Keyboard layout as part of language documentation: the case of the Meꞌphaa and Chinantec keyboards

Hugh Paterson III
SIL International and University of North Dakota[1]
Hugh.Paterson@sil.org

Abstract

Successful writing systems today depend on electronic input methods which can be easily used for producing printed or electronic material. This paper explores keyboard design issues involved in designing two keyboards for two different established orthographies. Both orthographies are based on Latin scripts and cover a total of five minority languages in Mexico (four languages in the Meꞌphaa genus and Sochiapam Chinantec [cso]). The design issues consider:

1. *Technical differences encountered across major computer operating systems (OS X and Windows)*
2. *Computer culture issues like the keyboard layout of the national language*
3. *Key stroke frequency of language specific segments*
4. *Unicode/non-Unicode issues related to character composition*

Designing a Unicode keyboard for data input allowed native speakers of Meꞌphaa to have a greater involvement in the data collected by feeding the documentation team typed data and texts in addition to providing oral data. Early adaption of digital input methods may prove to better meet the needs of both language community and researcher. By giving the language community a keyboard for their orthography the minority language speakers were given the opportunity to enter into, and use, new technological domains with their language.

[1] Pre-Print Draft for comment. and presentation at CRASSH Conference Language Endangerment: Methodologies and New Challenges, July 6th 2012, Cambridge, UK. New technologies and Please direct all comments to Hugh (me) at Hugh@thejourneyler.org or hugh.paterson@sil.org

Hugh Paterson III. 2012. Keyboard layout as part of language documentation: the case of the Meꞌphaa and Chinantec keyboards. CRASSH Conference Language Endangerment: Methodologies and New Challenges, July 6th 2012, Cambridge, UK. Short URLs: http://hugh.thejourneyler.org/?p=1328. & http://bit.ly/KZ2yZM

Acknowledgments

I would like to thank Bob Hallissy from NRSI for help understanding Unicode and converting non-Unicode texts to Unicode; Mark Weathers for his extensive knowledge of the use of technology in writing Meꞌphaa; Steve Marlett for inviting me to work with him on documenting the Meꞌphaa genus; Wilfrido Flores for giving me the opportunity to serve him in the creation of a Chinintec keyboard; John Brownie, the creator of Ukelele a program which makes the designing of keyboards layouts on OS X a really easy thing to do; Kevin Cline for explaining to me how MSKLC works and for creating the MSKLC keyboards for Meꞌphaa; Tim Sissel of SIL Mexico Branch for is assistance with the history of keyboard design in SIL's language program involvement in Mexico; Kirby O'Brien for his involvement in creating some of the graphics used to demonstrate some of the ideas of this paper; Becky Paterson for proof reading and helping an oral processor like me reduce my thoughts to keystrokes. SIL International Americas Area for sponsoring my involvement in the Meꞌphaa genus Language Documentation project and funding this paper. I take full responsibility for all errors.

I. Introduction

The changing face of communication has seen some leaps and bounds in technology in the last few years. Today, more than ever before, interpersonal communication and the transmission of knowledge is bound to a digital medium of some kind. Living in a digital world requires that digital content have digital input. For textual input this generally means keyboards. For persons working to textually communicate in minority languages, the challenge comes when languages use Unicode characters in their orthographies, which are not part of the standard set of English characters (ASCII). The challenge is not encoding the characters in Unicode but rather accessing or inputing the characters in an efficient and ergonomic manner.

Several kinds of solutions have been created to address this challenge in minority language communication. These solutions can be generalized into two categories based on the tasks that the problem solver was trying to address. These tasks can be described as *open ended* and *defined* tasks.[2]

Defined tasks are those which have a specific goal or output in mind. Some examples of defined tasks in language documentation might be: the needs of language documenters to keyboard transcriptions of data, to build dictionaries and grammars, to write meta-data, or to build informed-consent forms in minority languages. It is important to note that keyboard layouts are often created as a by-product in pursuit of a defined task.

The other kind of task would be the open ended task where the goal is to affect social change; for the software product to spread virally along the lines of social networks. This is often the goal in language revitalization, multi-lingual education and community literacy efforts. In these kinds of projects, the goal from the outset of the project is to affect social change by encouraging the social use of a given language in current and new domains (or genres) of use. This open ended goal is by its nature concerned with language vitality.

When our concern is the continued use of a language and the impact that our work will have on a given language community, it is imperative that we consider bringing the use of that language into digital mediums. As the domain of digital-language-use becomes more important

---

[2] I am particularly looking at minority languages, not major languages. Major languages mostly have keyboards. These keyboards and keyboard layouts have histories based in the respective history of computing in those particular languages.

globally, the *User Experience*[3] around input methods for minority languages must improve. If text input solutions are not found for minority language writers, then the message to them becomes “you have a language which is impossible to type”. The impact will be for language shift to occur to a majority language which can function in the digital medium (Zheltov 2005 p. 10).[4]

Orthography development is an issue of the endangered language movement regardless of whether it is approached through applied linguistics or language documentation. At the 2011 annual meeting of the Linguistics Society of America there was a symposium on Orthography development. While the presenters acknowledged various non-linguistic factors in orthography development the impact of modern input methods in orthography development seems to be left unacknowledged[5]. Input methods, such as keyboards, are inextricably connected to orthographies. In Chuxnabán Mixe [pxm], computer keyboard layout has even affected decisions related to orthography development (Jany 2011). This is to say, input choices and what was visible on the keyboard affect orthography development. While orthographies have always had a relationship to the technologies used to implement them such as the availability of characters on the type-wheel of a typewriter, what is interesting about Jany and the Chuxnabán Mixe is that all the Unicode characters and fonts were available to them as mentioned by Cahill & Karan (2008: section 6). However, the basic problem was still one of *User Experience* (Jany 2010a: 5, Jany 2010b: 235-6) centering around input methods. Jany (2010b) suggests that orthographies should use characters which appear on standard keyboards:

> *A second important non-linguistic factor in the development of an orthography for an oral language is ease of use with computers and new media. With the world-wide web reaching even the remotest areas of the world and expanding in use, it becomes clear that a new orthography should be designed in a way so that its graphemes are readily available on standard keyboards. This will not only facilitate the language documentation process, it will also encourage its use with new media and possibly in new domains.*

I disagree. Orthography design should not be subject to mechanical imperialism. Technological development is fluid and responsive. Challenges like this should drive us to innovate. Today’s touchscreen technology will become more widely available. Touch screen virtual keyboards offer the opportunity to work around limitations imposed by physical

---

[3] User Experience is a term derived form the information technology and web-design professions. Perhaps the best definition of this term I have found is from Michael Commings (2010): *User Experience Design is the art and science of integrating all the various elements that comprise an interactive system so that:*

1. *The user's needs, limitations, goals, desires, expectations, are served.*
2. *The publishing organization's objectives are served as a result of serving the user's (#1).*
3. *The whole is greater than the sum of the parts.*

[4] Minimally, this shift would occur in the use of the language as it is connected with digital technology. However, it can be imagined that the impact may be part of pressures having a broader impact on language shift.

