is, Bernard argues that social practice of literacy and writing do not coincide. This observation coincides with early observations about internet use and content production. That is there is something called the 20/80 or 90/9/1 rule__()__. ---(explain rule)--- In recent time this production rule has been rejected in favor of the _____ content producer model__()__. However, what appears to be even absent from the communications literature which talks about content producers is the change in modality of the production of content.

In the late 1990s and early 2000s when the 20/80 was embraced, content production was a challenge, even in majority languages. Devices for producing content were bulkier, tools were not as refined, platforms for disseminating content were clunker and less integrated across the various devices that content consumers used. In general this had the indirect consequence of filtering content producers to a sub-set of the population. Such an observation can lead us to the conclusion that the more difficult or complicated it is to encode (including entexualizing) a message the fewer people or the less often a person is liable to expend the effort to encode that message. This is the basic advantage of user experience design analysis ---(as further discussed in section 2.5)---. User experience design allows designers to look at any interaction and look at pain points, or points of confusion across the task and refine the methods of completing the task so that fewer barriers exist to accomplishing the task. One might say that user experience design allows us to work towards making tasks as intuitive as possible.

---

[16] What it means to "be literate" varies from society to society. The UN staticits on literacy generally measure literacy based on the ability to read a simple sentece - regardless of langauge in which the sentence is written. __(Cite UN statisitcs)__

*The digital Revolution hits social practice - oral phone, texting, and typing*

Disambiguate from "literacies" as in the sense of "competencies". We are talking about the ability to read.__()__

Claim: modality affects written communication.
Claim: medium affects how we read. -interactive mediums affect what we are looking for in the flow of visual processing.

Claim: in our context of language development we are often concerned with getting the orthography correct.
Claim: mediums and context also affect how we entextualize our language.

We see this in analogue in short hand (Barkhuus & Polichar 2011)

The importance of literacy is not lost on the linguists of the endangered language movement. However, the language documentation literature (the same literature which is also mixed with language development issues) is not as succinct about

**As an example of a relatively new form of immersive research, the discipline of ethnography emerged out of social anthropology: that is, white men studying black natives in the jungle, in an attempt to understand and control them. Today, we are the natives, caricatured in the interests of research. (Perks 2003)**

Writing in that movement __(Weber, Wroge & Yoder 2007)__

(Laponce 2004) - Read for comment on keyboarding

*The new social practice hits minority languages*

The challenges of developing and orthography.
But now that comes simultaneously with the challenges of CMC
Use Twitter and Facebook use to show evidence

The map of European languages here: http://www.eupedia.com/europe/maps_of_europe.shtml is contested but gives a point of reference for talking about twitter.

*The Technology of Writing (using keyboard layouts)*

The technology of writing has changed several times over the last two centuries. Quill and ink gave way to the fountain pens and ball point pens. These instruments of hand writing are still in use today but with the invention of the personal computer and the mobile phone, digital communication has supplanted methods of communication dependent on tools for handwriting. With the rise of digital communication the times a keyboard layout is used in a communicative event increases - creating a higher functional

dependency on the device. This serves to highlight the necessity of having keyboard layouts to enable communication. As a precursor to discussing keyboard layouts and design issues applied to them, it is important to acknowledge several additional factors which bear on keyboard layouts and the technology of writing: (1) the keyboard layout is just part of the total language interface on a digital device, (2) keyboards are just one part of the text input solution, and (3) there is a wide variety of physical keyboard device types.

*Text input and the keyboarding contexts*

What is the correlation of these maps to UI's and Voice to text services on these devices. What is the impact on language choice?

*Keyboards*

Overview of physical keyboards
http://www.sis.uta.fi/~pi52316/g/node6.html

Physical v.s virtual
Mobile v.s. stationary

what do they enable people to do?

*What has been done in other languages currently?*
## Roman Script
*Navajo [nav]*
https://itunes.apple.com/us/app/navajo-keyboard/id577841547?mt=8
http://navajonow.com/2011/07/14/navajo-font-vs-navajo-keyboard/
https://www.youtube.com/watch?v=TIaShnHXVmQ
http://www.languagegeek.com/dene/dine/dine_bizaad.html
http://www.languagegeek.com/dene/keyboards/Keymaps/Navajo.pdf
http://en.wikipedia.org/wiki/Talk:Navajo_language#Keyboards
(T. S. Lee & McLaughlin 2001) not sure if this citation is relevant.
*German [deu]*
*Polish [pol]*
*Italian [ita]*
## Non-Roman Script
## General discussion
Asian: (Nandasara & Mikami 2009)
*Chinese []*
Keyboard/Input solution: (2005)
*Japanese [jpn]*
(Morita 1985)
*Korean [kor]*
Script, orthography and phonology: (King 1996) ; Keyboard/Input solution: (Kim 1992)
*Arabic [ara]*
*Greek [ell]*
Keyboard/Input solution: (Marinaras & Lyritzis 1993)
*Armenian*
http://www.hyspell.com/

*Thai [tha]*

Gupta, Renu & Virach Sornlertlamvanich. 2007. Chapter 12 - Text Entry in South and Southeast Asian Scripts. In I. Scott MacKenzie & Tanaka-Ishii Kumiko (eds.), Text Entry Systems, 227-49. Burlington: Morgan Kaufmann.

Radhakrishnan, T., J. W. Atwood & S. G. Krishnamoorthy. 1983. A multilingual input/output device for Indian scripts. International Journal of Man-Machine Studies 19.2: 137-46. http://www.sciencedirect.com/science/article/pii/S0020737383800522

*Bangla []*

two resources (Sattar, Pathan & Ali 2004) (Seddiqui, Hassan, Hossain & Islam)

*Urdu [urd]*

(Afzal & Hussain 2001) - Urdu computing standards

(Ali 1920) - Urdu orthography

(Rehman & Qureshi 2011) - interface design

Several other urdu papers.

*Devanagari []*

Script, orthography and phonology: (Millar & Glover 1981, Pandey 2007); Keyboard/Input solution:

*Sinhala []*

Keyboard/Input solution: (Nandasara 2009)

*Tibetan []*

(Sen, et al. 2005) - Mobile and syllable based

(Wei-lan 2007) Intelligent input software

## Poly-orthography (Regional)

*South Africa*

(Bailey 2007)

*Nigeria*

*Amazigh [tzm]*

akufi.org

http://www.akufi.org/en/tools/downloads.html

*Orthographies*

Orthographies and the connection between an orthography and its keyboard.

explain 'the technology stack'

Do orthographies get designed around the the technology or are they designed around the contrasts evident in the language?

*Typing behaviors and text input behaviors*

three arguments for a better keyboard.

right-left symmetry, center row action, speed of input, fits law.

Is socially acknowledged as needing a better solution: Permanent link to this comic: http://xkcd.com/1284/
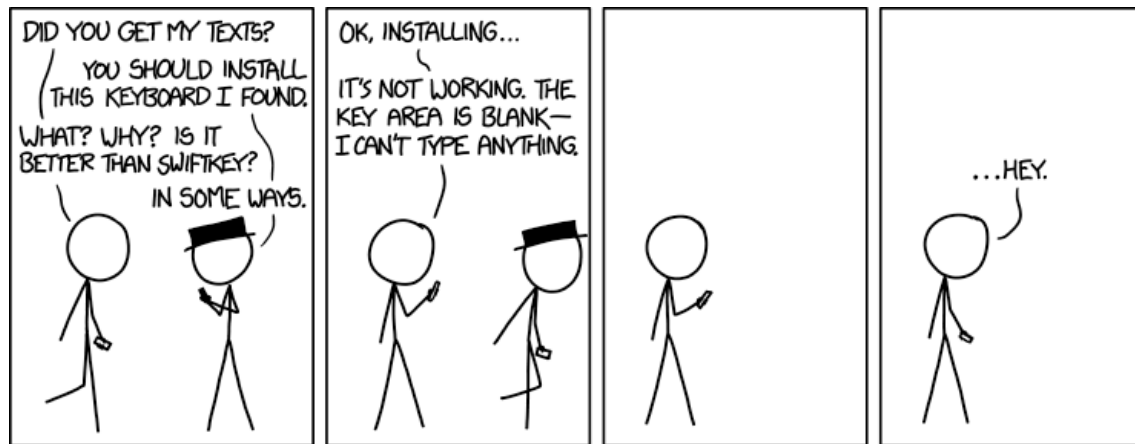
Poly-language keyboards: (Bailey 2007)

Interactive Written Discourse as an Emergent Register (Ferrara, Brunner & Whittemore 1991)
map the orthography to the keyboard: (Jany 2010b)

Kalasha keyboard (Cooper 2005)

keyboard design and operation: (Alden, Daniels & Kanarick 1972)

(Modiano 2001)- 340

**Some sociolinguists perceive this process as linguistic imperialism, pointing out that government agencies and private enterprises, primarily in the UK and the US, export educational materials and operate language schools as a way to extend their 'sphere of influence'. Braj Kachru proposes that one way to safeguard the cultural integrity of the nonnative speaker is to promote those indigenized varieties of English which are established forms of intranational communication (see Kachru 1982). For Europe, however, where the ideology of integration motivation, near-native proficiency, and educational standards based on 'prestige' varieties is accepted and practised, the impositions of Anglo-Americanization are only beginning to be discussed. One hears of 'McDonaldization'. Nevertheless, European integration, and the use of English as the unofficial language for European affairs, is forcing EU citizens to come to terms with Anglo-American 'linguistic imperialism'.**

Keyboarding is important for minority languages (Trosterud 2012)

Cite 'is there a role for linguists anymore'?(Clifton 2013, Ottenheimer 2009, Sebba 2007, 2009)

What is the relationship between an orthography and the keyboard layout?

*Current design process for keyboards*

Something about different views on these relationships what is the current process? the need for a framework

## The Digital Revolution with text input hits Language Documentation

Keyboarding and transcription: (Bowern 2008: 80)

**Your transcription system should be easy to type. There are free Unicode IPA fonts, for example, which can be mapped to keyboards. There are other systems, such as web input, drop-down boxes and scroll through menus, but as soon as you are typing even small amounts of data extra keystrokes or mouse clicks slow down data entry considerably. Ease of typing may affect your choice of transcription system. For example it may be that your language has a vowel system with primary realization of phonemes like this:**

$$\text{(4) i u}$$

$$\text{ɛ ɔ}$$

$$\text{ɐ}$$

**In this case, a, e and o would be good alternatives for ease of data entry over ɐ, ɛ and O (once you're sure of the phonemes). There is more information about things to consider in transcription systems in §5.1.**

**page 196: It's tempting to create orthographies which use only those characters on standard QWERTY keyboards and which use digraphs for any phonemes not covered by the roman alphabet. Many digraphs can make the writing system cumbersome, though. Some of the Kimberley language orthographies have more than half their phonemes represented by digraphs (cf. Kimberley Language Resource Centre 1999). This adds considerably to the length of words and makes reading daunting and difficult for people with limited literacy.**

**Page 197: On the other hand, using entirely roman characters (or the equivalent if you're not using a roman-based writing system) makes documents considerably more portable and avoids font problems, which may otherwise be extremely time-consuming, and require less computer knowledge on the part of users; they do not have to use keyboard mapping software, for example.**

## The Digital Revolution with text misses Language Development

There are two primary points which evade the current linguistics literature. The first is how is a linguists supposed to map a keyboard layout; this is quickly followed by the question: what is the impact of such choices? The second is what is the new relationship between the created object, including its interactional elements and the connecting society?

Support this section with Cahill (Cahill 2011, Cahill & Karan 2008, Esizmetor 2009, MacKenzie & Zhang 2001)

(Whitworth & Ahmad 2013)

Targeting Translation - Counterinsurgency and the Weaponization of Language
Windows 8 now allows writing in 7,000 plus languages: http://blogs.windows.com/windows/b/windowsexperience/archive/2014/02/05/over-7-000-languages-just-1-windows.aspx

## *The internet impacts language use and perception*

L33t speak

## *Keyboarding as gesture, linguistic performance and User Experience*

*Is keyboarding linguistic performance?*

(Jucker & Dürscheid 2012)
Are Instant Messages Speech? (Hunsinger, Klastrup & Allen 2010: 1-21)
In general though some model of what is "language" and what is "linguistics" must be espoused to.

## 1.5. The Role and impact of design

*What is Design?*

What is good design? Is there a science behind design? can we know?

*Elements of Design*

support this section with (Dul, et al. 2004) (J. J. Garrett 2011) (Vitsœ 2012) (Borchers 2000)

*What is impacted?*

Tone and diacritic orthographies

<--- note from Paul--->

**That's really interesting! I would have appreciated your paper at the time - I was working from intuition to combine linguistic info with technical best practice for keyboards and Unicode.**

**Some major factors I found:**

**- It is impossible to underestimate the ability level of older generation typists! They need training (hard to organise for underresourced languages) and/or high ease of use to be productive. Younger people don't find typing so confusing.**

**- People are creatures of habit. Existing well known ways of typing local letters/ diacritics need to be used as a base for a development path, even if they don't fit with standard practice for keyboards internationally. That means slow movement from interim solutions towards (ideally) standards-compliant keyboards integrated with OS'es like Windows.**

**- Some edge case details of the orthographies are still formally undecided, leading to possibly superfluous characters, and that's likely to be a problem for other minority languages as well.**

**- Other languages are also likely to have the same problem of which dialect features or obsolete orthographic features are supported and which are out of scope.**

**- Virtual phone layouts have become relatively very important.**

**- The need to be able to type the majority language is important, for example French mixed in with Berber. That's likely to be a problem for other languages too.**

**- The older generation often insist on no change to the key positions for accessing the majority language. Sometimes this is because they in fact type more often in the majority language than their own! That constrains the minority language support to be squeezed in around the majority language layout, at least initially.**

**- A related issue is that the layout needs to be related carefully to the printed key tops of standard local keyboards, since they are everywhere in cybercafes, homes, schools etc. and since the minority language is underresourced, they are unlikely to be replaced with new dedicated physical keyboards.**

**- The lack of widespread writing of the minority language means that there is little initial pressure for fast typing (design with few keypresses). That comes only later and from younger, demanding users.**

**- The question of typing using deadkeys or by adding diacritics Unicode-style arises for languages with lots of diacritics, especially where there are no precomposed letters.**

**Paul.**

## *1.6. The Role and impact of technical social systems*

*Technical social systems*

When writing about sociological behavior in cyberspace, Hammer (2005) writes:

**Cyberspace is a 'space' that, according to deconstructivist theorists, is changing not only the way that people communicate on an everyday level but changing human identity itself.**

The ability for humans to massively interact via digital interfaces allows the human nature and sociological results of this behavior to manifest itself in new mediums. This behavior is not technological, but it is enabled by technology. The results and new behaviors introduced by limitations, or design features in these technical communications systems are not technology but still the behavior of humans. However, to posit that technology is changing human behavior is a recognition of the pervasiveness of technical systems within human existence. Since the popularization of the smart phone, technical social systems are more relevant and pervasive in the global communications ecology and continue to become more relevant and pervasive. It is important to distinguish technical social systems from other computing systems, and how they relate to concepts like *information technology* or *computers*. It is also important to distinguish concepts like *technology* and *socio-technical systems*. In their book *The Social Design of Technical Systems* (2013: 6-25), Whiteworth and Ahmad align the history of computing with von Bertalanffy's *General System's Theory* (1968) to derive a hierarchical taxonomy wherein computing systems can be studied as mechanical systems, software systems, human systems, and social systems. Whiteworth and Ahmad (2013: 6) divide the history of computing in to the following segments:

- 1950-1960s First computer
- 1970s Mainframe-based computing
- 1980s Client Server interactions
- 1990s World Wide Web and Internet based computing
- 2000s Social-computing with platforms like FaceBook, YouTube and twitter

These divisions (though also by decade) mark significant developments in digital communications; highlighting significant developments in how humans integrated these technologies into their social behaviors. As we establish importance of the keyboard layout in social behavior, it also becomes important to establish keyboard layouts relative to other concepts in the technology domain. Whiteworth and Ahmad present their derived hierarchical taxonomy as follows:

| Name | Scope |
|---|---|
| Socio-technical system (STS) | Community + HCI(s) |
| Human-Computer Integration (HCI) | Person + IT(s) |
| Information Technology (IT) | Software + Device(s) |
| Technology | Any device (or tool) |

It is through this taxonomy that we can see that text input is an issue to both the socio-technical system and the human-computer interaction layers of the applications of technology. This is important as in the words of Bryan Rieger (2011), a user experience designer: "It's not the device people are after, it is all the things the device enables…" Often, the device is only one means to resolving sociological needs.