[5] Mike Cahill does mention keyboards in his proposal (Cahill 2011), but only mentions typewriters in currently available documents from the symposium. A formal publication is expected (Cahill 2012, pc.). It might address some of these issues.

keyboards. Limiting an orthography to the characters which are easily accessed on a keyboard can still be problematic for several reasons:

1. Keyboards vary around the world
2. Some languages have more phonological distinctions than standard keyboards offer and benefit from having “non-standard” characters.
3. This approach to orthography development forgets the purpose behind invention of the keyboard. - It was created with certain character-key combinations to serve specific languages and their orthographies, not to serve as a repository of characters from which orthography designers can choose.
4. Leaders in the field of orthography development recommend that linguistic symbols be used when possible. (Cahill & Karan 2008)
5. There is a cognitive-psychological relationship between the sounds we hear and symbols we use. Orthographies and literacy do have an effect on the way we recall and relate to oral language (Perre, Pattamadilok, Montant & Ziegler 2009, Ziegler, Ferrand & Montant 2004, Ziegler & Muneaux 2007).

Good design would be to innovate new technology to the point that the technology is invisible and the needs of the community are served. In all probability, when communities like Chuxnabán Mixe are served with well designed solutions, the solutions are likely to also benefit other language communities.

One inspirational and innovate approach is the story of the South African Keyboard (Bailey 2007), created by `translate.org.za`. The designers chose to not be subjected to the confines of technology and used the opportunity to create a keyboard layout which addressed the specific needs of one language, Venda [ven] and still served the needs of multiple languages in South Africa.

II. The two keyboards in focus

This paper explores the design process for two keyboard layouts which serve a total of five languages in México. The first keyboard layout serves the four Meꞌphaa languages ([tcf],[tpx], [tpl],[tpc]). The second keyboard layout serves Sochiapam Chinantec [cso][6]. Both language situations use Latin scripts. In each section below we will look at the physical devices, software, orthographies and the language development environments to help define the complexities of the tasks these keyboards need to accomplish. As we will see, the extensive use of diacritics and punctuation marks create their share of challenges.

Meꞌphaa

The original purpose of my involvement in the keyboard design for Meꞌphaa was to facilitate text creation and the typing of texts as part of the NEH funded project *Documenting the Meꞌphaa 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 was already in use by several Meꞌphaa

[6] Three letter codes are the current (as of 2012) ISO 639-3 codes for these languages. (ISO 639-3:2007)

writers, including some bilingual teachers in the Meꞌphaa 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 Meꞌphaa 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)

[7]

- Access to characters from the Meꞌphaa orthography
- The design standard (ISO v.s ANSI) of the physical keyboards[8]

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

[8] Keyboards sold in the U.S.A. are more often ANSI, whereas keyboards sold in Mexico are more often ISO.

[9]

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 font[10]. 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 correctly in 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 in the document. However, using the hacked font created documents which were composed in Unicode but where the characters would not display or print as the intended Unicode code points. 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 X. To create these keyboards we used Ukelele

---

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

[10] 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.

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[11] were used this also affected the behavior of the keyboard and the typing experience. This is discussed more fully in sections three and four.

[12]

[11] 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.

[12] To be converted to caption: Meꞌphaa keyboard with no shift state.

[13]

Sochiapam Chinantec

The second keyboard layout discussed in this paper was designed for Sochiapam Chinantec. A Keyman (Durdin & Durdin 2011) keyboard already existed and was in use by several people involved in a community literacy program. Wilfrido Flores, a native Chinantec speaker and writer, and one of the program facilitators asked me to make it possible to type Chinantec on his MacBook Pro. What resulted was a copy of the Keyman keyboard[14], working on OS X and an ANSI keybord

---

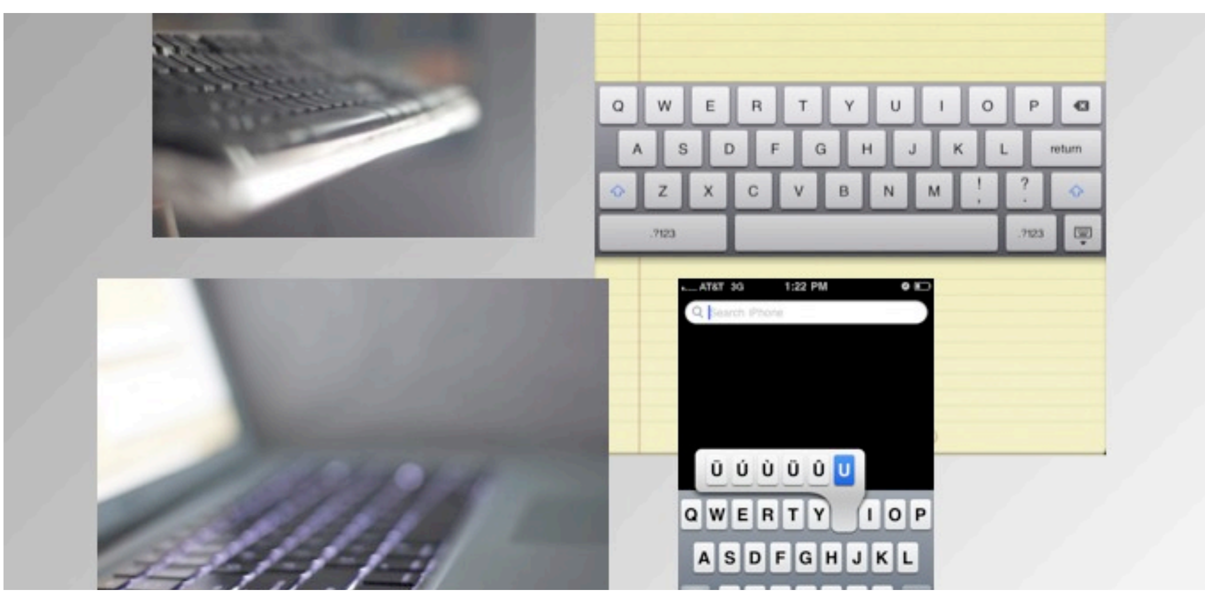
[13] To be converted to caption: Meꞌphaa keyboard with shift depressed.

[14] Keyman only works on Windows operating systems and is pay-ware (fee based license).

In section four I will discuss how the orthographies of these languages relate to the keyboard layouts and various challenges users will encounter as we consider keystroke frequency and diacritic accessibility. Before diving into the factors related to orthographies, we will review some of the goals, options and challenges in dealing with the physical objects we call keyboards and their virtual representations, keyboard layouts.

III. Keyboards, keyboard layouts and design

It is important here to distinguish between the terms *keyboard* and *keyboard layout*. This will help guide this discussion as we define reasonable solutions for electronic input in minority languages.