Keyboards __(discuss in more detail in section___)__ position of keyboard layouts. Keyboard layouts are software which permit and require a specified human-computer interaction, their parameters of

function are defined based on the physical technology, but they are fundamentally a human-computer interaction element which enables people to participate in socio-technical systems. As we look back on the history of computing, we can see the rise of importance of the keyboard with the rise in importance of the socio-technical use of computing. With relationships in mind, we can make the association to language development that if a language community is to develop their language from an EGIDs level 6 to something higher perhaps like a level EGIDS level 4 then, there are going to be multiple challenges when we look at socio-technical systems but one of the first is going to be text input

This becomes important as we look at the relevance and position of keyboard layouts to the overall scope of technology applications and the relevance of technology applications to language development efforts.

as we look at the relevance and position of keyboard layouts to the overall scope of technology applications and the relevance of technology applications to language development efforts.

__(INSERT IMAGES AND DIAGRAMS)__

## *The Keyboard*

What is the position of the keyboard in these social technical systems?
What is a keyboard?

# *1.7.  The objectification of languages*

What does object culture and anthropology have to do with this?

## *Objectification of the language*

Support this section with David Nathan and Anthropology literature
(Dobrin, Austin & Nathan 2007)

## *Object Culture*

(Pearce 1989)
**Objectification of the Orthography**
**Objectification of the Keyboard Layout**

# 2. Methods

## *2.1.  (Methodology) Methodological considerations*

Attaining a suitable Persian Keyboard layout using an evolutionary strategy (Kazem & Naghsh 2011)
What methodology will I follow and why?
Keyboard assumptions:
> Physical v.s virtual
> mobile v.s Stationary
> "ergonomic"
Orthographies get designed around the technology
Typing behavior
> See discussion under REAL data.
> What typing behavior does the keyboard dictate?

What is the REAL typing behavior?

What should be the real typing behavior i.e. DVORAK vs. QWERTY

Current design process of orthographies and keyboards

As described above

## *Phonology overview*

This is an overview of what people have said, not what I believe about it.

## *Corpus Study*

To this end, semantically equivalent texts in three to six languages will be compared based on the following parameters. The texts will be of existing translated material that is currently in circulation. The orthography of a given language will be assessed in terms of the phonological/phonemic distinctions made in the language. Available text input methods (keyboards/ keyboard layouts) will be assessed in terms of minimal inefficiencies, i.e., the number of keystrokes required to input a certain character and thereby create the typed text. The input method comparison will be achieved by mapping each instance of each character in the text to the keystrokes required to produce all characters that appear in the sample text.

## *Other studies which have used biblical texts as the basis of research*

A Fine-Grained Model for Language Identification (Improving Non English Web Searching)

McFarland 1984 KJV  use in word counts.  Warns against Bias based on KJV in Text.

Bible as a parallel corpus: (Trushkina 2006)  (Kilgarriff, et al. 2014)  parallel corpus.(Forsyth & Sharoff 2014)

## *Statistics*

How to count and why it counts

## *Ergonomics*

(Alden, Daniels & Kanarick 1972) (J. Duncan & Ferguson 1974) (David 1985)

## *Alternative layouts*

Look at this website and their work: http://www.eatoni.com/wiki/index.php/Main_Page

## Learning for the first time

## The problem of secondary learning or training effects on relearning

(Anderson, Mirka, Joines & Kaber 2009) (Fagarasanu, Kumar & Narayan 2005)

## Input errors

(Chen, Yesilada & Harper 2010) (MacKenzie & Soukoreff 2002) (Norman 1981)

## *Alternative testing methods*

mention heat mapping and visualizations

Considerations for evaluations (J. Lee & McKay 2010)

*Real Data*

**Risks**
**Real use benefits and data creation habits**

*Economics*

**What are the locally distributed key layouts?**
**What are the locally distributed glyphs?**

## *2.2. Orthography text samples and analyzed keyboard layouts*

These texts samples are for visual reference when considering compared orthographies. It is assumed that these are also the kind of text which would be typed by computer users who are also users of the respective languages. These short samples are all of the same section of the book of James from the Christian New Testament. In the analysis conducted, the whole text of the book of James was used in each case. In contrast to presenting the whole text of each language here, visual samples are presented, and the whole text is made available in Appendix II. The languages used in this study were chosen based on personal experience and accessibly of comparable texts, through a wide geographical distribution of languages was sought. Me'phaa, Chinantec, Spanish, and English are used and spoken in Mexico, and are processed together because they have an overlapping socio-linguistic situation. English, Ezaa, Bekwarra, Cishingini, Igbo and Okphela are used and spoken in Nigeria. The groups of languages from Mexoci and Nigeria both form the bulk of the minority languages considered in this thesis. However, It is also necessary to consider languages which have larger populations of users and have more developed resources for computer-mediated communication. Specific challenges with text input, spelling and diacritic restoration are often evaluated with european languages. So, English, German, French and Italian as languages used and spoken in Europe are considered together. English and Navaho are languages used and spoken in the United States and are compared. Russian and Mongolian are languages used in Mongolia, both using a cyrillic script.

### A. Me'phaa Text Sample

A̱ ngui̱n', tsáan' ninimba̱'la' ju̱ya̱á Jesús, gaju̱ma̱'la' rí phú gagi juwala' ído̱ rí nanújngala' awúun mba'a inii gajmá. Numuu ndu̱ya̱á mála' rí ído̱ rí na̱'nga̱'la' inuu gajmá, nasngájma ne̱ rí gakon rí jañii a̱kia̱n'la' ju̱ya̱á Ana̱'ló', jamí na'ne ne̱ rí ma̱wajún gúkuála'. I̱ndo̱ó má' gí'maa rí ma̱wajún gúkuála' xúgíí mbi'i, kajngó ma̱jráan'la' jamí ma̱'ne rí jañii a̱kia̱n'la', asndo rí náxá'yóo nitháan rí ja'yoo ma̱nindxa̱'la'. [I̱yii' rí ni'tháán Santiági̱o 1:2-4 (Carrasco Zúñiga & Weathers 2008-2010)]

### B. Sochiapam Chinantec Text Sample

Hnoh² reh², ma³hiún¹³ hnoh² honh² li̱³ua³ cáun² hi³ quiunh³² náh², quí¹ la³ cun³ hi³ má²ca³li̱³ ñíh¹ hnoh² jáun² hi³ ti̱³ jlánh¹ bíh¹ re² łi̱²ti̱n² tsú² hi³ jmu³ juenh² tsɨí³, nɨí¹juáh³ zia³² hi³ cá² lau²³ ca³ti̱²¹ hi³ taunh³² tsú² jáun² ta²¹. Hi³ jáun² né³, chá¹ hnoh² cáun² honh², hi³ jáun² łi̱¹³ li̱³ti̱n² hnoh² re² hi³ jmúh¹³ náh² juenh² honh², hi³ jáun² hnoh² łi̱¹³ łn³ náh² tsá² má²hún¹ tsɨí³, tsá² má²ca³hiá² ca³táunh³ ca³la³ tán¹ hián² cu³ti̱³, la³ cun³ tsá² tiá² hi³ li̱³hniauh²³ hí¹ cáun² ñí¹con² yáh³. [Jacobo Jmu² Cáun² Sí² Hi³ Ca³ti̱n¹ Tsá² *Judíos, Tsá² Má²tiáunh¹ Ñí¹ Hliáun³ 1:2-4 (La Liga Bíblica 2009)]

## C. Spanish Text Sample

Hermanos míos, gozaos profundamente cuando os halléis en diversas pruebas, sabiendo que la prueba de vuestra fe produce paciencia. Pero tenga la paciencia su obra completa, para que seáis perfectos y cabales, sin que os falte cosa alguna. [Santiago 1:2-4 Reina-Valera 1995 (RVR1995) ]

## D. English Text Sample

Dear brothers and sisters, when troubles come your way, consider it an opportunity for great joy. For you know that when your faith is tested, your endurance has a chance to grow. So let it grow, for when your endurance is fully developed, you will be perfect and complete, needing nothing. [James 1:2-4 New Living Translation (NLT 2007)]

## E. Ezaa Text Sample

Ụnwunna mu, nteke iwhe adata byakfutaru unu; g'ọ dụlee, unu wojeru iya l'ọo iwhe ẹswa byarụ unu. Unu makwarụ-a l'ọo nteke l'aadata unu nke ekwekwe, unu kweru bẹ unu l'e-shi nwụta ọtaru iwhe nshi. Unu talẹkwawho nshi tafụ iya l'ishi nggẹ unu dụkota ree, dzukwaawho oke; t'ọ bọ dụ iwhe l'a-whọdu l'ẹhu unu.  [Ẹkwo-Ozhi, Jémusu Deru 1:2-4 (Meier 1980)]

## F. Bekwarra Text Sample

Ebwiyaa, k'unyang ng'iyem atitye-atitye a shi n'amin ng'i kan achi-anaani inen na ngin, amin è chi r'irinen k'irityem, k'ucheche dee amin e nyie dee, k'unyang ng'iyem abin a tyung dee achi-anaani inen i yi ang'áchìchī nga, i sha irityem inen k'i bya ha. Amin è ye k'úchú bi irityem he, k'amin è chi r'iyem woo ab'e sha uni k'i giri, k'amin è waa abo chaa iyem achaani fo re. [Ileta Ang'ijems a fuo 1:2-4__ (waiting on SIL Feedback)__]

## G. Cishingini Text Sample

Ạja'ạ ạ vạ, ya'in mạzạnga rạkạ vu dạ ayin a na baci kukondo kakau ku tạwại wạ dạ. I yeve tạ i ili i nan lo yi tạ o kukondo uneki wạ ạdu u dạ. Kukondo ku nan lo ka ku da ka kuzuwa dạ i lulukpạ. Lulukpại ali kukondo ku dạ ku laza. Ayin a na baci ku lazai, kạtạ yo okpo uma a na i derere. Ayin a nan lo, ili i na baci i nambai yi tạ ạ kuciyạ i da. [Akaka a Yakubu 1:2-4 __(waiting on SIL Feedback)__]

## H. Okphela Text Sample

Inyọghuo-mhẹ, leli ọli shi emhi oghẹlẹ ini imufẹ ọdọda e bhale debẹ. Irari khi u lẹsẹ khi ini a da mu irudunga oyẹ fẹ bino ne, iregbemie oyẹ o ya mẹ asha fu-a shi ọ. Zobẹ ni iregbemie oyẹẹ o gbe akanya oyọli pfo, ni a mhọli iwomhẹloe, ni o gba pfo, ni a khi eni ingeli enete ẹa baku. [Ebe ni IJemhisi Ọ kẹkẹ 1:2-4 __(waiting on SIL Feedback)__]

## I. Igbo Text Sample

Guanu ya n'ọṅù nile, umu-nnam, mẹ̀ ọ bula unu dabàra n'ọwunwa di iche iche; ebe unu mara na nnwaputa nke okwukwe-unu nāluputa ntachi-obi. Ma ka ntachi-obi luzue ọlu-ya, ka unu we tozue okè dizu kwa nma, buru ndi ọ dighi ihe ọ bula fọduru nke unu nēnweghi. [Jemes 1:2-4 (Bible Nso 2010)]

### J. German Text Sample

Liebe Brüder und Schwestern! Betrachtet es als Grund zur Freude, wenn euer Glaube immer wieder hart auf die Probe gestellt wird. Denn durch solche Bewährungsproben wird euer Glaube fest und unerschütterlich. Bis zuletzt sollt ihr so unerschütterlich festbleiben, damit ihr in jeder Beziehung zu reifen Christen werdet und niemand euch etwas vorwerfen kann oder etwas an euch zu bemängeln hat. [Jakobus 1:2-4 Hoffnung für Alle (HFA) (Biblica 2002)]

### K. French Text Sample

Mes frères, considérez-vous comme très heureux quand vous avez à passer par toutes sortes d'épreuves ; car, vous le savez, si votre foi résiste à l'épreuve, celle-ci produit la persévérance. Mais veillez à ce que votre persévérance se manifeste pleinement, afin que vous soyez parfaits, sans défaut, qu'il ne vous manque rien. [Lettre de Jacques 1:2-4 (Société biblique française 2000)]

### L. Italian Text Sample

Fratelli miei, considerate una grande gioia quando venite a trovarvi in prove svariate, sapendo che la prova della vostra fede produce costanza. E la costanza compia pienamente l'opera sua in voi, perché siate perfetti e completi, di nulla mancanti. [https://www.bible.com/bible/122/jas.1.nr06] _()_

### M. Russian Text Sample

С великою радостью принимайте, братия мои, когда впадаете в различные искушения, зная, что испытание вашей веры производит терпение; терпение же должно иметь совершенное действие, чтобы вы были совершенны во всей полноте, без всякого недостатка. [Иакова 1:2-4 (Zhuromsky n. d.)]

### N. Mongolian Text Sample

 Ахан дүүс минь ээ! Янз бүрийн бэрхшээл сорилт тулгарвал, түүнийгээ баяр баясгалантай зүйлд тооцож байгтун! Учир нь гэвэл, сорилт туршилтыг тийнхүү давсаныхаа үр дүнд итгэл чинь хөдлөшгүй бат болно гэдгийг та нар мэдэж байгаа билээ. Хөдлөшгүй бат итгэлээ хадгалан, түүнийхээ үр дүнг амьдралдаа бүрэн дүүрэн харуулж байгтун! Тийнхүү та бүхэн нас бие гүйцэх мэт итгэлээрээ өсч бойжихдоо хүнд байвал зохих сайн сайхан зан чанараар огт дутахгүй болно. [ЯАХОБЫН БИЧСЭН ЗАХИДАЛ 1:2-4 (Bible Society of Mongolia 2014)]

### O. Navajo Text Sample

Shik'isóó danohłíinii, ał'ąą át'éego nanihídahodi'nitaahgo, baa hózhǫ́ǫgo baa nitsídaahkees, háálá nihe'oodlą' nabídahonitaahígíí bee nízaadgóó ha'jólní áhooł'įįhgo nihił béédahózin. Áko bíni'dii ha'ahóní nihinaalnishgo ałtso ła'yilééh, ákót'éego nihe'oodlą' bee ts'ídá t'áá ákóne' ádanoht'éego índa t'áadoo bee nihich'į' anáhóót'i'ígóó dahinohnáa doo. [James 1:2-4 (American Bible Society 2000)]

# 3. (Results)The data to be explored

What languages will the compared data be from and why.

## 3.1.  Use Case Mexico

*Chinantec*

   (Foris 1973, 1994, 2000)

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*

**Phonology**
    *Phoneme Chart*
**Orthography**
    *Orthography Chart*
**Keyboard Layout**
    *Keyboard layout diagram*
**Statistics from Text Analysis**
    *Summary of characters used in the text*
    *Characters not used in the text*
    *Percentage of use*
    *Combinational use of characters*
**Observations - brief**

*Spanish*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*

**Phonology**
    *Phoneme Chart*
**Orthography**
    *Orthography Chart*
**Keyboard Layout**
    *Keyboard layout diagram*
**Statistics from Text Analysis**
    *Summary of characters used in the text*
    *Characters not used in the text*
    *Percentage of use*
    *Combinational use of characters*
**Observations - brief**

*English*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*

**Phonology**
    *Phoneme Chart*
**Orthography**
    *Orthography Chart*
**Keyboard Layout**
    *Keyboard layout diagram*
**Statistics from Text Analysis**
    *Summary of characters used in the text*
    *Characters not used in the text*
    *Percentage of use*

*Combinational use of characters*

## Observations - brief

*Me'phaa*

The original purpose of my involvement in the keyboard design for Me'phaa was to facilitate text creation and the typing of texts as part of the NEH funded project *Documenting the Me'phaa Genus* (Marlett 2010 NEH-DEL: FN-50079-10). I particularly facilitated technology use on OS X and worked with a fellow team member, Kevin Cline, who facilitated technology use on Windows based operating systems. An existing keyboard layout was already in use by several Me'phaa writers, including some bilingual teachers in the Me'phaa speaking region. Since some of these writers were also going to be involved in the text collection and text creation process for the language documentation project, it was decided to use the existing keyboard layout as a starting point. In this way the documentation project would maximize the continuity from the previous typing experience of contributors.