Keyboards

Conceptually, a keyboard could be a separate device from the computer, however in some cases, such as laptop keyboards, the keyboard is an inseparable part of the computing experience. In both cases a keyboard is a physical object. However, the difference between separable and inseparable leads to an important design distinction. With separate, detachable keyboards many more design options are available, including: designing, building and marketing a completely new device. This is in fact what Lancor Technologies of Nigeria did when it designed the Konyin keyboard to facilitate typing in a variety of Nigerian languages (Lancor Technologies 2012, MarketWire 2007, Nzeshi 2006). However, a solution like the Konyin keyboard does not solve portability goals and compactness which many computer users have.[15] Laptops, netbooks and tablets are devices which fill the portability needs of users. While there are many laptops in the world due to the nature of the computer industry in the last decade, the second quarter of 2011 saw for the first time more tablets ship from manufacturers than netbooks (ABI Research 2011). At the same time there has been a decrease in the sales of new laptops (Pinola 2011). This is important in the world of keyboard design for two reasons:

1. Tablets generally have virtualized, touch screen keyboards (keyboard layouts) rather than physical keyboards.
2. Tablet OSes like iOS and Android can offer context sensitive keyboard choices to users. This adds more thought and design questions to the equation.[16]

Another way that a keyboard can be conceptualized is as a virtual keyboard, where each button on the physical keyboard only has a relative meaning. Virtual keyboards may or may not be bound to a physical keyboard. i.e. tablets v.s laptops. All three kinds of keyboards - attached, detached, and virtual - have a keyboard layout.

Keyboard Layouts

Keyboard layouts are definitions of which symbols and characters are found in which order and how they are mapped to a given key on the keyboard. This includes any dead keys and modifier keys used to reach characters. There are at least three important aspects to keyboard layouts which deserve some mention:

1. Keyboard layout as it affects the build of the physical keyboard
2. Keyboard layout as it affects the virtual keyboard
3. Keyboard layout as it affects virtual keyboard on tablets

The layout of keys on a physical keyboard is often governed by one of three industrial standards: ISO (ISO/IEC 9995-2:2009), ANSI (ISO/IEC 9995-3:2010), and JIS (JIS X 6002:1980). These standards not only define how many keys are on a physical keyboard but also where on the keyboard, and the general shape in which they occur and also which characters

---

[15] A physical keyboard device also does not solve the viral social distribution aims of language documenters, i.e., a physical keyboard can not be passed across an online social network from one user to another.

[16] An example of a context sensitive input is when the user is being asked to input a telephone number that they are presented with a keyboard with numeric digits.

appear in which order and which are printed on the surface of the keys.[17] Manufacturers use these standards and consider them as they produce keyboarding experiences for their products. The ISO keyboard is mainly produced on computers being marketed around the world including markets in Europe, Central and South America. The ANSI keyboard is most popular in the United States of America; the JIS keyboards are almost exclusive to Japan. However, not all vendors produce items which follow an international or national standard for keyboard layout or for physical keyboard production.[18]

The differences between these industrial standards bring out two major differences which affect users of keyboards:

1. The quantity of keys on the keyboard
2. The position of keys (the keyboard layout)

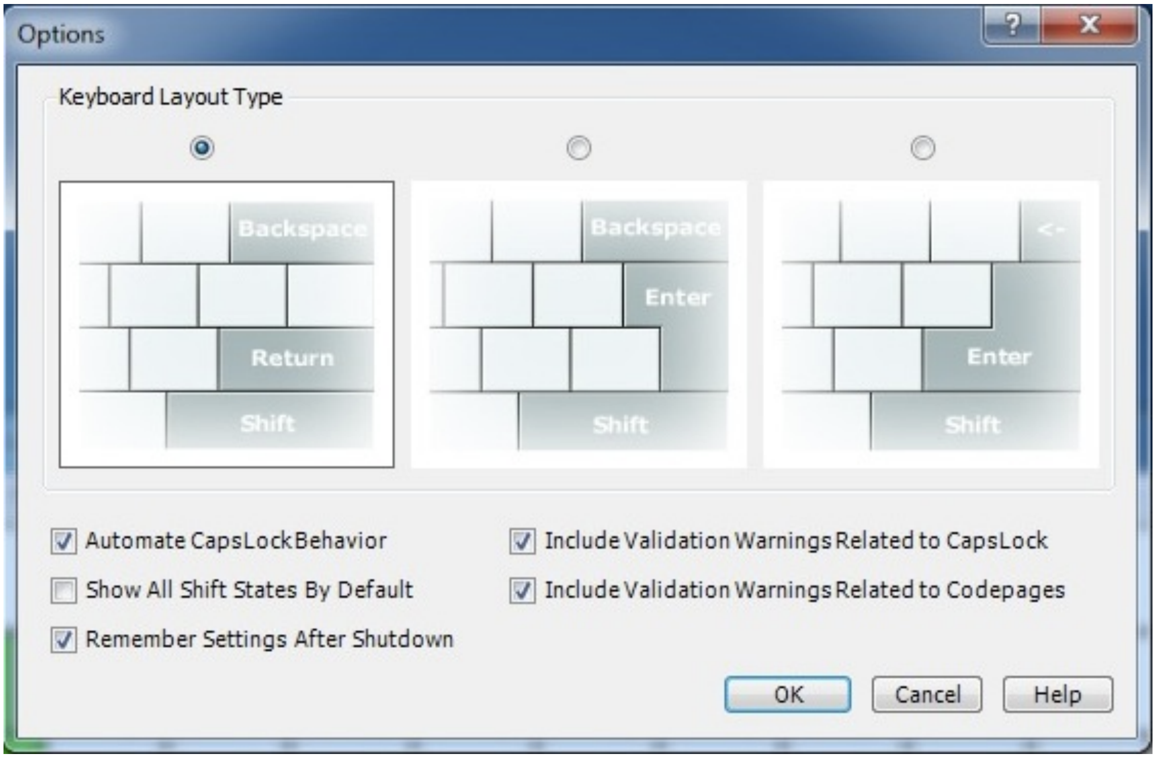
With respect to the quantity of keys ISO keyboards have one more key than ANSI keyboards while JIS keyboards have two more than ANSI in their compact forms. Full size keyboards have more keys; on the order of 110 (ISO), 109 (ANSI), and 112 (JIS)[19].

With respect to the position of the keys and the keyboard layout, there are several noticeable differences between the ISO and the ANSI keyboards.

[17] This distinction in manufacturing is also the source defining keyboards by how many keys they possess. Keyboards produced by Apple, Inc. in the ISO, ANSI and JIS layouts may be labeled as having 110, 109, and 112 keys or in their compact forms 79, 78, and 80 keys respectively. I have also seen key full sized keyboards which range from 101-109 keys as they are built to different design standards and for use in different markets.

[18] An example of this is Apple laptops which are marketed in Britain use an ANSI keyboard but have a hybrid layout which is not ANSI, ISO or BSI (BS 4822:1994). There appears to be a shift in BSI to adopt the ISO/IEC 9995-2:2009 standard (BSI 2012).

[19] These numbers represent what are on Apple produced keyboards. They do not represent what the various industrial standards specify.



These differences include the shape of the enter/return key, the presence of an additional key between the left shift key and the "Z" key on a QWERTY layout (10 vs.11 keys between the shift keys). The shape of the enter/return key on an ANSI keyboard is rectangle whereas the shape of the enter/return key on an ISO keyboard is "L" shaped. Three of the other minor differences are:

1. The placement of the vendor key
2. The presence or absence of auxiliary keys like function-keys or number pad keys or brightness control keys/buttons.
3. The presence of additional modifier keys (more common in JIS keyboards)

20

[21]Practically speaking this means that different computers and/or keyboards have different quantities of keys which have different physical arrangements, producing different characters when struck by the user. This becomes relevant when we design *virtual keyboards* or *keyboard layouts*.

---

[20] To become caption: ISO keyboard made by Apple, Inc. with a German QWERTZ keyboard layout.

[21] To become caption: ANSI keyboard made by Apple with a U.S.A. QWERTY keyboard layout.

Virtual keyboard layouts are software which provide users with an alternative keyboard layout to what is printed on the surface of the keys. All keyboards require software which interprets keystrokes and outputs characters[22]. Major operating systems such as Windows, Ubuntu, OS X, Android and iOS contain built in methods of having several virtual keyboard layouts installed on a single device for use at different times. This means they have built in functionality for facilitating the switching of keyboard layouts without switching the physical keys (if a particular device even has physical keys).

Not having physical keyboards is the is one of the hallmarks of tablets. Touchscreen technology has elevated the importance of keyboard layouts. By making the keyboard layout the continuity to the user's other typing experiences. By not having physical keys this also liberates the keyboard layout from certain constraints like the physical size, quantity, or arrangement of

keys.[23] Also by the nature of using on screen buttons rather than physical keys, touch screen keyboards are able to do things which were not possible with standard keyboards. One example of this is the "hold" gesture. On iOS devices if the key is held then selected character's related to that character become visible for selection. This behavior was also recently introduced in OS X 10.7 "Lion".

[22] For a more technical discussion on these processes consult Hosken (2001a, b).

[23] To be made a Caption: iPad keyboards, while appearing to be QWERTY are neither ANSI nor ISO. They use context to provide the user the most relevant characters.

[24]This behavior, where the duration of a keystroke has an impact on tasks in applications, has also been replicated in applications like Google Chrome.[25] This is to say that keyboard gestures can serve multiple functions and can be accessed by both application developers and core OS functions. Keyboard designers for minority languages should add this option to their tool belt when designing a keyboarding experience. Apple uses this approach for when keyboarders need to access infrequently accessed characters.

So, keyboard layouts are often affected by industrial standards, affect physical keyboard production, can be implemented by the OS regardless of what is printed on the keys of physical keyboards, serve to bring an element of continuity to computing tasks for tablet users, and are facing a new level of change as they are virtualized and contextualized on tablets. However, as important as it is to know all this, this knowledge does not directly help us answer the following questions:

- *How does the person tasked with designing a new keyboard layout design a good layout?*
- *How do they build it?*
- *By what criteria should they evaluate their own work?*

Design

The words *Usability* and *Design* each suffer from a very unfortunate ambiguity. Usability in a very raw sense means is a tool usable. Just because every tool can be used as a hammer, does not mean that every tool should be shaped like a hammer. Nor does it mean that every tool should be

[24] To be made a Caption: The "hold" gesture on an iOS device showing how an accented character can be selected.

[25] This can be seen on Google Chrome Browser version 20.0.1132.47 running on OS X 10.6.8 in the "Hold to quit" dialogue.

used as a hammer. Just because a keyboard layout can be used does not mean that it is a good layout. The term design in computing also suffers a similar fate to the term usability. If some computer tool does something, it does so because it was designed to do so. The software was not generated by accident. Mere existence does not mean that the computer tool is esthetically pleasing or that it creates a sexy or desirable impression on its user (Anderson 2006, 2009, 2011). An impression of such a nature that the user wants to come back to the software and use that software again, while that software actually fits the functional needs of its users .

As language documenters and linguists, when we build digital solutions, like keyboard layouts, we need to consider the lasting effect on the communities to whom we are providing these products. It is our moral and professional obligation not just to seek out solutions but to seek out great solutions. In the manufacturing industry, manufactures are often held accountable for the effects of their products on the users of their products. When we offer our linguistic and technical expertise to communities of minority language speakers and writers, we need to not just design solutions; we need to offer well designed solutions. Just because we create something which is usable and useful does not mean we have created something desirable. And when the community does not want to use that created input method, our answer should not to simply say: “well they do not have enough desire”. The interesting thing about keyboard layouts is that they are not just products, they are also experiences as well. Each keystroke in its place is a pattern created in an attempt to implement the orthography. It creates an experience that writers’ fingers will potentially encounter multiple times a day. This physical interaction is only part of the User Experience, and should not be overlooked in the design process. Other parts of the User Experience deal with the keyboard layout as software. So the keyboard layout should be considered and designed as software as well.

There are several marks of good design. Dieter Rams has put forward ten principals of good design (Vitsœ 2012). Here I will relate to four of them.

Good design:

- Makes a product useful
- Makes a product understandable
- Is unobtrusive
- Is thorough down to the last detail

Rams suggests that a product is not useful if it does not also meet certain criteria: aesthetically, functionally, and psychologically. Because we are talking about keyboard layouts, aesthetics are physically dictated by the physical keyboard, or the manufacturer of the OS on tablets. However, criteria for function and the psychological relationship with the keyboard layout both are very available to keyboard layout designers (linguists and language documenters).

Some of the functional criteria are obvious, the keyboard layout must be able to implement the orthography of the target language. Another very useful thing for the keyboard layout to accomplish is to type in the majority language’s orthography. In the cases of Meˈphaa and Sochiapam Chinantec this would mean being able to type Spanish as well. It is important to notice the directionality of composition. Typing a document in Spanish and adding a few words or sentences in Meˈphaa is drastically different from typing a Meˈphaa document and adding a few words or sentences in Spanish. Even if the orthographies are “similar” in that they both use Latin scripts and try and show social affinity by “looking similar” there is still a difference in terms of the User Experience when composing the document. For instance in the texts analyzed in the next section, the Meꞌphaa glyph < á > is used 880 times whereas in the same content written in Spanish the same glyph is only used 59 times. A keyboard which takes into account the complexity of inputing a complex glyph should also take into account the frequency that that glyph is accessed. To input 59 < á > glyphs in Spanish on the Spanish ISO keyboard requires one to make 118 keystrokes. Alternatively, on a standard OS X ANSI U.S. QWERTY keyboard, one would need to make 177 keystrokes to form the same 59 < á > glyphs. Using the Meꞌphaa keyboard layout we created it still only takes 118 keystrokes to produce the 59 < á > glyph in Spanish. However, if one were writing the same content in the Meꞌphaa language one would need to use 880 < á > glyphs, this would require 1,760 keystrokes. At this point the keyboard layout designer needs to ask: *Is the Spanish keyboard layout an efficient option for typing Meꞌphaa?* Concerning psychological factors, we need to consider how much work it is to produce each character and how that impacts a person’s desire to type in a given language. Psychological factors also include User Experience and the psychological composition of characters represented by complex characters like accented characters representing tone and stress. In terms of User Experience, we must also consider placement of a frequently typed character on the keyboard layout. In the Meꞌphaa text analyzed the LATIN SMALL LETTER SALTILLO U+A78C is used 1,189 times. This accounts for 7% of all characters used in the text and is the second most common non-complex character. This character on the Meꞌphaa keyboard layout is placed at one of the farthest places on the keyboard for the little finger to reach[26]. This distance can have an effect on a typists speed. The Saltillo is also a character which is not in the Mexican-Spanish orthography[27]. The combination of these two factors makes it more compelling to type in Spanish rather than in Meꞌphaa. The ordering of input on the diacritic marks is another User Experience factor pointed out by Hosken (Hosken 2001 section 5.2). It turns out that this is significant for both the Meꞌphaa and and Sochiapam Chinantec keyboard layouts. Both keyboards use dead keys to assign diacritic marks to base characters. What is important is the ordering of striking the diacritic. Should the dead key (or what might be socially viewed by users as a “tone mark key”) be struck first and then the base (known as the dead key method) or should the base be struck first and then the modifier key (known as the operator key method)? While both may be valid ways to consider input there are several issues related to User Experience which need to be considered.