The pre-existing keyboard layout was created by Mark L. Weathers, and a team of Me'phaa speakers who have been involved in a long standing language development project. Their design process was organic, but was influenced by the following factors:

- The keyboard commonly used in Mexico (Spanish ISO)



17

- Access to characters from the Me'phaa orthography
- The design standard (ISO v.s ANSI) of the physical keyboards18

---

[17] To be converted to caption: Spanish ISO keyboard layout. The most commonly used keyboard layout in Mexico - No keys depressed. Modifier keys in orange or shaded.

[18] Keyboards sold in the U.S.A. are more often ANSI, whereas physical keyboards sold in Mexico are more often ISO. Laptop manufacturers and physical keyboard producers like Dell and Belkin often alter products slightly by changing the physical Keyboard depending on the market audience to which the device will be shipped.

19

An additional challenge encountered in the Me'phaa design case was presented by the use of a custom font. The community had a custom font made for them many years ago. This font was then converted (some fewer years ago) to Unicode code points. However, the conversion to Unicode was incomplete. Several glyphs were intentionally mismatched by hacking the font20. Through the hacking process, the Unicode code points no-longer represented the intended Unicode glyphs. This was intentionally done for two reasons:

1. To make the expected glyphs appear because there was not a Unicode code point for them.
2. To make input from Spanish ISO and QWERTY ANSI laptop keyboards appear to create the correct glyphs when using the font with written Me'phaa texts.

The goal for hacking the font was to make it possible for people in the community to type on their own computers in Me'phaa. From a typist's perspective all that was needed was to have the "Me'phaa Font" and use that font when rendering the document. However, using the hacked font created documents which were composed in Unicode and therefore "Unicode Compliant" but where the characters would not display or print as the intended Unicode code points. Socially, this means that if a document creator controls the complete life of the document from creation to publication and printing, that the glyphs used by the printer device would match the expected graphemes. However, if a Me'phaa writer takes a document written with the "Me'phaa font" to a local printer - say to create a banner or poster - that the Me'phaa unique glyphs would not print as the document creator intended. The larger impact on the community is that there is now an expectation that the "Me'phaa font" does not work when printing but does work on you computer. For documents created as part of the NEH funded language documentation project, it was decided that Unicode compliance was necessary. This meant designing keyboard layouts which would produce the expected input and also map the glyphs to their correct Unicode code points. This functionality was needed across several platforms including: Windows XP, Windows 7 and Mac OS

---

[19] To be converted to caption: The Spanish ISO keyboard layout laid out on an ANSI keyboard rather than an ISO keyboard.

[20] Hacking as defined by Priest (2004) was common practice prior to the adoption of current best practice of using Unicode. In the Me'phaa case, the rationale for hacking is partially because a capital Saltillo (LATIN CAPITAL LETTER SALTILLO U+A78B) was needed for the orthography and was only available as a PUA character until March 2008 when it was added to Unicode version 5.1.0.

X. To create these keyboards we used Ukelele version 2.1.9 (Brownie 2012) and Microsoft Keyboard Layout Creator (MSKLC) version 1.4 (Microsoft 2012).

    The resulting product for the language documentation project was a keyboard layout which was based on the keyboard layout for Spanish ISO, but was set to work on ISO keyboards and ANSI keyboards. This layout was consistent across the various operating systems (OSes). Because several dead keys[21] were used this also affected the behavior of the keyboard and the typing experience. This is discussed more fully in sections three and four.

https://airspace.leapmotion.com/apps/bettertouchtool/osx



22

*Comparative analysis of these keyboard layouts*

**Multilingual comparison section summary**

    When evaluating the *User Experience* around a keyboard layout one of the questions to ask is: *How accessible to typists (users) are the most common characters they will be trying to access?* This question touches on several aspects of the user experience in that it pushes designers to think about the orthography, the tactile experience, and the frequency of characters used in the language. In many regards this is a similar question which Dvorak asked when he developed an alternative to the QWERTY keyboard layout. However, in the case of minority languages there is the added complication that there are often characters in the minority languages which are not included in the national language or function differently than they do in the national language.

**About the texts**

    To answer the above question I compared four texts, one in each:

---

[21] Dead keys are keys which are struck and released prior to the final key so that in some way a modification or alteration of the normal final character is achieved.

[22] To be converted to caption: Me'phaa keyboard with no shift state.

- English (NLT 2007)
- Spanish (RVR 1995)
- Sochiapam Chinantec (Unknown 2009)
- Me'phaa (Carrasco Zúñiga et al. 2008).

The texts were translations of the book of James found in the Christian scriptures. It was chosen because it was readily available across the set of target languages and by using translated texts it gave a text based language sample which maintains a semantic continuity across the various texts.

To perform the analysis I counted the characters used in each text with perl script (Hallissy 2005) and then also counted the words in the text using a small text counting program (Peter Bilak & Rossum 2003). These results were then adjusted as needed to account for complex characters and composite characters. Each character was assigned a multiplier based on how many keystrokes were used by that particular keyboard to produce each character. The resulting counts are available in the charts below.

| Corpora Statistics | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Me'phaa | | Sochiapam Chinantec | | Spanish | | English | |
| Character Count | Word Count | Character Count | Word Count | Character Count | Word Count | Character Count | Word Count |
| 16618 | 2856 | 19610 | 4506 | 12147 | 2165 | 10432 | 2575 |
| Keystrokes | 22454 | Keystrokes | 27515 | Keystrokes | 12715 | Keystrokes | 13575 |
| Keystroke to character Efficiency | 74.01% | Keystroke to character Efficiency | 71.27% | Keystroke to character Efficiency | 95.53% | Keystroke to character Efficiency | 76.85% |

In this presentation of the data, a keystroke is understood to be the pressing of a button on the keyboard even if it does not independently return visible output. A space break is counted as a keystroke. A capital letter ⟨ A ⟩ on an English QWERTY keyboard would require two keystrokes: shift + a. This of method of counting keystrokes is different than the method employed by NRSI (Constable 2001) because they do not count the use of modifier keys like shift to be two key strokes. A character for the purpose of this chart is understood to be what is technically considered an orthographic character plus any graphemic variations ( ⟨ a ⟩ and ⟨ A ⟩ are the same orthographic character but are different graphemes each with their different grapheme counts). An orthographic character may be a composed character (requiring more than one unicode code point) ⟨ í ⟩, a composite character (having more than one grapheme but only one unicode code pint) ⟨ á ⟩, or a base character ⟨ n ⟩. In terms of user experience an orthographic character should be considered what an minority language writer would consider one "letter". In Sochiapam Chinantec each number representing tone ⟨ ¹ ,² , ³ ⟩ is considered a letter. In Me'phaa the vowels with their tone markings ⟨ á , a̲ ⟩ are considered separate letters from the base ⟨ a ⟩ (vowels without tone markings - indicating a mid-tone). Word count as indicated in the chart is understood to be the orthographic word as indicated by space markers. The keystroke to character efficiency percentage is the orthographic character count divided by the number of keystrokes to achieve that character count.

In general the higher the number the better the keyboard layout for the target orthography. With an ideallic 100% there would be a one-to-one correspondence between orthographic graphemes and keystrokes ( The graphemes ⟨ a ⟩ and ⟨ A ⟩ would have separate keys and would each require one keystroke to create.). But notice that even with English we operate at a 76.85% efficiency. Generally, a statistician would ask for a T-Test or an ANOVA analysis to determine if there is a significant difference indicated by the keystroke to character efficiency ratio between the various orthography implementations

and the various keyboard layouts for each language. The challenge with this additional level of analysis, is that even if we decide that 1.00% is a significant difference there is no explicit margin of what the difference is significant for. That is, *are we measuring the probability that a user will not want to type in their language with their language's keyboard layout or just not type in their language?* Currently, in the minority languages of Me'phaa and Sochiapam Chinantec there are relatively few authors, fewer typists and even fewer yet who have access to either of the current solutions offered in this paper, and those who are typists, have an immense desire to type in their language regardless of the challenges involved. Therefore measuring efficiency and correlating it to social attitudes based on control groups of minority language authors is relatively impossible at this time. If we were to attempt measure some sort of cross-language-relatedness, two remaining a challenges need to be accounted for:

1. An authors' perception that there is only way to accomplish the task of (digital) authoring in the minority language (either by use of the keyboard or by use of a hacked font).
2. The task of authoring is only being completed by persons which have determined that any way to perform the task, regardless of the effort involved, is sufficient.

Either of these dynamics above makes the input from these typists as unbiased users suspect. However, one might ask, *what is the efficiency ratio even useful for?* In leu of a better metric, the efficiency ratio useful to designers as a metric in understanding efficiency when we compare various solutions both internal to one language and across languages. Because the metric is in and of itself is not a good indicator of overall impact on writing ability for small groups of motivated authors we need to look to other metrics and indicators as well.

To further develop the analysis and to understand what the differences between the actual percentages mean in terms of tactile experience the characters in each language were divided into three groupings: *punctuation*, *characters without diacritics,* and *characters with diacritics*. In conjunction to this division by character type, a heat map image of the frequency of the used keys was created using a javascript browser app called Heatmap Keyboard (Wied 2011). __((how does this section fit here?)In traditional keyboarding theory there are two conditions of well designed keyboard layouts which are claimed to be important to ergonomics and efficiency of typing and text input:

1. The most common characters are accessible to the typist from the home row, where fingers sit while at rest __(I don' like how not each of these citations is explained (Fitts 1954, MacNeilage 1964, Vella, Vigouroux & Gorce 2009))__.
2. A regular rhythm can be created between the strokes of the right hand and those of the left (Dealey & Dvorak 1933, Dvorak 1935). )__

For traditional keyboard experiences, the heatmap images present a better visualization of the how the keyboard layout affects the tactile portion of the *User Experience.* When considering modern touchscreen-context sensitive devices, keyboard layouts are often adjusted relative to the task options the user is permitted to attempt. Therefore some alterations may need to be considered if designers are targeting those kinds of devices. Generally, even with those devices the alternative keyboard layouts are alterations of the the main layout not a redesign from an unfamiliar layout. Further discussion on context sensitive keyboard layouts is beyond the scope of this paper, but in general should be pursued because of the social penetration of small electronic devices - even among minority language speaking populations.

__(The characters have also been divided by how they are accessed i.e. if they require a modifier key or a dead key to be accessed by the user.)__

## About the compared orthographies

To understand the test results it might be helpful if some clarifications were made regarding how graphemes in the compared orthographies look the same but behave differently and the orthographies handle some similar phonological properties like tone and stress in very different ways. A discussion of the graphemic representation of tone in Me'phaa and Sochiapam Chinantec can be found in section 3 of

this paper under the the discussion on design. Aside from indicators of tone and stress, there are several unique graphemes or characteristics of each orthography which are relevant.

1. orthographic depth

for a light treatment of [tcf] tone consult (Suárez 1983:25-26) or for a more recent discussion on tone in [], which also uses this keyboard layout, but was not analyzed, consult __((Black 2005, Cline 2013))__

2. Not yet mentioned orthographic characters
diaeresis and barred i accent mark = stress

__(orthographic depth, Saltillo and diaeresis and barred i accent mark = stress)__ It should also be noted that both of the minority language orthographies referenced here are somewhat in a state of flux, or rather there is not necessarily an academy or strong literary tradition which prevents writers or educators from changing how they orthographically represent their language. Because of this there may be slight innovations over time in the way that writers of these languages express themselves. Marlett (enero 2012a) provides a good summary overview for the various ways the Me'phaa language has been expressed in popular and academic literature. Data for the comparison below comes from a small book titled *Xó nitháán mè'phàà: Cómo se escribe el tlapaneco* which was designed to explain to Me'phaa speakers how to write in Me'phaa (Asociación para la Promoción de Lecto-Escritura Tlapaneca 1988: 11).

Similarly, Fortis (2000) provides a great explanation in English for how the Sochiapam orthography works, whereas a brief introduction is available in Spanish from the introduction to the Chinantec New Testament (Unknown 2009b).

The chart below shows a composite of the phonetic graphemes used across the four languages. Referencing this chart may be helpful if trying to determine if graphemes absent in other charts are simply absent from the text or if they are not used in a particular language.

**Comparison of Graphemes**

| | a | b | c | d | e | f | g | h | i | ɨ | j | k | l | m | n | ñ | o | p | q | r | s | t | u | v | w | x | y | z | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| eng | a | b | c | d | e | f | g | h | i | | j | k | l | m | n | | o | p | q | r | s | t | u | v | w | x | y | z | | | | | | |
| spa | a | b | c | d | e | f | g | h | i | | j | k | l | m | n | ñ | o | p | q | r | s | t | u | v | w | x | y | z | ´ | | | | | |
| tcf | a | b | c | d | e | f | g | h | i | | j | k | l | m | n | ñ | o | p | q | r | s | t | u | v | w | x | y | ' | ´ | | | | | _ |
| cso | a | b | c | d | e | f | g | h | i | ɨ | j | | l | m | n | ñ | o | p | q | r | s | t | u | | | | y | z | ´ | 1 | 2 | 3 | .. | |

## Punctuation

Punctuation marks make up a relatively small, but necessary, portion of each text. They enable the minority language writer to be able to communicate clearly. In the compared texts the character count for punctuation marks range from 2.71% - 4.13% of the total orthographical characters used. The frequency of use and keyboard placement of these characters can have a profound impact on keystroke usage. Make punctuation marks too difficult to access and complexity is introduced in every sentence. Make the punctuation marks too easy to access and other characters which are more frequent in the language are made harder to access therefore also introducing complexity potentially at every word. In each of the analyzed texts, the punctuation marks reference sentence or clausal level information. However, this may

not always be the case as some orthographies use punctuation marks in a semiographic manner to represent the subject of the sentence or the tense of a verb (Roberts 2011: 86, 88).

Across analyzed texts there are three noteworthy variations in the kinds of punctuation marks used: The use of inverted punctuation marks, those marks used to signify quotations and those signifying questions. In the Me'phaa text ⟨ << , >> ⟩ are used to signify quotes. It is assumed that these are supposed to be encoded as ⟨ « , » ⟩following formal Spanish orthography conventions as discussed in section 3 under design. Chinantec follows English standard orthography practices and uses quote marks ⟨ " , " ⟩. Unlike the other texts, no question marks appear in the Me'phaa text because they have a question particle and therefore do not need a question mark ⟨ ? ⟩. English does not use inverted punctuation marks.

| Punctuation Usage | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Characters | | Me'phaa | | Sochiapam Chinantec | | Spanish | | English | |
| Unicode Value | Glyph | Number of occurrences in the text | Percentage comprising the whole text | Number of occurrences in the text | Percentage comprising the whole text | Number of occurrences in the text | Percentage comprising the whole text | Number of occurrences in the text | Percentage comprising the whole text |
| U+002C | , | 216 | 1.3% | 288 | 1.4% | 183 | 1.8% | 148 | 1.4% |
| U+002E | . | 177 | 1.0% | 131 | 0.6% | 96 | 1% | 148 | 1.4% |
| U+003E** | > | 32 | 0.2% | - | - | - | - | - | - |
| U+003C** | < | 32 | 0.2% | - | - | - | - | - | - |
| U+00BB | » | - | - | - | - | 15 | 0.1% | - | - |
| U+00AB | « | - | - | - | - | 15 | 0.1% | - | - |
| U+201C | " | - | - | 15 | 0.1% | - | - | 13 | 0.1% |
| U+201D | " | - | - | 15 | 0.1% | - | - | 13 | 0.1% |
| U+003B | ; | 0 | 0.0% | 28 | 0.1% | 17 | 0.2% | 4 | 0% |
| U+003A | : | 20 | 0.1% | 22 | 0.1% | 17 | 0.2% | 5 | 0% |
| U+00A1 | ¡ | 2 | 0.0% | 6 | 0% | 4 | 0% | - | - |
| U+0021 | ! | 2 | 0.0% | 6 | 0% | 4 | 0% | 9 | 0.1% |
| U+005D | ] | 1 | 0.0% | 2 | 0% | 0 | 0% | 0 | 0% |
| U+005B | [ | 1 | 0.0% | 2 | 0% | 0 | 0% | 0 | 0% |
| U+0029 | ) | 1 | 0.0% | 0 | 0% | 0 | 0% | 0 | 0% |
| U+0028 | ( | 1 | 0.0% | 0 | 0% | 0 | 0% | 0 | 0% |
| U+003F | ? | - | - | 14 | 0.1% | 23 | 0.2% | 23 | 0.2% |
| U+00BF | ¿ | - | - | 14 | 0.1% | 23 | 0.2% | - | - |
| U+2019 | ' | 0 | 0.0% | 0 | 0% | 0 | 0% | 36 | 0.3% |
| U+2014 | — | 0 | 0.0% | 0 | 0% | 0 | 0% | 6 | 0.1% |
| U+002D | - | 0 | 0.0% | 0 | 0% | 0 | 0% | 2 | 0% |
| U+002A | * | 0 | 0.0% | 3 | 0% | 0 | 0.0% | 0 | 0.0% |
| Sum | | 485 | 2.92% | 546 | 2.78% | 397 | 4.13% | 407 | 3.90% |
| Percentage of keystrokes in the whole text used to create punctuation marks | | 2.41% | | 2.37% | | 4.25% | | 2.95% | |
| Number of keystrokes used to create all punctuation marks | | 543 | | 651 | | 541 | | 435 | |
| Punctuation character keystroke efficiency | | 89.32% | | 83.87% | | 73.38% | | 93.56% | |

There are several figures and ratios which are important to keep in mind as we consider the other sections of the orthographies and the typist's access to these characters. If we consider the total quantity of punctuation characters in their breakdowns for usage we see that there are basic consistencies across the texts i.e. there are approximately 15 quotes in each of the texts. This should be expected from a

translated text. However, we also see that there is also some variations like the the number of full sentences as indicated by periods in each text. This variation can be expected to reflect factors of naturalness, style in translation and also the grammar of the target language. As a keyboard layout designer, an important figure is the punctuation character keystroke efficiency percentage. This tells us how accessible the punctuation graphemes are to the the typist at any given time. We notice from the figures above that a English language typist using a QWERTY keyboard has a high percentage of access at any given time, whereas a Spanish language typists using a Spanish ISO keyboard layout must use more punctuation characters and has a more involved experience accessing those characters (more keystrokes). This gives that typists a lower efficiency rate for punctuation characters. These ratios are important to keep in mind and evaluate as the designer considers the placement of punctuation characters in relation to characters with and without diacritics.

## Characters without Diacritics

In the compared texts the characters without diacritics are all consonants. And all consonants are without diacritics, with the exception of ⟨ ñ ⟩. Characters without diacritics differ from base characters in that base characters are characters without diacritics but can receive diacritics. Therefore users of the orthography must decided if that base character should receive a diacritic to produce its correct form or not. The orthographic character ⟨ ñ ⟩ is included in both sections as a character with a diacritic and a character without a diacritic. This is in part because it is a consonant and it is beneficial, in these languages to consider it with the other consonants. It is also in part because, the English QWERTY keyboard treats this character as a character with a diacritic, however the spanish ISO keyboard gives it a single key for this character elevating it the status of a full character. As the discussion in section 3 under design points out, the status of this character is open to interpretation by orthography users.

| Consonants (Characters without diacritics) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Characters | | Me'phaa | | Sochiapam Chinantec | | Spanish | | English | |
| Unicode Value | Glyph | Number of occurrences in the text | Percentage comprising the whole text | Number of occurrences in the text | Percentage comprising the | Number of occurrences in the | Percentage comprising the whole text | Number of occurrences in the text | Percentage comprising the whole text |
| U+0042 | B | - | - | - | - | 1 | 0% | 14 | 0.10% |
| U+0062 | b | 283 | 1.70% | 140 | 0.70% | 159 | 1.60% | 159 | 1.50% |
| U+004C | C | - | - | 12 | 0.10% | 8 | 0.10% | 8 | 0.10% |
| U+0063 | c | 6 | 0% | 500 | 2.50% | 317 | 3.20% | 178 | 1.60% |
| U+0044 | D | 2 | 0% | 68 | 0.30% | 21 | 0.20% | 11 | 0.10% |
| U+0064 | d | 220 | 1.30% | 81 | 0.40% | 456 | 4.60% | 464 | 4.30% |
| U+0046 | F | - | - | - | - | - | - | 15 | 0.10% |
| U+0066 | f | 5 | 0% | - | - | 76 | 0.80% | 242 | 2.20% |
| U+0047 | G | 17 | 0.10% | 1 | 0% | - | - | 44 | 0.40% |
| U+0067 | g | 442 | 2.60% | 31 | 0.20% | 96 | 1% | 197 | 1.80% |
| U+0048 | H | - | - | 58 | 0.30% | 21 | 0.20% | 14 | 0.10% |
| U+0068 | h | 305 | 1.80% | 1,835 | 9.10% | 115 | 1.10% | 571 | 5.30% |
| U+004A | J | 25 | 0.10% | 47 | 0.20% | 4 | 0% | 10 | 0.10% |
| U+006A | j | 466 | 2.70% | 751 | 3.70% | 41 | 0.40% | 21 | 0.20% |
| U+004B | K | 6 | 0% | - | - | - | - | 1 | 0% |
| U+006B | k | 295 | 1.70% | - | - | - | - | 65 | 0.60% |
| U+0043 | L | - | - | 12 | 0.10% | 14 | 0.10% | 21 | 0.20% |
| U+006C | l | 330 | 1.90% | 549 | 2.70% | 466 | 4.70% | 398 | 3.70% |
| U+004D | M | 8 | 0% | 2 | 0% | 4 | 0% | 2 | 0% |
| U+006D | m | 829 | 4.80% | 396 | 2% | 283 | 2.80% | 195 | 1.80% |
| U+004E | N | 31 | 0.20% | 7 | 0% | 11 | 0.10% | 2 | 0% |
| U+006E | n | 1,606 | 9.40% | 1,496 | 7.40% | 580 | 5.80% | 596 | 5.50% |
| U+00D1 | Ñ | - | - | 1 | 0% | - | - | - | - |
| U+00F1 | ñ | 116 | 0.70% | 128 | 0.60% | 23 | 0.20% | - | - |
| U+0050 | P | 5 | 0% | - | - | 21 | 0.20% | 2 | 0% |
| U+0070 | p | 44 | 0.30% | 13 | 0.10% | 221 | 2.20% | 141 | 1.30% |
| U+0051 | Q | - | - | 19 | 0.10% | 2 | 0% | - | - |
| U+0071 | q | - | - | 149 | 0.70% | 115 | 1.10% | 2 | 0% |
| U+0052 | R | - | - | 2 | 0% | 1 | 0% | 3 | 0% |
| U+0072 | r | 470 | 2.70% | 77 | 0.40% | 658 | 6.60% | 661 | 6.10% |
| U+0053 | S | 5 | 0% | 7 | 0% | 27 | 0.30% | 21 | 0.20% |
| U+0073 | s | 251 | 1.50% | 502 | 2.50% | 800 | 8% | 711 | 6.60% |
| U+0054 | T | 14 | 0.10% | 40 | 0.20% | 8 | 0.10% | 16 | 0.10% |
| U+0074 | t | 359 | 2.10% | 699 | 3.50% | 369 | 3.70% | 839 | 7.70% |
| U+0056 | V | - | - | - | - | 6 | 0.10% | - | - |
| U+0076 | v | - | - | - | - | 113 | 1.10% | 126 | 1.20% |
| U+0057 | W | - | - | - | - | - | - | 11 | 0.10% |
| U+0077 | w | 167 | 1% | - | - | - | - | 267 | 2.50% |
| U+0058 | X | 43 | 0.30% | - | - | - | - | - | - |
| U+0078 | x | 353 | 2.10% | - | - | - | - | 7 | 0.10% |
| U+0059 | Y | - | - | - | - | 6 | 0.10% | 25 | 0.20% |
| U+0079 | y | 126 | 0.70% | 38 | 0.20% | 127 | 1.30% | 320 | 3% |
| U+005A | Z | - | - | 2 | 0% | - | - | - | - |
| U+007A | z | - | - | 73 | 0.40% | 34 | 0.30% | 4 | 0% |
| U+A78B | ' | 1 | 0% | - | - | - | - | - | - |
| U+A78C | ' | 1,189 | 7% | - | - | - | - | - | - |

## Characters with Diacritics

2.1.1 history of the Orthography. Include history of typing

Diacritic first then base with the exception of low tone in Me'phaa

to note here that there are two different keyboard layouts for Spanish.

Basic orthography explanation xV xC xTones

The test and what was done and why

| Me'phaa base characters | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | A | e | E | i | I | o | O | u | n | N |
| Number of total Low tones *(use of Combining Macron Below U+0331 plus base)* | 875 | 16 | 226 | 0 | 198 | 5 | 214 | 0 | 300 | n/a | n/a |
| Number of total uses of acute marks ⟨ ´ ⟩ or tilde ⟨ ˜ ⟩ *(composite characters using base)* | 880 | 9 | 19 | 0 | 830 | 4 | 222 | 0 | 268 | 116 | 0 |
| Number of times the base glyph is used without modification | 1195 | 93 | 100 | 1 | 502 | 21 | 185 | 2 | 590 | 1606 | 31 |
| Number of total base characters | 2950 | 118 | 345 | 1 | 1530 | 30 | 621 | 2 | 1158 | 1722 | 31 |

## Chinantec

The orthography and typing situation is different in in Sochiapam Chinantec than in Me'phaa. While both are tonal languages Sochiapam Chinantec expresses its tones as superscript numbers after a vowel. This means that the tone marks are not combining. However, this does not mean that Sochiapam Chinantec does not have diacritics. In deed it does. They simply represent stress which also needs to be marked in the orthography. (Unknown 2009b)

| Sochiapam Chinantec base characters | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | A | e | E | i | I | ɨ | o | O | u | U | n | N |
| Number of total Low tones *(use of Combining Macron Below U+0331 plus base)* | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| Number of total uses of diaeresis. (use of ⟨ ¨ ⟩ ) | | | 4 | | | | | | | | | | |
| Number of total uses of acute marks ⟨ ´ ⟩ or tilde ⟨ ˜ ⟩ *(composite characters using base)* | 1356 | 0 | 126 | 0 | 467 | 0 | 403 | 130 | 0 | 251 | 0 | 128 | 1 |

| Sochiapam Chinantec base characters | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | A | e | E | i | I | ɨ | o | O | u | U | n | N |
| Number of times the base glyph is used without modification | 715 | 2 | 159 | 0 | 1046 | 3 | 304 | 316 | 1 | 1054 | 5 | 1496 | 7 |
| Number of total base characters | 2071 | 2 | 289 | 0 | 1513 | 3 | 707 | 446 | 1 | 1305 | 5 | 1624 | 8 |

## History of orthography.
## Keyboard.
## Results of the Test.
## Comparison to National Language.

The Spanish Keyboard

| Spanish base characters | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | A | e | E | i | I | o | O | u | U | n | N |
| Number of total Low tones *(use of Combining Macron Below U+0331 plus base)* | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| Number of total uses of acute marks ⟨ ´ ⟩ or tilde < ˜ ⟩ *(composite characters using base)* | 59 | 0 | 65 | 1 | 59 | 0 | 41 | 0 | 9 | 0 | 23 | 0 |
| Number of times the base glyph is used without modification | 1077 | 15 | 1185 | 15 | 514 | 2 | 935 | 1 | 425 | 1 | 580 | 11 |
| Number of total base characters | 1136 | 15 | 1250 | 16 | 573 | 2 | 976 | 1 | 434 | 1 | 603 | 11 |

## U.S. English point of Reference.

Because this is closer than British English and there is a large migration population.

| English base characters | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | A | e | E | i | I | o | O | u | U | n | N |
| Number of total Low tones *(use of Combining Macron Below U+0331 plus base)* | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| Number of total High tones *(composite characters using base)* | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a |
| Number of times the base glyph is used without modification | 744 | 21 | 1232 | 2 | 609 | 22 | 1014 | 2 | 400 | 2 | 596 | 31 |

| English base characters | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | a | A | e | E | i | I | o | O | u | U | n | N |
| Number of total base characters | 744 | 21 | 1232 | 2 | 609 | 22 | 1014 | 2 | 400 | 2 | 596 | 31 |

Chinantec Keyboard layout on an ANSI keyboard



English Keyboard layout on an ANSI keyboard with an ANSI key layout.

---

[25] Me'phaa Keyboard layout on an ANSI physical keyboard with an ISO key layout.



[26] Spanish ISO Keyboard layout

---

[25] Me'phaa Full text heat map

[26] Spanish Full text

## *3.2. Use Case Navajo - English*

*English*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
>*Phoneme Chart*

**Orthography**
>*Orthography Chart*

**Keyboard Layout**
>*Keyboard layout diagram*

**Statistics from Text Analysis**
>*Summary of characters used in the text*
>*Characters not used in the text*
>*Percentage of use*
>*Combinational use of characters*

**Observations - brief**

*Navajo*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
>*Phoneme Chart*

**Orthography**
>*Orthography Chart*

**Keyboard Layout**
>*Keyboard layout diagram*

**Statistics from Text Analysis**
>*Summary of characters used in the text*
>*Characters not used in the text*
>*Percentage of use*
>*Combinational use of characters*

**Observations - brief**

*Section summary*

## *3.3. Use Case Nigerian Languages*

This is tentative on if I can get someone from Nigeria group to give me the text and the keyboard layout.

*English*

## Social Use setting of typing in the language
*Mulitlanguage use - current contexts?*

## Phonology
    *Phoneme Chart*

## Orthography
    *Orthography Chart*

## Keyboard Layout
    *Keyboard layout diagram*

## Statistics from Text Analysis
    *Summary of characters used in the text*
    *Characters not used in the text*
    *Percentage of use*
    *Combinational use of characters*

## Observations - brief

*Ezea*

## Social Use setting of typing in the language
*Mulitlanguage use - current contexts?*

## Phonology
    *Phoneme Chart*

## Orthography
    *Orthography Chart*

## Keyboard Layout
    *Keyboard layout diagram*

## Statistics from Text Analysis
    *Summary of characters used in the text*
    *Characters not used in the text*
    *Percentage of use*
    *Combinational use of characters*

## Observations - brief

*Okphela*

## Social Use setting of typing in the language
*Mulitlanguage use - current contexts?*

## Phonology
    *Phoneme Chart*

## Orthography
    *Orthography Chart*

## Keyboard Layout
    *Keyboard layout diagram*

## Statistics from Text Analysis
    *Summary of characters used in the text*
    *Characters not used in the text*
    *Percentage of use*
    *Combinational use of characters*

## Observations - brief

*Cishingini*

**Social Use setting of typing in the language**

*Mulitlanguage use - current contexts?*

**Phonology**

> *Phoneme Chart*

**Orthography**

> *Orthography Chart*

**Keyboard Layout**

> *Keyboard layout diagram*

**Statistics from Text Analysis**

> *Summary of characters used in the text*
> *Characters not used in the text*
> *Percentage of use*
> *Combinational use of characters*

**Observations - brief**

*Bekwarra*

**Social Use setting of typing in the language**

*Mulitlanguage use - current contexts?*

**Phonology**

> *Phoneme Chart*

**Orthography**

> *Orthography Chart*

**Keyboard Layout**

> *Keyboard layout diagram*

**Statistics from Text Analysis**

> *Summary of characters used in the text*
> *Characters not used in the text*
> *Percentage of use*
> *Combinational use of characters*

**Observations - brief**

*Ibgo*

**Social Use setting of typing in the language**

*Mulitlanguage use - current contexts?*

**Phonology**

> *Phoneme Chart*

**Orthography**

> *Orthography Chart*

**Keyboard Layout**

> *Keyboard layout diagram*

**Statistics from Text Analysis**

> *Summary of characters used in the text*
> *Characters not used in the text*
> *Percentage of use*
> *Combinational use of characters*

**Observations - brief**

*Section summary*

## *3.4. Use Case European languages*

*English*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
> *Phoneme Chart*

**Orthography**
> *Orthography Chart*

**Keyboard Layout**
> *Keyboard layout diagram*

**Statistics from Text Analysis**
> *Summary of characters used in the text*
> *Characters not used in the text*
> *Percentage of use*
> *Combinational use of characters*