This is an example of an accent ˊ

---

[26] On a physical ANSI keyboard it is a little bit further than on a physical ISO keyboard.

[27] The Saltillo also does not appear graphically on the physical keyboard. This is not a major challenge, but means that the user will have to learn to strike a key which does not return an input which matches the key top.

One of the concerns that Hosken (2001) brings up is that using the dead key method does not provide the user with any visual feedback, whereas there is always a visual change for every keystroke with the operator key method. It would appear that this has be addressed in OS X by rendering the diacritic without a base prior to the base being struck. Additionally, if the diacritic is one which can become part of a composed character, it is backgrounded in yellow. However, as of Windows 7, the behavior of Windows is still to not show the diacritic before the key for the base is struck. Therefore on that OS, Hosken's concern obtains.

In his design principles, Rams says that "Good design makes a product understandable" (Vitsœ 2012). He refines this by saying that the product needs to explain itself.[28] For keyboards this can apply in several ways.

It used to be the norm that most minority language writers are also new computer users (or had no exposure at all to computers). With globalization and digitization of communication, there are many more endangered language speakers who use or have access to computers. One challenge of using a virtual keyboard, and even more so for those who are new to typing, is the miss-match between what is printed on the keys of the keyboard and what the keyboard layout returns as output.

One evidence of this confusion was found in the Meꞌphaa text analyzed. Spanish uses a character called Guillemets < « » > to serve as quote marks. These are known in Unicode as < « > LEFT-POINTING DOUBLE ANGLE QUOTATION MARK U+00AB and < » > RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK U+00BB. The Meꞌphaa orthography also calls for the same style quote marks. When I was analyzing the Meꞌphaa text it was discovered that instead of using Guillemets the typist used a double set angle brackets. The best explanations for this kind of mistake is that the typist saw what was printed on the key and make logical mistake that a set of two angle brackets must be the same as

[28] It clarifies the product's structure. Better still, it can make the product talk. At best, it is self-explanatory.

a Guillemet.[29] One way to avoid this kind of mistake is to make the angle bracket key when hit twice in a row output a Guillemet. However, this method is more divergent from the Spanish ISO keyboard, and begs the question: *What will a Meꞌphaa writer do when they are writing a Meꞌphaa text on a Spanish keyboard?*

Another way to make keyboard layouts are more understandable is to design them to behave the same way across multiple operating systems. This provides continuity to users when they switch computers or operating systems and maximizes opportunities for social, peer based learning. Both continuity and learnability are important factors which affect the adoption of software. When language use in the digital medium is in question, adoption of software is essential to the success of language revitalization.

In implementing the Meꞌphaa keyboard layout one of the questions which was asked was: *Could use the vendor key as a dead key?* The result was “no, we could not use it”. It is typical for applications running on OS X to use the vendor key, also known as the command (⌘) key, as an application level shortcut key. This behavior is also not uncommon on Windows and Linux machines where the control key is used in leu of the vendor key. On Windows based machines, the vendor key is used to bring up the Windows menu. Therefore using the vendor key as a modifier key can be somewhat problematic and is not a optimal design choice.

A third way that a keyboard needs to explain itself is through the cognitive associations it invokes on its users. Should all characters be accessed the same way?[30] Not all orthographies use the same characters to represent the same sounds or ideas (phonetic or phonological representation as understood by the orthography users). Kutsch Lojenga (2011) offers an example from Yaka [axk] and Sango [sag] in the Central African Republic:

> *Occasionally, different accents are used, e.g. when the circumflex is used for H tone, as is done in YAKA (Bantu C.10, spoken in C.A.R.), where the choice of tone marks had to conform to the system used in the widely-known lingua franca Sango, by using a circumflex for H tone. It may not be elegant for a linguist, but it works.*

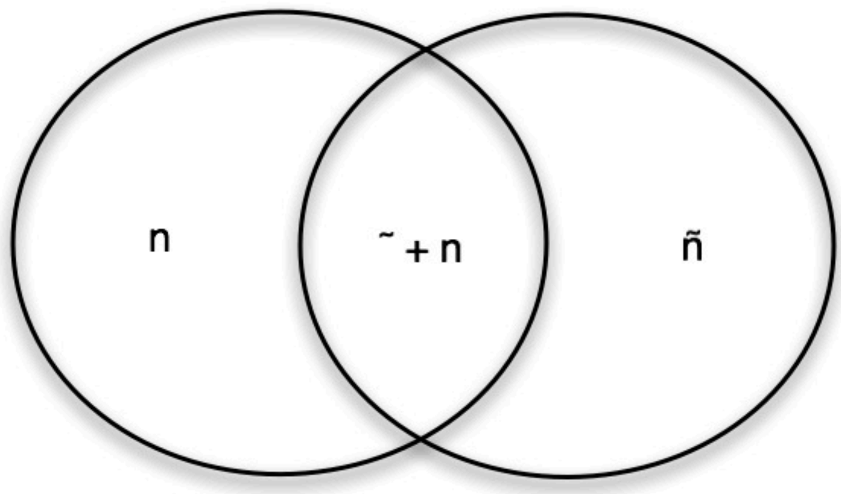
In Sochiapam Chinantec and Spanish the use of < ñ > is in a similar relationship. In Sochiapam Chinantec < ñ > represents a velar nasal whereas in Spanish it represents a palatal nasal.[31] However, in terms of the character composition, and tactile input of that character the question becomes *How is the < ˜ >diacritic related to the base < n > and does that relationship*

[29] Since the creation of the analyzed text the project has switched to using English style quotes and use ANSI+0147 & ANSI+0148 < “ ” > rather than the respective equivalent Unicode points U+201C & U+201D. (Mark Weathers, pc)

[30] So should all characters be accessed the same way in keyboard layouts? More importantly if they are not accessed in the same ways then how should they be?

[31] A similar relationship exists in Meꞌphaa with the acute accent. In Meꞌphaa the acute accent represents tone, where as in Spanish it represents stress.

parallel the semantically salient ideas about the phonemes these glyphs represent?



How do speakers conceptualize the graphical element
and relate it back to the sound it represents?

One option in designing the Meꞌphaa and Sochiapam Chinantec keyboard layouts was to remove the < ñ > from having its own dedicated key and make the < ˜ > a diacritic which was then accessed through a dead key or even the same dead key other diacritics in the language were accessed through.