**Observations - brief**

*French*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
> *Phoneme Chart*

**Orthography**
> *Orthography Chart*

**Keyboard Layout**
> *Keyboard layout diagram*

**Statistics from Text Analysis**
> *Summary of characters used in the text*
> *Characters not used in the text*
> *Percentage of use*
> *Combinational use of characters*

**Observations - brief**

*German*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
> *Phoneme Chart*

**Orthography**
> *Orthography Chart*

**Keyboard Layout**
> *Keyboard layout diagram*

**Statistics from Text Analysis**
> *Summary of characters used in the text*
> *Characters not used in the text*

*Percentage of use*
*Combinational use of characters*

**Observations - brief**

*Itialian*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
*Phoneme Chart*
**Orthography**
*Orthography Chart*
**Keyboard Layout**
*Keyboard layout diagram*
**Statistics from Text Analysis**
*Summary of characters used in the text*
*Characters not used in the text*
*Percentage of use*
*Combinational use of characters*
**Observations - brief**

*Polish*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
*Phoneme Chart*
**Orthography**
*Orthography Chart*
**Keyboard Layout**
*Keyboard layout diagram*
**Statistics from Text Analysis**
*Summary of characters used in the text*
*Characters not used in the text*
*Percentage of use*
*Combinational use of characters*
**Observations - brief**

*Section Summary*

## *3.5. Use Case Mongolia*

*Russian*

**Social Use setting of typing in the language**
*Mulitlanguage use - current contexts?*
**Phonology**
*Phoneme Chart*
**Orthography**
*Orthography Chart*
**Keyboard Layout**

*Keyboard layout diagram*

**Statistics from Text Analysis**

    *Summary of characters used in the text*

    *Characters not used in the text*

    *Percentage of use*

    *Combinational use of characters*

**Observations - brief**

*Mongolian*

**Social Use setting of typing in the language**

*Mulitlanguage use - current contexts*

**Phonology**

    *Phoneme Chart*

**Orthography**

    *Orthography Chart*

**Keyboard Layout**

    *Keyboard layout diagram*

**Statistics from Text Analysis**

    *Summary of characters used in the text*

    *Characters not used in the text*

    *Percentage of use*

    *Combinational use of characters*

**Observations - brief**

*Section summary*

# 4. Discussion and Application

## 4.1.  Frameworks for applied design

*Framework for keyboard layout research in minority languages*

This study, entitled The Application of Linguistic Knowledge to the Keyboard Layout Design Process for Minority Languages, offers (1) a framework for evaluating keyboard layout designs as they are created for minority language communities, (2) a model for testing a particular user group's reception of a given keyboard layout, and (3) a framework for engagement and distribution of digital products for text entry.

*Framework for digital viability of a language - EGIDS+*

What *things* are needed in a language to make it viable in the digital sphere?

What are the steps needed to reach EGIDS level 5 in a SUM model?

Overview of other keyboard layout work

(Harvey 2013)

*Framework for aligning considerations in the design of keyboard layouts*

## *4.2.  Discussion of relevant application of results*

### *Application in Language Documentation*

There are 3 reasons why linguists should pay attention to Keyboard layouts:
1.  Keyboarding is an expression of language and reveals some very unique ways that the human body expresses communicative thought.
2.  the language teacher should consider the use of language in all mediums not just oral
3.  language development : linguist :OT/Constructions based on cognition and synapses, Language documenter, technologists.

### *Application in Language Development*

### *Orthographies with overt marking of Tone- Cognitive effect measurement of orthographies*

(Roberts 2009, 2011) (Crofts 1976) (Bird 1999, 2001) (Hollenbach 2007) (Kutsch Lojenga 2011) (Zaretsky, Kraljevic, Core & Lencek 2009)

On overtly marking tone
<--Note from David Roberts -->

### *Reading - Typing activation of spelling instead of writing activation of spelling*

### *Writing - The way characters are composed has an effect on the understanding of the language*

Text input methods in the twenty-first century have the potential to entice or discourage speakers of minority languages in the use of their languages in digital mediums. Popular and prolific writing systems in the twenty-first century are crucially tied to electronic input methods that can be easily used in communicating ideas in written orthographic form including the production of printed or electronic reading material. Although, literacy systems have always contained a human interaction element as part of the writing method, in the twenty-first century these methods involve an electronic-tactile medium, e.g. the keyboard. Twenty-first century literacy in social contexts (or as social practice) almost always contains a digital element, e.g. SMS, e-mail, web-forums, Twitter, Facebook, etc.

Keyboard layout design is an important stepping stone to linguistic expression in the digital age. The production of written language is dependent on cognitive processes that access language competencies. This linguistic information is embodied in the message (discourse and syntax) and encoded via the orthography (phonological and lexical information). The design of orthographies and how these orthographies encode linguistic information affect the mechanics of language expression in written form.

### *Application in Cognitive linguistics*

### **Typing as communicative Gesture**

Content from Mondana:

Face-to-face interactions

- the dominant mode of every day interaction throughout human history
(Levinson & Brown)

Wilkins 1999, 2004

Classification of gesture.  Mueller 1998, Efron 1972, Ekman & Friesen 1969,

iconics
- Depicting aspects of a concrete
Metaphorics

Deictics

Beats

Semantics

## Gesture as Language
## Impacts of typing on the understanding of language

Functional load
What is Functional load and is it called upon in the text input process?

## *4.3.  Future direction*

# 5. Bibliography

2005. Chinese. In Sarmad Hussain, Nadir Durrani & Sana Gul (eds.), *PAN Localization: Survey of Language Computing in Asia*, 37-46. Lahore, Pakistan: Center for Research in Urdu Language Processing; National University of Computer and Emerging Sciences.

2010. Jemes. *Bible Nso*. Lagos, Nigeria: Bible Society of Nigeria.

Afzal, M. & S. Hussain. 2001. Urdu computing standards: development of Urdu Zabta Takhti (UZT) 1.01. Paper presented at Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century. Proceedings. IEEE International,2001.

Ajo, Frances, Valérie Guérin, Ryoko Hattori & Laura C. Robinson. 2010. Native speakers as documentors: A student initiative at the University of Hawai'i at Mānoa. In Lenore Grenoble & N. Louanna Furbee (eds.), *Language Documentation: Practice and vaules*, 275-88. Amsterdam / Philadelphia: John Benjamins Publishing Company.

al-Ashanti, 'AbdulHaq. 2013. Boko Harām - In Nigeria: Its Beginnings, Principles and Activities in Nigeria. SalafiManhaj.com. <Accessed: 11. June 2014>. http://download.salafimanhaj.com/pdf/SalafiManhaj_BokoHaram.pdf

Alden, David G., Richard W. Daniels & Arnold F. Kanarick. 1972. Keyboard Design and Operation: A Review of the Major Issues. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 14.4: 275-93. http://hfs.sagepub.com/content/14/4/275.abstract

Ali, A. Yusuf. 1920. Note on Urdu Orthography. *Bulletin of the School of Oriental Studies, University of London* 1.3: 29-34. http://www.jstor.org/stable/607588

American Bible Society. 2000. James. *Navajo Bible*. American Bible Society.

Amery, Rob. 2009. Phoenix or Relic? Documentation of Languages with Revitalization in Mind. *Language Documentation & Conservation* 3.2: 138-48. http://hdl.handle.net/10125/4436

Anderson, Allison M., Gary A. Mirka, Sharon M. B. Joines & David B. Kaber. 2009. Analysis of Alternative Keyboards Using Learning Curves. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 11.

Asociación para la Promoción de Lecto-Escritura Tlapaneca. 1988. *Xó nitháán mè'phàà: Cómo se escribe el tlapaneco*. Malinaltepec, Guerrero: Asociación para la Promoción de Lecto-Escritura Tlapaneca.

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

Barkhuus, Louise & Valerie E. Polichar. 2011. Empowerment through seamfulness: smart phones in everyday life. *Personal & Ubiquitous Computing* 15.6: 629-39.

Barton, David. 1994. *Literacy: an introduction to the ecology of written language*. Oxford: Blackwell.

Bergqvist, Henrik. 2012. Review of Language Documentation: Practice and Values. *Himalayan Linguistics* 11.1: 24-7. <Accessed: 11. April 2014>.

Bernard, H. Russell. 1996. Language preservation and publishing. In Nancy Hornberger (ed.), *Indigenous literacies in the Americas: Language planning from the bottom up* (Contributions to the Sociology of Language 75), 139-56. Berlin: Walter de Gruyter.

von Bertalanffy, Ludwig. 1968. *General System Theory: Foundations, Development, Applications*, Revised edn. George Braziller, Inc.

Bible Society of Mongolia. 2014. The Book of James. *Bibli [Bible in Mongolian]*. Ulaanbaatar, Mongolian: Bible Society of Mongolia.

Biblica, Inc. 2002. Jakobus. *Hoffnung für Alle*. Basel; Gießen: Brunnen Verlag.

Bielenberg, Brian. 1999. Indigenous language codification: Cutural effects. In Jon Reyhner, Gina Cantoni, Robert N. St. Clair & Evangeline Parsons Yazzie (eds.), *Revitalizing Indigenous Languages*, 103-12. Flagstaff, AZ: Northern Arizona University.

Bird, Steven. 1999. When Marking Tone Reduces Fluency: An Orthography Experiment in Cameroon. *Language and Speech* 42.1: 83-115. http://las.sagepub.com/content/42/1/83.abstract

Bird, Steven. 2001. Orthography and identity in Cameroon. *Written Language & Literacy* 4.2: 131-62. http://dx.doi.org/10.1075/wll.4.2.02bir

Bird, Steven & Gary F. Simons. 2003. Seven dimension of portability for language documentation and description. *Language* 79.3: 557-82. http://www.jstor.org/stable/4489465

Birken-Silverman, Gabriele. 1997. The Role of Dialects in Language Planning, Codification and Standardization: The Case of Corsican. *Dialectologia et Geolinguistica* 1997.5: 31-64. http://dx.doi.org/10.1515/dig. 1997.1997.5.31

Black, Cheryl A. 2005. An autosegmental analysis of Meʔphaa (Tlapanec) noun inflection. In Rosemary Beam de Azcona & Mary Paster (eds.), *Conference on Otomanguean and Oaxacan Languages, March 19-21, 2004* (Survey of California and other Indian languages Report 13), 1-14. Berkeley: University of California at Berkeley.

Blench, Rodger. 2012. Research and development of Nigerian minority languages. In Rodger Blench & Stuart McGill (eds.), *Advances in Minority Language Research in Nigeria*, vol. 1 (Kay Williamson Educational Foundation African Language Monographs 5), 1-17. Köln: Rüdiger Köppe Verlag.

Boerger, Brenda H. 2007. Natqgu Literacy: Capturing Three Domains for Written Language Use. *Langauge Documentation & Conservation* 1.2: 126–53. <Accessed: 04. May 2014>. http://hdl.handle.net/10125/1601

Borchers, Jan O. 2000. Interaction Design Patterns: Twelve Theses.

Bowern, Claire. 2008. *Linguistic fieldwork : a practical guide*. Houndmills, Basingstoke, Hampshire [England]; New York: Palgrave Macmillan.

Brinckwirth, Anton T. 2012. *Implementation and Outcomes of an Online English-Portuguese Tandem Language Exchange Program Delivered Jointly Across a U.S.-Brazilian University Partnership: A Case Study*. Ph.D dissertation, Virginia Commonwealth University, Richmond, Virginia.

Brownie, John. 2012. Ukelele: Mac OS X Keyboard Layout Editor. Version: 2.1.10. Computer program. <Accessed: 12 June 2012>. http://scripts.sil.org/ukelele

Cahill, Michael C. 2011. Non-Linguistic Factors in Orthographies. In Ethan Poole (ed.), Final proposal for Developing Orthographies for Unwritten Languages 3 hour Symposium, Organizers: Michael Cahill (SIL International) and Keren Rice (University of Toronto) (Series), 5-6. Linguistic Society of America. <Accessed: 30 June 2012>. http://www.lsadc.org/info/documents/2011/annual-meeting/final-proposals/cahill-rice.pdf

Cahill, Michael C. & Elke Karan. 2008. Factors in Designing Effective Orthographies for Unwritten Languages. *SIL Eletronic Working Papers* 001. Dallas, Tx.: SIL International. <Accessed: 14 January 2013>. http://www.sil.org/silewp/abstract.asp?ref=2008-001

Cahill, Michael C. & Keren Rice (eds.). 2014. *Developing orthographies for unwritten languages* (Publications in Language Use and Education 6). Dallas, Tx: SIL International, Global Publishing.

Carrasco Zúñiga, Estanislao & Mark L. Weathers. 2008-2010. Santiago (James). Ms., Pre-Publication Draft of Bible Portion.

Chang, Debbie. 2010. *TAPS: Checklist for Responsible Archiving of Digital Language Resources*. M.A. dissertation, Graduate Institute of Applied Linguistics, Dallas, Tx.

Chelliah, Shobhana Lakshmi & Willem Joseph de Reuse. 2011. *Handbook of descriptive linguistic fieldwork*. Dordrecht; New York: Springer.

Chen, Tianyi, Yeliz Yesilada & Simon Harper. 2010. What input errors do you experience? Typing and pointing errors of mobile Web users. *International Journal of Human-Computer Studies* 68.3: 138-57. http://www.sciencedirect.com/science/article/pii/S1071581909001578

Christensen, Carol A. 2004. Relationship between orthographic-motor integration and computer use for the production of creative and well-structured written text. *British Journal of Educational Psychology* 74.4: 551-64.

Clifton, John M. 1994. Stabel Multilingualism in a Small Language Group: The case fo Kaki Ae. *Language and Linguistics in Melanesia* 25.1: 107-24.

Clifton, John M. 2013. Orthography as Social Practice: Is There a Role for the Linguist? In Nicholas Ostler (ed.), *Proceedings of the 17th FEL Conference*, 189-90. England: Ogmios : Foundation For Endangered Languages.

Cline, Kevin L. 2013. *The Tone System of Acatepec Me'phaa*. M.A. dissertation, University of North Dakota, Grand Forks.

Coleman, Ross. 2011. Field, File, Data, Conference: Towards New Modes of Scholarly Publication. In Nick Thieberger, Linda Barwick, Rosey Billington & Jill Vaughan (eds.), *Sustainable data from digital research: Humanities perspectives on digital scholarship*, 57-68. Melbourne: Custom Book Centre, University of Melbourne.

Constable, Peter. 2001. Understanding characters, keystrokes, codepoints and glyphs: encoding and working with multilingual text. In Melinda Lyons (ed.), *Implimenting Writting Systems: an Introduction*, 9-33. Dallas: SIL International.

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

Crasborn, Onno. 2010. What Does "Informed Consent" Mean in the Internet Age?: Publishing Sign Language Corpora as Open Content. *Sign Language Studies* 10.2: 276-90. http://muse.jhu.edu/journals/sign_language_studies/v010/10.2.crasborn.html

Crasborn, Onno, Micha Hulsbosch, Lari Lampen & Han Sloetjes. 2014. New Multilayer Concordance Functions in ELAN and TROVA. *Proceedings of the Tilburg Gesture Research Meeting [TiGeR 2013]*.

Crofts, Marjorie. 1976. Must tone always be written in a tonal language? *The Bible Translator* 27. 127-34.

Csató, Éva Á. & David Nathan. 2003. Multimedia and the documentation of endangered languages. In Peter K. Austin (ed.), *Language Documentation and Description* (1), 73-84. London: The Hans Rausing Endangered Languages Project, School of Oriental and African Studies.

David, Paul A. 1985. Clio and the Economics of QWERTY. *The American Economic Review* 75.2: 332-7. http://www.jstor.org/stable/1805621

Dealey, William L. & August Dvorak. 1933. Why and Wherefore of Typewriting Errors. *The Journal of Business Education* 9.3: 19-20. <Accessed: 2013/01/05>. http://dx.doi.org/10.1080/08832323.1933.9959010

DeJong, David H. 1998. Is Immersion the Key to Language Renewal? *Journal of American Indian Education* 37.3: 31-46. <Accessed: 10. June 2014>. http://jaie.asu.edu/v37/V37S3imm.htm

Detey, Sylvain & Jean-Luc Nespoulous. 2008. Can orthography influence second language syllabic segmentation?: Japanese epenthetic vowels and French consonantal clusters. *Lingua* 118.1: 66-81. http://www.sciencedirect.com/science/article/pii/S0024384107000733

Dobrin, Lise M., Peter K. Austin & David Nathan. 2007. Dying to be counted: the commodification of endangered languages in documentary linguistics. In Peter K. Austin, Oliver Bond & David Nathan (eds.), *Proceedings of Conference on Language Documentation and Linguistic Theory*, 59-68. London: SOAS.