Unicode composite and base characters with consonants	
ñ	n + ˜
LATIN SMALL LETTER N WITH TILDE U+00F1	LATIN SMALL LETTER N U+006E + COMBINING TILDE U+0303

This would serve to free up a key in the layout for a more common character, and also serve to bring consistency to the input of characters with diacritics. In both cases it was decided to leave the < ñ > key as it appears on the Spanish ISO keyboard layout. But this example serves to point out that such considerations should be made.

If a keyboard layout is to be intuitive to its users then there should be a parallel between the graphical representation of sounds and the way the glyphs are generated through the fingers. There should also be some internal cohesion regarding how composite characters are created by a given keyboard layout. This too is part of designing the tactile experience aspect of the keyboard layout.

An example of internal cohesion can be seen in the Meꞌphaa keyboard layout in the way that tone is marked. Meꞌphaa has three levels of tone which are indicated in the orthography. High tones are marked with an acute accent above the vowel < ´ >, mid tones are unmarked, and low tones are marked with a *COMBINING MACRON BELOW U+0331* < _ >. The use of the macon below gives the visual effect of an underline below the vowel. The Meꞌphaa keyboard layout dedicates one dead key for high tone and another dead key for low tone. By giving each tone mark its own dead key the keyboard layout creates a symmetry in user experience for how a tone can be marked on each vowel.

The symmetry in the Sochiapam Chinantec keyboard was not as simple to implement. This is due to some limitations in one of the operating systems the keyboard was being implemented on and the way that characters are coded in Unicode. Understanding the how Unicode allows for the target characters to be created will help us see where there is symmetry and where there is not.

In Meꞌphaa the letter < a > can be used by itself, with a low tone mark or with a high tone mark. In every case that < a > is combined with a low town mark two Unicode characters are needed: the base character < a > and the combining macron below diacritic < _ >. However, when a high tone is used there are several ways these could be encoded: as < a > plus < ´ > or as a single character < á >.

Unicode composite and base characters with vowels		
a	a + _	a + ́ or á
LATIN SMALL LETTER A U+006	LATIN SMALL LETTER A U+0061 + COMBINING MACRON BELOW U+0331	LATIN SMALL LETTER A U+0061 + COMBINING ACUTE ACCENT U+0301
		LATIN SMALL LETTER A WITH ACUTE U+00E1

In the Meꞌphaa case the available options in Unicode do not make a difference for the implementation of a symmetrical input method. However, in the Sochiapam Chinantec keyboard layout there is an important difference. Sochiapam Chinantec, like Meꞌphaa is a tonal language. However, the orthography does not mark the tone on the vowel, but rather with numbers at the end of the syllable (Fortis 2000, Unknown 2009). It does mark a type of stress on the vowels with an acute accent (Mugele 1982). Symmetry does not become a problem until we try and implement a stressed barred ɨ <ɨ́ >. Unicode does not have a composite character for LATIN SMALL LETTER I WITH STROKE AND ACUTE. This means that the character needs to be a series of at least two Unicode code points and it could potentially be coded as three code points.

The conceptual construction of a Character		
ɨ	í + ̵	ı + ̵ + ́
LATIN SMALL LETTER I WITH STROKE U+0268	LATIN SMALL LETTER I WITH ACUTE U +00ED	LATIN SMALL LETTER DOTLESS I U+0131
	COMBINING SHORT STROKE OVERLAY U+0335	COMBINING SHORT STROKE OVERLAY U+0335

The conceptual construction of a Character		
ɨ	í +̵	ı +̵ + ́
		COMBINING ACUTE ACCENT U+0301

The challenge comes because the keyboard layout editor from Microsoft for Windows, MSKLC, will not allow the building of keyboard layouts which provide the stringing of input with one keystroke. That is, if I want to input three or more characters with one keystroke, it can not be done.[32] This behavior is desirable for using decomposed characters. i.e. if I have a base and two diacritics then I would have three unicode characters. With OS X, a dead key can be used to enter another state of the keyboard, in which when the correct key is struck the desired series of Unicode characters is input. However, with MSKLC it is impossible to replicate this behavior. Rather, the dead key must be used to insert the combining diacritic and then the next key is used to insert the base. In this manner all of the necessary diacritics for Meˈphaa were addressed.

However, for the Sochiapam Chinantec no solution was found for accented barred <ɨ́ >. This is because by nature this is a composed character. If we were to represent symmetry with the other characters in the orthography, we would have to move from barred i < ɨ >, LATIN SMALL LETTER I WITH STROKE U+0268 to barred i with acute. This is not possible with Unicode because the barred i would need to be dotless, rather than combining above the dot. Alternatively, one could add the diacritic COMBINING SHORT STROKE OVERLAY U+0335 to the base character < í > LATIN SMALL LETTER I WITH ACUTE U +00ED but this method on Windows would require a fourth dead key for the stroke overlay (a dead already exits for acute, dieresis, and tone) and this dead key would not match the behavior of the keyboard layout for adding a stress mark to the other vowels. If implemented as a stressed < í > plus a stroke overlay. This would also not fit the way that indigenous writers thing about the vowel; as being barred i < ɨ > plus stress.

Rams' third point about design which is applicable to keyboard layouts is *Good design is unobtrusive.* The limitation of MSKLC to be able to produce keyboards which are flexible and meet the needs of writers has opened the door of opportunity for third party application developers to come up with creative solutions. These solutions add complexity to the User Experience and also add complexity to the deployment of keyboard layout files. Up to this point the discussion has focused on using tools which create files which work and are installed within the framework of the OS not needing third party software. Some of these third party solutions include:

- Keyman (Durdin & Durdin 2011)
- Inkey (InKey 2012)

Both of these solutions allow for the editing of custom keyboard layout files but require their software to be active and running on the computer to use the keyboard layouts.

[32] This truly seems counterintuitive. If a reader has a comment on accomplishing this within MSKLC please do share. But to the best of my ability I have not found a way to accomplish this behavior.

Number of total Low tones *(use of Combining Macron Below U+0331 plus base)*	875	16
Number of total High tones *(composite characters using base)*	880	9
Number of times the base glyph is used without modification	1195	93
Number of total base characters	2950	118

What do we need to consider if we are going to design keyboards for minority languages?

Do we want to use of products created through our efforts to spread? What is the impact we as catalysts in language vitality situations want to see? What is the level of complexity we want to bring to writing in another language?

FOSS principals
Distribution mechanism (Raymond 2012)

National language keyboard layout (is there one? what is "official" what is "common" can it be better? What is better?)

How the character is composed in the mind of the writer (is it a toneme or is it a vowel with a tone? or is it a high tone vowel?)
The Actual Composition of characters

Unicode/non-Unicode issues as they relate to character composition.

Design

From a *User Experience* perspective, there are four "levels" to design in keyboards. These include:

1. What the physical keyboards produce when the keys are pressed.
2. The physical layout of button parts of the keyboard.
3. What is printed on the face of the keys.
4. How the character is accessed.

There are several layers to a keyboard

Generally speaking, there are three arrangements of physical keys on keyboards.

This keyboard, with respect to its physical keys, inter-changeable, with the following exception.

Good design would not be designing a usable keyboard (physical). Good design would be creating a keyboard layout which ergonomically and psychologically meets the needs of users.

IV. The Experiment

So in looking at *User Experience* one of the questions is: *How accessible to typists (users) are the most common characters they will be trying to access?* To answer this question I took four texts[33], one in each English (NLT 2007), Spanish (RVR 1995), Sochiapam Chinantec (Unknown 2009) , and Me'phaa (Carrasco Zúñiga et al. 2008). I counted the characters and then mapped the characters to the keystrokes required to produce those texts with the keyboards mentioned in this paper.

[33] The book of James was used as it was readily available across the set of target languages and by using a translated text it gave a sample which maintains a semantic continuity.

About the texts

Text 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	20416	4506	9611	2165	10432	2575
Keystrokes	20588	Keystrokes	27515	Keystrokes	12715	Keystrokes	13575
Keystroke to character Efficiency	80.72%	Keystroke to character Efficiency	74.20%	Keystroke to character Efficiency	75.59%	Keystroke to character Efficiency	76.85%

To analyze the texts the characters were divided into three groupings: punctuation, characters without diacritics, and characters with diacritics. Though for *User Experience* analysis the characters could have 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 orthographies

(Fortis 2000, Unknown 2009b)

Explaining the differences in orthographies.

Citations for Me'phaa orthography (Asociación para la Promoción de Lecto-Escritura Tlapaneca 1988, Marlett enero 2012a)

Citation for orthographical Tone typology (Roberts 2011)

Punctuation

Punctuation Usage									
Characters		**Me'phaa**		**Sochiapam Chinantec**		**Spanish**		**English**	
Unicode Value	**Mark**	**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	1.30%	288	1.4%	183	1.8%	148	1.4%
U+002E	.	177	1.00%	131	0.6%	96	1%	148	1.4%
U+003E**	>	32	0.20%	-	-	-	-	-	-
U+003C**	<	32	0.20%	-	-	-	-	-	-
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%	28	0.1%	17	0.2%	4	0%
U+003A	:	20	0.10%	22	0.1%	17	0.2%	5	0%
U+00A1	¡	2	0%	6	0%	4	0%	-	-
U+0021	!	2	0%	6	0%	4	0%	9	0.1%
U+005D	]	1	0%	2	0%	0	0%	0	0%
U+005B	[	1	0%	2	0%	0	0%	0	0%
U+0029	)	1	0%	0	0%	0	0%	0	0%
U+0028	(	1	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%	36	0.3%
U+2014	—	0	0%	0	0%	0	0%	6	0.1%
U+002D	-	0	0%	0	0%	0	0%	2	0%

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 High tones *(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

2.2 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	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)*	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

2.2.1 History of orthography.

2.2.2 Keyboard.

2.2.3 Results of the Test.

Consonants (Characters without diacritics)									
Characters		**Meꞌphaa**		**Sochiapam Chinantec**		**Spanish**		**English**	
Unicode Value	**Mark**	**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+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	Ꞌ	1	0%	-	-		-	-	-
U+A78C	ꞌ	1,189	7%	-	-	-	-	-	-

2.3 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 High tones *(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

2.3.1 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
Number of total base characters	744	21	1232	2	609	22	1014	2	400	2	596	31

V. Why is it that User Experience Counts? *Usability matters.*

Various solutions

OS Based Solutions

Ukelele - os x
MSKLC - Windows
https://github.com/simos/keyboardlayouteditor - Linux

Third Party Solutions

Keyman
Inkey
http://code.google.com/p/inkey-keyboard-creator/

New Input styles

There are a variety of new input methods which are available on a variety of platforms.
Context sensitivity

iOS and OSX Lion

Android

Swipe keyboard http://www.swype.com/

iOS and drawing Characters

VI. Conclusions

Language documenter's work we need to not only consider the recording documenting and typing of language and cultural data, but we need to also be mindful that this means that our work does have impact.

While some might say that we should make our observations without changing the observed, or at least attempting to not change the observed. A project which does not include

of the

What do we need to consider if we are going to design keyboards for minority languages?

Target characters in the writing environment (orthography and other languages of frequent use)
Frequency of accessibility for each target character
Cross-platform consistency of the user experience, and typist's solution
Operating System being used
FOSS principals
Distribution mechanism
Implementation of solution

National language keyboard layout (is there one? what is “official” what is “common” can it be better? What is better?)
How the character is composed in the mind of the writer (is it a toneme or is it a vowel with a tone? or is a i high tone vowel?)
The Actual Composition of characters

A. Meꞌphaa Text Sample

A̱ ngui̱nꞌ, tsáanꞌ ninimba̱ꞌlaꞌ ju̱ya̱á Jesús, ga̱ju̱ma̱ꞌlaꞌ rí phú gagi juwalaꞌ ído̱ rí nanújngalaꞌ awúun mbaꞌa inii gajmá. Numuu ndu̱ya̱á málaꞌ rí ído̱ rí na̱ꞌnga̱ꞌlaꞌ inuu gajmá, nasngájma ne̱ rí gakon rí jañii a̱kia̱nꞌlaꞌ ju̱ya̱á Ana̱ꞌlóꞌ, jamí naꞌne ne̱ rí ma̱wajún gúkuálaꞌ. I̱ndo̱ó máꞌ gíꞌmaa rí ma̱wajún gúkuálaꞌ xúgíí mbiꞌi, kajngó ma̱jráanꞌlaꞌ jamí ma̱ꞌne rí jañii a̱kia̱nꞌlaꞌ, asndo rí náxáꞌyóo nitháan rí jaꞌyoo ma̱nindxa̱ꞌlaꞌ. [I̱yi̱iꞌ rí niꞌtháán Santiágo̱ 1:2-4]

B. Sochiapam Chinantec Text Sample

Hnoh^{2} reh^{2}, ma^{3}hiún^{13} hnoh^{2} honh^{2} lɨ^{3}ua^{3} cáun^{2} hi^{3} quiunh^{32} náh^{2}, quí^{1} la^{3} cun^{3} hi^{3} má^{2}ca^{3}lɨ^{3} ñíh^{1} hnoh^{2} jáun^{2} hi^{3} tɨ^{3} jlánh^{1} bíh^{1} re^{2} lɨ́^{2}tɨn^{2} tsú^{2} hi^{3} jmu^{3} juenh^{2} tsɨ́^{3}, nɨ́^{1}juáh^{3} zia^{32} hi^{3} cá^{2} lau^{23} ca^{3}tɨ^{21} hi^{3} taunh^{32} tsú^{2} jáun^{2} ta^{21}. Hi^{3} jáun^{2} né^{3}, chá^{1} hnoh^{2} cáun^{2} honh^{2}, hi^{3} jáun^{2} lɨ́^{13} lɨ^{3}tɨn^{2} hnoh^{2} re^{2} hi^{3} jmúh^{13} náh^{2} juenh^{2} honh^{2}, hi^{3} jáun^{2} hnoh^{2} lɨ́^{13} lɨ́n^{3} náh^{2} tsá^{2} má^{2}hún^{1} tsɨ́^{3}, tsá^{2} má^{2}ca^{3}hiá^{2} ca^{3}táunh^{3} ca^{3}la^{3} tán^{1} hián^{2} cu^{3}tí^{3}, la^{3} cun^{3} tsá^{2} tiá^{2} hi^{3} lɨ^{3}hniauh^{23} hí^{1} cáun^{2} ñí^{1}con^{2} yáh^{3}. [Jacobo Jmu^{2} Cáun^{2} Sí^{2} Hi^{3} Ca^{3}tɨn^{1} Tsá^{2} *Judíos, Tsá^{2} Má^{2}tiáunh^{1} Ñí^{1} Hliáun^{3} 1:2-4]