Dodd, Vikram. 2013. Youth crime commissioner Paris Brown stands down over Twitter row of Work. In *The Guardian*. Online: theguardian.com. <Accessed: 10. June 2014>. http://www.theguardian.com/uk/2013/apr/09/paris-brown-stands-down-twitter/

Dorian, Nancy C. 1982. Defining the speech community to include its working margins. In Suzanne Romaine (ed.), *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold.

Dorian, Nancy C. 1987. The value of language-maintenance efforts which are unlikely to succeed. *International Journal of the Sociology of Language* 1987.68: 57-67. <Accessed: 05. May 2014>. http://www.degruyter.com/view/j/ijsl.1987.issue-68/ijsl.1987.68.57/ijsl.1987.68.57.xml

Dorian, Nancy C. 1993. A Response to Ladefoged's Other View of Endangered Languages. *Language* 69.3: 575-9. http://www.jstor.org/stable/416699

Dul, Jan, Henk de Vries, Sandra Verschoof, Wietske Eveleens & Albert Feilzer. 2004. Combining economic and social goals in the design of production systems by using ergonomics standards. *Computers & Industrial Engineering* 47.2–3: 207-22. http://www.sciencedirect.com/science/article/pii/S0360835204001196

Duncan, Daniel. 2013. Whistles in the Mist: Whistled Speech in Oaxaca. p. Video Run time: 26:46. Vimeo.com: The Southwest Center, University of Arizona. <Accessed: 23. April 2014>. http://vimeo.com/57291304

Duncan, Joan & D. Ferguson. 1974. Keyboard Operating Posture and Symptoms in Operating. *Ergonomics* 17.5: 651-62. <Accessed: 2013/07/02>. http://dx.doi.org/10.1080/00140137408931404

Dvorak, August. 1935. Developing Rhythm in Typewriting. *The Journal of Business Education* 10.6: 19-34. <Accessed: 2013/01/10>. http://dx.doi.org/10.1080/08832323.1935.9958667

Dwyer, Arienne M. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation* (Trends in Linguistics Studies and Monographs 178), 31-66. Berlin: Mouton de Gruyter.

Dwyer, Arienne M. 2010. Models of suddessful collaboration. In Lenore Grenoble & N. Louanna Furbee (eds.), *Language Documentation: Practice and vaules*, 3-24. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Dwyer, Arienne M. 2012. Tools and techniques for endangered-language assessment and revitalization. In Trace Foundation (ed.), *Minority Language in Today's Global Society*, 2 vols, vol. 1 (Trace Foundation Lecture Proceeding). New York: Trace Foundation.

Editorial. 2013. The maze of impact metrics. *Nature* 502.7471: 271. http://www.nature.com/news/the-maze-of-impact-metrics-1.13952

Eisenlohr, Patrick. 2004. Language Revitalization and New Technologies: Cultures of Electronic Mediation and the Refiguring of Communities. *Annual Review of Anthropology* 33.1: 21-45. http://www.jstor.org/stable/25064844

Elia, Antonella. 2006. Language learning in tandem via Skype. *Reading Matrix: An International Online Journal* 6.3: 269-80. http://connection.ebscohost.com/c/articles/24393892/language-learning-tandem-via-skype

Esizmetor, David Oshorenoya. 2009. What Orthography for Naijá? Paper presented at Conference on Nigerian Pidgin, University of Ibadan, Nigeria. http://www.ifra-nigeria.org/IMG/pdf/Esizimetor_2010b.pdf

Esposito, John L. (ed.) 1995. *Women. In The Islamic World: Past and Present* (Oxford Islamic Studies Online). Oxford, UK: Oxford University Press.

Fagarasanu, Mircea, Shrawan Kumar & Yogesh Narayan. 2005. The training effect on typing on two alternative keyboards. *International Journal of Industrial Ergonomics* 35.6: 509-16. http://www.sciencedirect.com/science/article/pii/S016981410400215X

Ferguson, Charles Albert. 1959. Diglossia. *Word : journal of the Linguistic Circle of New York.* 15.1: 325-40.

Ferguson, Charles Albert. 1968. Language development. In A. Fishman Joshua, Charles Albert Ferguson & Jyotirindra Das Gupta (eds.), *Language problems of developing nations*, 27-35. New York: John Wiley & Sons, Inc.

Ferrara, Kathleen, Hans Brunner & Greg Whittemore. 1991. Interactive Written Discourse as an Emergent Register. *Written Communication* 8.1: 8-34. http://wcx.sagepub.com/content/8/1/8.abstract

Fishman, Joshua A. 1964. Language Maintenance and Language Shift as a Field of Inquiry. A Definition of The Field and Suggestions for its Further Development. *Linguistics* 2.9: 32. http://www.degruyter.com/view/j/ling.1964.2.issue-9/ling.1964.2.9.32/ling.1964.2.9.32.xml

Fishman, Joshua A. 1967. Bilingualism With and Without Diglossia; Diglossia With and Without Bilingualism. *Journal of Social Issues* 23.2: 29-38. http://dx.doi.org/10.1111/j.1540-4560.1967.tb00573.x

Fishman, Joshua A. 1991. *Reversing language shift theoretical and empirical foundations of assistance to threatened languages* (Multilingual Matters 76). Clevedon: Multilingual Matters, Ltd.

Fishman, Joshua A. 2013. Language maintenance and language shift as a field of inquiry: A definition of the field and suggestions for its further development. *Linguistics* 51.Jubilee: 9-10.  <Accessed: 29. April 2014>. http://www.degruyter.com/view/j/ling.2013.51.issue-jubilee/ling-2013-0038.xml

Fitts, Paul M. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47.6: 381-91.

Foris, David Paul. 1973. Sochiapan Chinantec Syllable Structure. *International Journal of American Linguistics* 39.4: 232-5. http://www.jstor.org/stable/1264825

Foris, David Paul. 1994. *A grammar of Sochiapan Chinantec*. Ph.D. dissertation, The University of Auckland (New Zealand), New Zealand.

Foris, David Paul. 2000. Developing a Sochiapan Chinantec orthography: linguistic factors and sociolinguistic results. Paper presented at Bilingualism at the Ends of the Earth, Depatment of General and Applied Linguistics, University of Waikato, Hamilton, New Zealand, (24th -26th November 2000). <Accessed:

Forsyth, Richard S. & Serge Sharoff. 2014. Document dissimilarity within and across languages: A benchmarking study. *Literary and Linguistic Computing* 29.1: 6-22. http://llc.oxfordjournals.org/content/early/2013/02/05/llc.fqt002.abstract

Fortis, David. 2000. Developing a Sochiapan Chinantec orthography: linguistic factors and sociolinguistic results. Paper presented at Bilingualism at the Ends of the Earth, Depatment of General and Applied Linguistics, University of Waikato, Hamilton, New Zealand, (24th -26th November 2000). <Accessed:

Fortunati, Leopoldina & Jane Vincent. 2014. Sociological insights on the comparison of writing/reading on paper with writing/reading digitally. *Telematics and Informatics* 31.1: 39-51. http://www.sciencedirect.com/science/article/pii/S0736585313000087

Furbee, N. Louanna. 2010. Language Documentation: Theory and practice. In Lenore Grenoble & N. Louanna Furbee (eds.), *Language Documentation: Practice and vaules*, 3-24. Amsterdam / Philadelphia: John Benjamins Publishing Company.

van Galen, Gerard P. 1991. Handwriting: Issues for a psychomotor theory. *Human Movement Science* 10.2–3: 165-91. http://www.sciencedirect.com/science/article/pii/016794579190003G

Garrett, Andrew. 2011. An Online Dictionary with Texts and Pedagogical Tools: The Yurok Language Project at Berkeley. *International Journal of Lexicography* 24.4: 405-19. http://ijl.oxfordjournals.org/content/24/4/405.abstract

Garrett, Jesse James. 2011. *The elements of user experience : user-centered design for the Web and beyond*. Berkeley, CA: New Riders.

General Assembly of the State of North Carolina. 1831. An Act to Prevent all Persons From Teaching Slaves to Read or Write, The use of Figures Excepted. *Acts passed by the General Assembly of the State of North Carolina [1830-1831]*, 11. Raleigh, North Carolina: Lawrence & Lemay.

Gonzalez, Andrew. 1999. The Language Planning Situation in the Philippines. In Robert B. Kaplan & Richard B. Jr Baldauf (eds.), *Language Planning in Malawi, Mozambique and the Philippines* (Multilingual Matters 113), 133-71. UK: Multilingual Matters, Ltd.

Goossens, Michel , S. P. Q. Rahtz & Frank Mittelbach. 1997. *The LaTex graphics companion : illustrating documents with TeX and PostScript* (Addison-Wesley series on tools and techniques for computer typesetting). Reading, Ma.: Addison-Wesley.

Grenoble, Lenore A. & Lindsay J. Whaley. 2006. Orthography. *Saving languages : an introduction to language revitalization*, 137-59. Cambridge University Press.

Gudschinsky, Sarah C. 1957. *Handbook of Literacy*, rev. edn. Santa Ana, CA: Summer Institute of Linguistics.

Guérin, Valérie. 2008. Writing an endangered language. *Language Documentation & Conservation* 2.1: 47–67. http://hdl.handle.net/10125/1804

Hale, Ken, et al. 1992. Endangered Languages. *Language* 68.1: 1-42. http://www.jstor.org/stable/416368

Hallissy, Bob. 16 November 2005. Unicode Character Count utility. Version: 0.3. Computer program. <Accessed: 30 December 2012>. http://scripts.sil.org/UnicodeCharacterCount

Hammer, Anita. 2005. Weaving Trickster: Myth and Tribal Encounters on the World Wide Web. In Eric W. Rothenbuhler & Mihai Coman (eds.), *Media Anthropology*, 260-8. Thousand Oaks; London; New Delhi: SAGE Publications Ltd.

Harvey, Chris. 2013. Languagegeek.com. <Accessed: 23 January 2013>. http://www.languagegeek.com/keyboard_general/all_keyboards.html

Hassane, Moulaye. 2008. Ajami in Africa: the use of Arabic script in the transcription of African languages. In Shamil Jeppie & Souleymane Bachir Diagne (eds.), *The Meanings of Timbuktu*, 109-22. Cape Town, South Africa: CODESRIA / HSRC Press.

Himmelmann, Nikolaus P. 1998. Documentary and Descriptive Linguistics (full version). *Linguistics* 36.1: 161-95. <Accessed: 2 January 2011>. http://www.uni-muenster.de/imperia/md/content/allgemeine_sprachwissenschaft/dozenten-unterlagen/himmelmann/linguistics98.pdf

Hollenbach, Barbara E. 2007. Tone languages and language (un)awareness. *SIL Forum for Language Fieldwork* 2007.004: http://www.sil.org/resources/publications/entry/7773

Hunsinger, Jeremy, Lisbeth Klastrup & Matthew Allen (eds.). 2010. *International Handbook of Internet Research*. Dordrecht; Heidelberg; London; New York: Springer.

International Organization for Standardization. 2007. ISO 639-3:2007 Codes for the representation of names of languages. Part 3: Alpha-3 code for comprehensive coverage of languages. ISO 639-3:2007. <Accessed: 12 June 2012>. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39534

Jany, Carmen. 2010a. Issues and challenges in the orthography development of Chuxnabán Mixe. Paper presented at Conference of Endangered Languages and Cultures of Native America 2010, University of Utah in Salt Lake City,  April 9-10, 2010.  <Accessed: 12 June 2012>. http://www.humis.utah.edu/humis/docs/organization_919_1300140441.pdf

Jany, Carmen. 2010b. Orthography Design for Chuxnabán Mixe. *Language Documentatlon & Conservatlon* 4.1: 231-53.  <Accessed: 12 June 2012>. http://hdl.handle.net/10125/4481

Jucker, Andreas H. & Christa Dürscheid. 2012. The Linguistics of Keyboard-to-screen Communication. A New Terminological Framework. *Linguistik online* 56.6: 39-64. <Accessed: 07. June 2014>. https://bop.unibe.ch/linguistik-online/article/view/255

Kazem, Ali Asghar Pourhaji & Kian Nimgaz Naghsh. 2011. Attaining a suitable Persian Keyboad layout using an evolutionary strategy. *Proceedings of the 2011 International Confrence on Instrumentation, Measurement, Circuits and Systems (ICIMCS 2011), 12-13 December 2011*, 3 vols, vol. 3: Computer-Aided Design, Manufacturing and Management, 451-4. Hong Kong, China: ASME.

Kiesler, Sara, Jane Siegel & Timothy W. McGuire. 1984. Social psychological aspects of computer-mediated communication. *American Psychologist* 39.10: 1123-34.

Kilgarriff, Adam, et al. 2014. Corpus-based vocabulary lists for language learners for nine languages. *Language Resources and Evaluation* 48.1: 121-63. http://dx.doi.org/10.1007/s10579-013-9251-2

Kim, Kyongsok. 1992. A common approach to designing the Hangul code and keyboard. *Computer Standards & Interfaces* 14.4: 297-325. http://www.sciencedirect.com/science/article/pii/092054899290042C

King, Ross. 1996. Korean Writing. In Peter T. Daniels & William Bright (eds.), *The World's Writing Systems*, 218-27. New York; Oxford: Oxford University Press.

Kipp, Darrell. 2007. Swimming in words. *Cultural Survival Quarterly* 31.2: 36-43.

Kluge, Angela. 2007. The Gbe Language Continuum of West Africa: A Synchronic Typological Approach to Prioritizing In-depth Sociolinguistic Research on Literature Extensibility. *Langauge Documentation & Conservation* 1.2: 82-2 5. http://hdl.handle.net/10125/1714

Kuhlmann, Annette. 1992. Collaborative Research among the Kickapoo Tribe of Oklahoma. *Human Organization* 51.3: 274-83. http://www.metapress.com/content/K3N65X64T5NX36X5

Kutsch Lojenga, Constance. 2011. Orthography and tone. Paper presented at Linguistic Society of America Annual Meeting, Pittsburgh, Jan. 6-9, 2011, Symposium on Developing Orthographies for Unwritten Languages. <Accessed: 30 June 2012>. http://www.sil.org/linguistics/2011LSASymposium/KutschLojenga.html

La Liga Bíblica. 2009. Sí² Quioh²¹ Santiago. JÚ¹ CHÚ³² QUIOH²¹ JESÚS TSÁ² LŧŃ³ CRISTO (El Nuevo Testamento en el chinanteco de Sochiapan) (Series), 525-33. La Liga Bíblica. <Accessed: 12 June 2012>. http://www.scriptureearth.org/data/cso/PDF/00-WNTcso-web.pdf

Labov, William. 1966. *The social stratification of English in New York City*. Washington: Center for Applied Linguistics.

Labov, William. 1972. *Sociolinguistic patterns* (University of Pennsylvania publications in conduct and communication 4). Philadelphia: University of Pennsylvania Press.

Labov, William. 2006. *The social stratification of English in New York City*. Cambridge [UK]; New York: Cambridge University Press.

Ladefoged, Peter. 1992. Another View of Endangered Languages. *Language* 68.4: 809-11. http://www.jstor.org/stable/416854

Laponce, J. A. 2004. Minority Languages and Globalization. *Nationalism and Ethnic Politics* 10.1: 15-24. <Accessed: 2014/04/06>. http://dx.doi.org/10.1080/13537110490450755

Lee, Joonseok & R. I. McKay. 2010. Comparing Typing Speeds for Keyboard Layouts Fairly and Efficiently. <Accessed: 09. June 2014>. http://www.cc.gatech.edu/~jlee716/papers/hci.pdf

Lee, Tiffany. S. & Daniel McLaughlin. 2001. Reversing Navajo language shift, revisited. In A. Fishman Joshua (ed.), *Can threatened languages be saved? Reversing language shift, revisited: A 21st century perspective* (Multilingual matters 116), 23-43. Clevedon, UK: Multilingual Matters Ltd.

Leonard, Wesley Y. & Erin Haynes. 2010. Making "collaboration" collaborative: An examination of perspectives that frame linguistic field research. *Langauge Documentation & Conservation* 4.1: 268-93. <Accessed: 29. April 2014>. http://hdl.handle.net/10125/4482

Lewis, M. Paul. 2010a. The Sustainable Use Model and the Expanded Graded Intergenerational Disruption Scale (EGIDS). Paper presented at International Language Assessment Community Meetings 2010 (ILAC VI), Penang, Malaysia. <Accessed: 18. April 2014>.

Lewis, M. Paul. 2010b. The Sustainable Use Model for Language Development. Paper presented at Payap University Linguistics Colloquium, Payap University, Chaing Mai Thailand, (22 September 2010). <Accessed: 18. April 2014>.

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

Lewis, M. Paul & Gary F. Simons. Forthcoming. *Perspectives on Community-Based Language Development*. Dallas, Tx.: SIL International.

Lexander, Kristin Vold. 2011. Texting and African language literacy. *New Media & Society* 13.3: 427-43. http://nms.sagepub.com/content/13/3/427.abstract

Liu, Ziming. 2005. Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of Documentation* 61.6: 700-12.

Llamzon, Teodoro A. 1977. A Requiem for Pilipino. In Bonifacio P. Sibayan & Andrew B. Gonzalez (eds.), *LanguagePlanning and the Building of a National Language: Essays in Honor of Santiago A.Fonacier on his Ninety-second Birthday*, 291-303. Manila: Linguistic Society of the Philippines.

Lüpke, Friederike. 2011. Orthography development. In Peter Austin & Julia Sallabank (eds.), *Handbook of endangered languages*, 312-36. Cambridge: Cambridge University Press.

Lyons, Melinda (ed.) 2001. *Implementing Writing Systems: an Introduction*, Preliminary edn. Dallas: SIL International.

MacKenzie, I. Scott. 2002. KSPC (keystrokes per character) as a characteristic of text entry techniques. *Proceedings of the Fourth International Symposium on Human Computer Interaction with Mobile Devices*, 195-210. Heidelberg, Germany: Springer-Verlag.

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

MacKenzie, I. Scott. 2013. Human-computer interaction an empirical research perspective http://public.eblib.com/EBLPublic/PublicView.do?ptiID=1110719

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.

MacKenzie, I. Scott & Kumiko Tanaka-Ishii. 2007. *Text entry systems mobility, accessibility, universality*. Amsterdam; Boston: Morgan Kaufmann.

MacKenzie, I. Scott & Shawn X. Zhang. 2001. An empirical investigation of the novice experience with soft keyboards. *Behaviour & Information Technology* 20.6: 411-8. <Accessed: 2013/01/06>. http://dx.doi.org/10.1080/01449290110089561

MacNeilage, Peter F. 1964. Typing Errors as Clues to Serial Ordering Mechanisms in Language Behaviour. *Language and Speech* 7.3: 144-59. http://las.sagepub.com/cgi/content/abstract/7/3/144

Mahmudweyssi, Parwin , Denise Bailey, Ludwig Paul & Geoffrey Haig. 2012. *The Gorani language of Gawrajū, a village of West Iran: Texts, grammar, and lexicon* (Beiträge zur Iranistik 35). Wiesbaden, Germany: Dr. Ludwig Reichert Verlag.

Maiman, Bruce. 2009. Civil servant fired for anti-Obama and 'racist' comments on Facebook page of Work. In *Examiner.com*. Online: Examiner.com. <Accessed: 10. June 2014>. http://www.examiner.com/article/civil-servant-fired-for-anti-obama-and-racist-comments-on-facebook-page

Major, Roy C. 2008. Transfer in Second Language Phonology: A Review. In Jette G. Hansen Edwards & Mary L. Zampini (eds.), *Phonology and second language acquisition* (Studies in Bilingualism (SiBil) 36), 63-93. Amsterdam; Philadelphia: John Benjamins Publishing.

Mangen, Anne, Bente R. Walgermo & Kolbjørn Brønnick. 2013. Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research* 58.0: 61-8. http://www.sciencedirect.com/science/article/pii/S0883035512001127

Marinaras, Nicholas & Kostas Lyritzis. 1993. Design of an alternative keyboard layout for the Greek language. *International Journal of Human-Computer Interaction* 5.3: 289-310. <Accessed: 2014/06/04>. http://dx.doi.org/10.1080/10447319309526069

Marlett, Stephen A. 2010. Documenting the Me'phaa Genus. Granted by: National Endowment for the Humanities - Documenting Endangered Languages - Fellowships. <Accessed: 21 June 2012>. https://securegrants.neh.gov/publicquery/main.aspx?f=1&gn=FN-50079-10

Marlett, Stephen A. enero 2012a. Las Transcripciones. In Stephen A. Marlett (ed.), Los Archivos Lingüísticos Me'phaa, vol. 13e (Mexico Eletronic Working Papers Series 13e). SIL International. <Accessed: 30 June 2012>. http://www.sil.org/mexico/workpapers/WP013-PDF/MephaaTranscripciones.pdf

Maslamani, Jaber Ali. 2013. Potential Integration of a Computer-Mediated Communication Platform into the Saudi EFL Classroom: A Synthetic Review *International Journal of Humanities and Social Science* 3.15: 73-8. http://www.ijhssnet.com/journals/Vol_3_No_15_August_2013/9.pdf

Maxwell, Judith M. 2010. Training graduate students and community members for native language documentation. In Lenore Grenoble & N. Louanna Furbee (eds.), *Language Documentation: Practice and vaules*, 255-74. Amsterdam / Philadelphia: John Benjamins Publishing Company.

McCarty, Teresa L. 2003. Revitalising Indigenous Languages in Homogenising Times. *Comparative Education* 39.2: 147–63.

Meier, Paul E. 1980. Ẹkwo-Ozhi, Jémusu Deru. *Ogbandzu owhuu, shi l'eka Chileke*. World Home Bible League and Nigeria Bible Translation Trust.

Microsoft. 2012. The Microsoft Keyboard Layout Creator. Version: 1.4. Computer program. <Accessed: 15 May 2012>. http://msdn.microsoft.com/en-us/goglobal/bb964665

Millar, B. & W. W. Glover. 1981. Synthesis of the Devanagari orthography. *International Journal of Man-Machine Studies* 14.4: 423-35. http://www.sciencedirect.com/science/article/pii/S0020737381800466

Mithun, Marianne. 1992. Orthography planning. In William Bright (ed.), *Oxford International Encyclopaedia of Linguistics*, 4 vols, vol. 3, 141-4. New York: Oxford University Press.

Modiano, Marko. 2001. Linguistic imperialism, cultural integrity, and EIL. *ELT Journal* 55.4: 339-47. http://eltj.oxfordjournals.org/content/55/4/339.abstract

Morita, M. 1985. Japanese Text Input System. *Computer* 18.5: 29-35.

Mosel, Ulrike. 2006. Fieldwork and community language work. In Jost Gippert, Himmelmann Nikolaus & Ulrike Mosel (eds.), *Essentials of language documentation* (Trends in Linguistics Studies and Monographs 178), 67-86. Berlin; New York: Mouton de Gruyter.

Mufwene, Salikoko S. 2002. Colonisation, Globalisation, and the Future of Languages in the Twenty-first Century. *International Journal on Multicultural Societies* 4.2: 162-93.

Mumin, Meikal. 2014. The Arabic Script in Africa: Understudied Literacy. In Meikal Mumin & Kees Versteegh (eds.), *The Arabic script in Africa : studies in the use of a writing system* (Studies in Semitic Languages and Linguistics 71), 40-76. Leiden; Boston: Brill.

Murdadā, Ahmad. 2012. Jama'at "Boko Harām": Nasha'tuhā, Mabadi'uhā wa A'māluhā fi Naygeeriyah. *Arabic language Islamic and African studies journal: al-Qira'āt* Rabī' al-Ākhir- Jumādā al-Ākhira 1433 AH/April – June 2012 CE.12: 12-25.

Nandasara, S. T. 2009. From the Past to the Present: Evolution of Computing in the Sinhala Language. *IEEE Annals of the History of Computing* 31.1: 32-45. <Accessed: 07. June 2014>. http://doi.ieeecomputersociety.org/10.1109/MAHC.2009.10

Nandasara, S. T. & Yoshiki Mikami. 2009. Guest Editors' Introduction: Asian Language Processing: History and Perspectives. *IEEE Annals of the History of Computing* 31.1: 4-7. <Accessed: 07. June 2014>. http://doi.ieeecomputersociety.org/10.1109/MAHC.2009.2

Nathan, David. 2010. Sound and unsound practices in documentary linguistics: towards an epistemology for audio. In Peter Austin (ed.), *Language Documentation and Description*, vol. 7, 262-84. London: SOAS.

Norman, Donald A. 1981. Categorization of action slips. *Psychological Review* 88.1: 1-15. http://psycnet.apa.org/journals/rev/88/1/1/

Otsuka, Yuko & Andrew Wong. 2007. Fostering the growth of budding community initiatives: The role of linguists in Tokelauan maintenance in Hawai'i. *Language Documentation & Conservation* 1.2: 240–56. http://hdl.handle.net/10125/1716

Ottenheimer, Harriet Joseph. 2009. Review of Spelling and society. *Langauge Documentation & Conservation* 3.2: 247-52. <Accessed: 11. April 2014>. http://hdl.handle.net/10125/4443

Pandey, Pramod. 2007. Phonology-orthography interface in Devanāgarī for Hindi. *Written Language & Literacy* 10.2: 145-62.

Patrick, Peter L. 1999. The Speech Community: Some Definitions. Paper presented at NWAVE-28, Toronto, Canada, (17 October, 1999). <Accessed: 09. June 2014>. http://orb.essex.ac.uk/lg/lg232/SpeechComDefs.html

Patrick, Peter L. 2008. The Speech Community. In J. K. Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*, 573-97. Malden, MA: Blackwell Publishing Ltd.

Pattamadilok, Chotiga, Laetitia Perre & Johannes C. Ziegler. 2011. Beyond rhyme or reason: ERPs reveal task-specific activation of orthography on spoken language. *Brain and Language* 116.3: 116-24. http://www.sciencedirect.com/science/article/pii/S0093934X10001999

Pearce, Susan M. 1989. Museum Studies in Material Culture. In Susan M. Pearce (ed.), *Objects in Structures*, 47-59. London ; New York: Leicester University Press.

Peck, Michaeleen, Eunice N. Askov & Steven H. Fairchild. 1980. Another Decade of Research in Handwriting: Progress and Prospect in the 1970s. *The Journal of Educational Research* 73.5: 283-98. http://www.jstor.org/stable/27539768

Pecos, Regis & Rebecca Blum-Martinez. 2001. The Key to Cultural Survival: Language Planning and Revitalization in the Pueblo de Cochiti. In Leanne Hinton & Kenneth Locke Hale (eds.), *The Green Book of Language Revitalization in Practice: Toward a Sustainable World*, 75-82. San Diego, California: Academic press.

Penfield, Susan D., et al. 2008. Community collaborations: best practices for North American indigenous language documentation. *International Journal of the Sociology of Language* 2008.191: 187-202. <Accessed: 24 January 2013>. http://www.degruyter.com/view/j/ijsl.2008.2008.issue-191/ijsl.2008.029/ijsl.2008.029.xml

Perks, Martyn. 2003. 'Ethnography' exposed. *Spiked-Online*. <Accessed: 06. April 2014>. http://www.spiked-online.com/Articles/00000006E039.htm

Perre, Laetitia, Chotiga Pattamadilok, Marie Montant & Johannes C. Ziegler. 2009. Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research* 1275. 73-80. http://www.sciencedirect.com/science/article/B6SYR-4W38RK8-2/2/0a8d826c71d21e995d826b7a15344cd2

Peter Bilak & Just van Rossum. 2003. LetterFrequencyMeter. Version: 0.9b2. Computer program. <Accessed: 30 December 2012>. http://www.type-applications.com/Letter_frequency_meter

Pike, Kenneth L., Thomas N. Headland & Marvin Harris (eds.). 1990. *Emics and Etics: The Insider/Outsider Debate* (Frontiers of Anthropology 7). SAGE Publications, Inc.

Pittman, Dean. 1948. *Practical Linguistics : A Textbook and Field Manual of Missionary Linguistics*. Cleveland, OH: Mid-Missions.

Porter, Jim. 2003. Why technology matters to writing: A cyberwriter's tale. *Computers and Composition* 20.4: 375-94. http://www.sciencedirect.com/science/article/pii/S8755461503000677

Poveda, David, Ana Cano & Manuel Palomares-Valera. 2005. Religious genres, entextualization and literacy in Gitano children. *Language in Society* 34.01: 87-115. http://dx.doi.org/10.1017/S0047404505050049

Powers, William K. 2009. Saving Lakota: Commentary on Language Revitalization. *American Indian Culture and Research Journal* 33.4: 139-49. <Accessed: 10. June 2014>. http://aisc.metapress.com/content/2832X77334756478

Priem, Jason, Dario Taraborelli, Paul Groth & Cameron Neylon. 2010. altmetrics: a manifesto. http://altmetrics.org/.http://altmetrics.org/manifesto

Priest, Lorna A. 2004. Transitioning a Vastly Multilingual Corporation to Unicode. Paper presented at 26th Internationalization and Unicode Conference, San Jose, CA, September 7-10, 2004. <Accessed: 09. May 2014>. http://scripts.sil.org/cms/scripts/render_download.php?format=file&media_id=IUC26_325&filename=IUC26_325.pdf

Provost of the University of Wisconsin. 2014. UW-Madison Department Planning Profiles. <Accessed: 11. May 2014>. https://apir.wisc.edu/deptplanningprof.htm

Pytlyk, Carolyn A. 2007. *Shared orthography: Do shared written symbols influence the perception of native-nonnative sound contrasts?* Masters of Arts dissertation, University of Victoria, Lin, Hua.

Pytlyk, Carolyn A. 2011. Shared Orthography: Do Shared Written Symbols Influence the Perception of L2 Sounds? *The Modern Language Journal* 95.4: 541-57. http://dx.doi.org/10.1111/j.1540-4781.2011.01244.x

Quakenbush, J. Stephen & Gary F. Simons. 2012. Looking at Austronesian Language Vitality through EGIDS and SUM. Paper presented at Paper presented at 12th International Confrence on Austronsian Linguistics (12-ICAL), Denpasar, Bali, (2-6 July 2012). <Accessed: 14. April 2014>. http://www-01.sil.org/~simonsg/presentation/12-ICAL.pdf

Redish, Laura & Orrin Lewis. 2009. Tano/Tewa Indian Language. <Accessed: 01. May 2014>. http://www.native-languages.org/tewa.htm

Rehman, B. & T. Qureshi. 2011. Urdu as interface design language - A novel approach. Paper presented at Electronics, Communications and Photonics Conference (SIECPC), 2011 Saudi International,24-26 April 2011.

Reiman, D. Will. 2010. Basic Oral Language Documentation. *Language Documentation & Conservation* 4:. 254-68. http://hdl.handle.net/10125/4479

Reina, Casiodoro de, Cipriano de Valera & Societies United Bible. 1995. Santiago. Santa Biblia : Reina-Valera 1995 (Series). Sociedades BÌblicas Unidas. <Accessed: 1 June 2012>. http://www.biblegateway.com/passage/?search=Santiago+1&version=RVR1995

Reyhner, Jon Allan & Louise Lockard (eds.). 2009. *Indigenous Language Revitalization: Encouragement, Guidance & Lessons Learned.* Flagstaff, AZ: Northern Arizona University.

Rice, Sally. 2011. Applied field linguistics: delivering linguistic training to speakers of endangered languages. *Language and Education* 25.4: 319-38. <Accessed: 2013/01/23>. http://dx.doi.org/10.1080/09500782.2011.577216

Rieger, Bryan & Stephanie Rieger. 2011. It's not the device people are after, it is all the things the device enables…. Paper presented at UX London 2011, Cumberland Hotel situated in Marble Arch, central London. <Accessed: 19. June 2014>. http://2011.uxlondon.com/programme/its-not-the-device/

Roberts, David. 2009. Visual Crowding and the tone orthography of African languages. *Written Language & Literacy* 12.1: 140-55. http://dx.doi.org/10.1075/wll.12.1.07rob

Roberts, David. 2011. A tone orthography typology. *Written Language & Literacy* 14.1: 82-108. <Accessed: 15 May 2012>. http://www.ingentaconnect.com/content/jbp/wll/2011/00000014/00000001/art00005

Romero, Nathan. n.d. Chochiti Keres: About Me and My Language: The politics of saving a vanishing language. <Accessed: 11. May 2014>. http://www.ling.hawaii.edu/ldtc/languages/keres/

Sackett, Kathleen D. & Linda A. Humnick. 2013. Evaluating Community-based Language Development Activities with the Sustainable Use Model: A Tsakhur Case Study. Paper presented at ICLDC'2013, Honolulu, Hawai'i, (1 March 2013). <Accessed: 15. April 2014>. http://hdl.handle.net/10125/26181

Sattar, Abdus, Al-Mukaddim Khan Pathan & Mohammad Ameer Ali. 2004. Development of an Optimal Bangla Keyboard Layout Based on Character and Fingering Frequency. *National Conference on Computer Processing of Bangla, 2004 Independent University, Bangladesh*, 38-46.

Schroeter, Ronald & Nick Thieberger. 2011. EOPAS, The Ethnoer Online Representation of Interlinear Text. In Nick Thieberger, Linda Barwick, Rosey Billington & Jill Vaughan (eds.), *Sustainable data from digital research: Humanities perspectives on digital scholarship*, 99-124. Melbourne: Custom Book Centre, University of Melbourne.

Sebba, Mark. 2007. *Spelling and society : the culture and politics of orthography around the world*. Cambridge, UK; New York: Cambridge University Press.

Sebba, Mark. 2009. Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1.1: 35-49. <Accessed: 2013/06/12>. http://dx.doi.org/10.1093/wsr/wsp002

Seddiqui, Hanif, Mohammad Mahadi Hassan, Sazzad Hossain & Nurul Islam. An optimal Bangla Keyboard Layout.

Seifart, Frank. 2006. Orthography development. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation* (Trends in Linguistics Studies and Monographs 178), 275-99. Berlin: Mouton de Gruyter.

Seltzer, Leslie J., Ashley R. Prososki, Toni E. Ziegler & Seth D. Pollak. 2012. Instant messages vs. speech: hormones and why we still need to hear each other. *Evolution and Human Behavior* 33.1: 42-5. <Accessed: 24. April 2014>. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3277914/

Sen, Lin, et al. 2005. A Tibetan input method based on syllable word for mobile phone. Paper presented at Embedded Software and Systems, 2005. Second International Conference on,16-18 Dec. 2005.

Shokri, Guiti, Carina Jahani & Hossein Barani. 2013. *When Tradition Meets Modernity: Five Life Stories from the Galesh Community in Ziarat, Golestan, Iran* (Studia Iranica Upsaliensia 21). Uppsala: Acta Universitatis Upsaliensis.

Shortis, Tim. 2007. Gr8 Txtpectations: The Creativity of Text Spelling. *EnglishDramMedia* 8. 21-6.

Simon, Ellen, Della Chambless & Ubiratã Kickhöfel Alves. 2010. Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences* 32.3: 380-94. http://www.sciencedirect.com/science/article/pii/S0388000109000412

Simons, Gary F. 2011. On defining language development. Paper presented at Poster presented at 2nd International Conference on Language Documentation and Conservation, University of Hawaii, 11–13 February 2011. <Accessed: 8 January 2013>. http://www.sil.org/%7Esimonsg/poster/ICLDC%202011%20poster.pdf

Société biblique française. 2000. Lettre de Jacques. *Bible Parole de Vie*. Villiers-le-Bel Société biblique française.

Stevenson, Nancy C. & Carol Just. 2012. In Early Education, Why Teach Handwriting Before Keyboarding? *Early Childhood Education Journal*. 1-8. http://dx.doi.org/10.1007/s10643-012-0565-2

Suárez, Jorge A. 1983. *La lengua tlapaneco de Malinaltepec* (Archivo de Lenguas Indígenas de México). México, D. F.: Universidad Nacional Autónoma de México.

Swadesh, Morris. 1971. *The origin and diversification of language*. Chicago: Aldine, Atherton.

Taylor, Jim. 2011. The Assessment of Research Quality in UK Universities: Peer Review or Metrics? *British Journal of Management* 22.2: 202-17. http://dx.doi.org/10.1111/j.1467-8551.2010.00722.x

Thieberger, Nicholas. 2009. Steps towards a grammar embedded in data. In Patricia Epps & Alexandre Arkhipov (eds.), *New challenges in typology: transcending the borders and refining the distinctions* (Trends in Linguistics. Studies and Monographs 217), 365-83. Berlin ; New York, NY: Mouton de Gruyter.

Thieberger, Nicholas & Simon Musgrave. 2007. Language documentation and ethical issues. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 4 (Language Documentation and Description), 26-37. London: SOAS.

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. http://septentrio.uit.no/index.php/nordlyd/article/view/2474/0

Trushkina, Julia. 2006. The North-West University Bible corpus: A multilingual parallel corpus for South African languages. *Language Matters* 37.2: 227-45. <Accessed: 2014/06/08>. http://dx.doi.org/10.1080/10228190608566262

Tyndale House Publishers. 2007. Book of James. From: *Holy Bible, New Living Translation*. Tyndale House Foundation. <Accessed: 1 June 2012>. http://www.biblegateway.com/passage/?search=James&version=NLT

UNESCO, Ad Hoc Expert Group on Endangered Languages. 2003. Language Vitality and Endangerment. Paper presented at International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Paris, France (10–12 March 2003). <Accessed: 7 January 2013>. http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf

Unknown. 2009a. JÚ[1] CHÚ[32] QUIOH[21] JESÚS TSÁ[2] LłŃ[3] CRISTO (El Nuevo Testamento en el chinanteco de Sochiapan). La Liga Bíblica. <Accessed: 12 June 2012>. http://www.scriptureearth.org/data/cso/PDF/00-WNTcso-web.pdf

Unknown. 2009b. Sí[2] Quioh[21] Santiago. JÚ[1] CHÚ[32] QUIOH[21] JESÚS TSÁ[2] LłŃ[3] CRISTO (El Nuevo Testamento en el chinanteco de Sochiapan) (Series), 525-33. La Liga Bíblica. <Accessed: 12 June 2012>. http://www.scriptureearth.org/data/cso/PDF/00-WNTcso-web.pdf

Unseth, Peter. 2011. The invention of Scripts in West Africa for Ethnic Revitization. In Joshua A. Fishman & Ofelia Garcìa (eds.), *Handbook of Language & Ethnic Identity*, vol. 2, 23-32. New York, NY [etc.]: Oxford University Press.

Vella, F., N. Vigouroux & P. Gorce. 2009. Fitts's law interpretation as ergonomic metrics for motor impaired users. *Computer Methods in Biomechanics and Biomedical Engineering* 12.sup1: 249-50. <Accessed: 2013/01/06>. http://dx.doi.org/10.1080/10255840903097871

Vendelin, Inga & Sharon Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116.7: 996-1007. http://www.sciencedirect.com/science/article/pii/S0024384105001038

Vincent, Audra. 2009. What are Native Tribes in the Northwest Doing to Revive Their Languages? *The McNair Scholars Journal of the University of Washington* 8.1: 347-61. <Accessed: 10. June 2014>. http://depts.washington.edu/uwmcnair/2009%20Journal.pdf#page=347

Vincent, Audra, Shannon Bischoff, Ivy Doak, Amy Fountain & John Ivens. 2013. The Coeur d'Alene Online Language Resource Center: A grassroots digital resource . Paper presented at 3rd International Conference on Language Documentation and Conservation (ICLDC), Honolulu, Hawai'i. <Accessed: 11. May 2013>. http://hdl.handle.net/10125/26156

Vitsœ. 2012. Dieter Rams: ten principles for good design. <Accessed: 27 June 2012>. http://www.vitsoe.com/en/gb/about/dieterrams/gooddesign

Walker, Andrew. 2012. What Is Boko Haram? In *Special Report 308*, pp. 1-16. Washington D.C.: United States Institute of Peace. June 2012. <Accessed: 17. June 2014>. http://www.usip.org/publications/what-boko-haram

Walker, Willard. 1984. The Design of Native Literacy Programs and How Literacy Came to the Cherokees. *Anthropological Linguistics* 26.2: 161-9. http://www.jstor.org/stable/30027501

Weber, Diana Dahlin, Diane Wroge & Joan Bomberger Yoder. 2007. Writers' Workshops: A Strategy for Developing Indigenous Writers. *Language Documentation & Conservation* 1.1: 77-93.

Webster, Anthony K. 2006. Keeping the Word: On Orality and Literacy (With a Sideways Glance at Navajo). *Oral Tradition* 21.2: 295-324. <Accessed: 07. May 2014>.

Wei-lan, Wang. 2007. Intelligent input software of Tibetan. *Computer Standards & Interfaces* 29.4: 462-6. http://www.sciencedirect.com/science/article/pii/S0920548906001279

White, John K. 1962. On the Revival of Printing in the Cherokee Language. *Current Anthropology* 3.5: 511-4. http://www.jstor.org/stable/2739696

Whitworth, Brian & Adnan Ahmad. 2013. *The social design of technical systems: building technologies for communities*. Aarhus, Denmark: Interaction Design Foundation.

Wied, Patrick. 2011. Heatmap Keyboard. Computer program. <Accessed: 31 January 2012>. http://www.patrick-wied.at/projects/heatmap-keyboard/

Woodbury, Anthony C. 2003. Defining Documentary Linguistics In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 1, 35-51. London: SOAS.

World Economic Froum. 2013. The Global Gender Gap Report. p. 397. Geneva, Switzerland: World Economic Forum. 24 October 2013. <Accessed: 16. June 2014>. http://www.weforum.org/reports/global-gender-gap-report-2013

Yabes, Leopoldo Y. 1977. History of Filipino as the common National Language. In Bonifacio P. Sibayan & Andrew B. Gonzalez (eds.), *LanguagePlanning and the Building of a National Language: Essays in Honor of Santiago A.Fonacier on his Ninety-second Birthday*, 341-55. Manila: Linguistic Society of the Philippines.

Yamada, Racquel-María. 2007. Collaborative Linguistic Fieldwork: Practical Application of the Empowerment Model. *Language Documentation & Conservation* 1.2: 257-82. <Accessed: 23 March 2014>. http://scholarspace.manoa.hawaii.edu/handle/10125/1717

Zamyatin, Konstantin. 2012. The Education Reform in Russia and its Impact on Teaching of the Minority Languages: An Effect of Nation-Building? *Journal on Ethnopolitics and Minority Issues in Europe* 11.1: 17-47. <Accessed: 12. June 2014>. http://www.ecmi.de/fileadmin/downloads/publications/JEMIE/2012/Zamyatin.pdf

Zaretsky, Elena, Jelena Kuvac Kraljevic, Cynthia Core & Mirjana Lencek. 2009. Literacy predictors and early reading and spelling skills as a factor of orthography: Cross-linguistic evidence. *Written Language & Literacy* 12.1: 52-81. http://dx.doi.org/10.1075/wll.12.1.03zar

Zhuromsky, Victor R. n. d. Иакова. *Russian New Testament by Victor R. Zhuromsky (RUVZ)*.

# Appendix I: Glossary of technical concepts and terms

Some technical terms and abbreviations used in this paper:

NRSI Terms

Terms as used by SIL's Non-Roaman Script Initiative (NRSI). NRSI maintains a glossary of technical terms at: http://scripts.sil.org/cms/scripts/page.php?cat_id=Glossary

**Glyph**: a shape that is the visual representation of a character. It is a graphic object stored within a font. Glyphs are objects that are recognizably related to particular characters and which are dependent on particular design (i.e. g, g and *g* are all distinct glyphs). Glyphs may or may not correspond to characters in a one-to-one manner. For example, a single character may correspond to multiple glyphs that have complementary distributions based upon context (e.g. final and non-final sigma in Greek), or several characters may correspond to a single glyph known as a ligature (e.g. conjuncts in Devanagari script). (For more information on glyphs and their relationship to characters, see ISO/IEC TR 15285.) http://scripts.sil.org/cms/scripts/page.php?cat_id=Glossary#writingsys

**Orthography**: a writing system for a given language.

**Writing system**: an implementation of one or more scripts to form a complete system for writing a particular language. http://scripts.sil.org/cms/scripts/page.php?cat_id=Glossary#writingsys

**Roman script**: the script based on the alphabet developed by the ancient Romans ("A B C D E F G ..."), and used by most of the languages of Europe, including English, French, German, Czech, Polish, Swedish, Estonian, etc. Also called Latin script. http://scripts.sil.org/cms/scripts/page.php?cat_id=Glossary#roman

**Grapheme**: anything that functions as a distinct unit within an orthography. A grapheme may be a single character, a multigraph, or a diacritic, but in all cases graphemes are defined in relation to the particular orthography. http://scripts.sil.org/cms/scripts/page.php?cat_id=Glossary#grapheme

**Orthographic Character**: a written symbol that is conventionally perceived as a distinct unit of writing in some writing system or orthography. http://scripts.sil.org/cms/scripts/page.php?cat_id=Glossary#orthochar

**Character**: In this paper unless otherwise specified, it is used synonymously with orthographic character.

**Composed Character**: see *complex character*.

**Composite Character**: also known as a pre-composed character. It is a single unicode point which represents a character which can be broken down into multiple other characters. ⟨ á ⟩ can be either a composed character consisting of both a the unicode points for ⟨ ´ ⟩ and ⟨ a⟩ or it can be a single character ⟨ á ⟩ and represented by a single unicode point.

**Complex Character**: is also referred to as a composed character; is a character which in order to be created must use more than one unicode code point. Sochiapam Chinantec's stressed barred i ⟨ɨ́ ⟩ is an example of this.

**Letter**: a non-technical term for a character in an orthography

**Base Character**: the main part of an orthographic character. the ⟨ a ⟩ in ⟨ á ⟩ would be the base character.

**Diacritic**: an accent or modifier to a base character

**Unicode Point (Unicode Value)**: is the unicode table designation for a glyph. It has the format of "U+ XXXX" where XXXX is some hexadecimal number.

**Keyboard Layout**: the abstract arrangement of characters and their relationship to the buttons of a physical or virtual keyboard.

**Keyboard (Physical):** part of the computer where text input is done.

**User Experience**: __()__

**Modifier Key**: http://en.wikipedia.org/wiki/Modifier_key

**Dead Key**: http://en.wikipedia.org/wiki/Dead_key Defined in the paper: this key is struck first and it alters the keyboard state so that the following key alters it output.

**Operator Key**: Defined in the paper: the base is struck first and then the modifier key

**Literacy as Social Practice:** __()__

**Writing as Social Practice**: __()__

**Literacy:** __()__

**Competencies:** __()__

**Socio-technical system (STS):** __()__

**Human-Computer Interation (HCI):** __()__

**Information Technology (IT):** __()__

**Technology:** __()__

**Text:** __()__

**Corpus**: __()__

**Macro Language**: In the ISO 639-3 code set there are two types of languages, single languages and macro-languages. There is alimited set of macro languages. An example is Arabic [ara]. Arabic is a macro language and can be refered used to refer to a variety of other languages which fall under the macro language like Algerian Arabic [arq] or Chadian Arabic [shu]. The macro language consept is a cary over from the ISO 639-2 standard.

**ISO 639-3**: An ISO standard employing three leter codes to identify language names. This is the current (2014) code practice for identifying languages in techincal (digital) contexts. For further discussion see: http://www-01.sil.org/iso639-3/scope.asp#M

**Design**:__()__

**User Experience Design**:__()__

**Txt:** Following (Shortis 2007) the term *Txt* is used to refer to the text used in SMS text messaging, instant messaging, internet chat, informal emails and social software

# Appendex II: Full text of analyzed texts

*English Full Text*

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

*Navajo Full Text*

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

*Me'phaa Full Text*

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

*Spanish Full Text*

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

*Chinantec Full Text*

How I got the text
When I got the text
Where I got the text
What I did to the text

How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Russian Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### German Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### French Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Italian Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Igbo Full Text

How I got the text
When I got the text

Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Mogolian Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Ezea Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Cishinini Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.

### Okphela Full Text

How I got the text
When I got the text
Where I got the text
What I did to the text
How it appears here
Full text.
Copyright holder as indicated:
Used by permission.