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)]

Bibliography

ABI Research. 20 October 2011. Media Tablets Eclipse Netbook Sales for the First Time in 2Q11. New York: ABI Research.http://www.abiresearch.com/press/3795-Media+Tablets+Eclipse+Netbook+Sales+for+the+First+Time+in+2Q11

Anderson, Stephen P. 2006. Creating Pleasurable Interfaces: Getting From Tasks to Experiences. Paper presented at Refresh '06. <Accessed: 27 January 2012>. http://www.slideshare.net/stephenpa/creating-pleasurable-interfaces-getting-from-tasks-to-experiences

Anderson, Stephen P. 2009. Seductive Interactions. Paper presented at Idea '09. <Accessed: 27 January 2012>. http://www.slideshare.net/stephenpa/seductive-interactions-idea-09-version

Anderson, Stephen P. 2011. Long After the Thrill: Sustaining Passionate Users. Paper presented at #IxD11. <Accessed: 27 January 2012>. http://www.slideshare.net/stephenpa/long-after-the-thrill-sustaining-passionate-users

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

British Standards Institution. 1994. BS 4822:1994 Specification for keyboard allocation of graphic characters for data processing. <Accessed: 27 June 2012>. http://shop.bsigroup.com/ProductDetail/?pid=000000000000116386

British Standards Institution. 2012. BS ISO/IEC 9995-2 AMD1. Information technology. Keyboard layouts for text and office systems. Part 2. Alphanumeric section. <Accessed: 27 June 2012>. http://shop.bsigroup.com/en/ProductDetail/?pid=000000000030255551

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. 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 & Elke Karan. 2008. Factors in designing effective orthographies for unwritten languages. *SIL Electronic Working Papers* 001. Dallas, Tx.: SIL International. <Accessed: 15 May 2012>. http://www.sil.org/silewp/abstract.asp?ref=2008-001

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

Commings, Michael. 16 August 2010. UX Design Defined: User Experience Design Diagram. uxdesign.com. <Accessed: 10 November 2011>. http://uxdesign.com/ux-defined

Durdin, Marc & John Durdin. 2011. Tavultesoft Keyman. Version: 7. Computer program. <Accessed: 15 May 2012>. http://www.tavultesoft.com/

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:

Hosken, Martin. 2001. An introduction to keyboard design theory: What goes where? In Melinda Lyons (ed.), *Implimenting Writting Systems: an Introduction*, Preliminary edn, 123-37. Dallas: SIL International.

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. <Accessed: 12 June 2012>. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39534

International Organization for Standardization. 2009. ISO/IEC 9995-2:2009 Information technology -- Keyboard layouts for text and office systems --. Part 2: Alphanumeric section. <Accessed: 12 June 2012>. http://www.iso.org/iso/catalogue_detail?csnumber=51644

International Organization for Standardization. 2010. ISO/IEC 9995-3:2010 Information technology -- Keyboard layouts for text and office systems -- Part 3: Complementary layouts of the alphanumeric zone of the alphanumeric section. <Accessed: 12 June 2012>. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=52869

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. http://www.humis.utah.edu/humis/docs/organization_919_1300140441.pdf

Jany, Carmen. 2010b. Orthography Design for Chuxnabán Mixe. *Language DocumentatIon & ConservatIon* 4.1: 231-53. http://hdl.handle.net/10125/4481

Jany, Carmen. 2011. Early Stages of Orthography Design: The Case of Chuxnabán Mixe. Paper presented at 2nd ICLDC, Hawaii. <Accessed: http://rocket.csusb.edu/~cjany/Conference_handouts/2nd%20ICLDC,%20Hawaii%202011.pdf

Japanese Industrial Standard. 1980. JIS X 6002:1980 Keyboard layout for information processing using the JIS 7 bit coded character set. <Accessed: 12 June 2012>. http://www.webstore.jsa.or.jp/webstore/Com/FlowControl.jsp?lang=en&bunsyoId=JIS+X+6002%3A1980&dantaiCd=JIS&status=1&pageNo=0

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

Lancor Technologies. 2012. Konyin Product website. <Accessed: 1 June 2012>. http://www.konyin.com/

MarketWire. 27 November 2007. Lagos Analysis Corp. (LANCOR) Files Lawsuit Against Nicholas Negroponte and OLPC Association for Patent Infringement. Negroponte's OLPC Accused of Unauthorized Use of LANCOR's Multilingual Keyboard Technology Invention in XO Laptops of Work. In. <Accessed: 1 June 2012>. http://www.marketwire.com/press-release/lagos-analysis-corp-lancor-files-lawsuit-against-nicholas-negroponte-olpc-association-796745.htm

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

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

Mugele, Robert Louis. 1982. *Tone and Ballistic Syllable in Lalana Chinantec*. Ph.D. dissertation, The University of Texas at Austin, United States -- Texas.

Nzeshi, Ownuka. 23 January 2006. Nigeria: Computer Keyboard With Nigerian Languages Invented of Work. In *AllAfrica*. Cape Town, Dakar, Lagos, Monrovia, Nairobi and Washington DC: AllAfrica. <Accessed: 12 June 2012>. http://allafrica.com/stories/200601230731.html

Open Source Project. 2012. Inkey Keyboard Creator. http://code.google.com/p/inkey-keyboard-creator/

Perre, Laetitia, Chotiga Pattamadilok, Marie Montant & Johannes C. Ziegler. 2009. Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research* 127573-80. http://www.sciencedirect.com/science/article/B6SYR-4W38RK8-2/2/0a8d826c71d21e995d826b7a15344cd2

Pinola, Melanie. 7 April 2011. Apple iPad Hammering Laptop Sales, Analysts Say. *PC World*. <Accessed: 1 June 2012>. http://www.pcworld.com/article/224628/apple_ipad_hammering_laptop_sales_analysts_say.html?tk=rel_news

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: http://scripts.sil.org/cms/scripts/render_download.php?format=file&media_id=IUC26_325&filename=IUC26_325.pdf

Raymond, Martin. 2012. ScriptSource: Making information on the world's scripts and languages accessible. Paper presented at Charting Vanishing Voices, 29-30 June 2012, Cambridge, UK. <Accessed:

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

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

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

Unknown. 2009a. JÚ^{1} CHÚ^{32} QUIOH^{21} JESÚS TSÁ^{2} LɨŃ^{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} LɨŃ^{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

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

Zheltov, Pavel V. 2005. Minority languages and computerization. The situation in the Russian Federation. *OGMIOS* 3.3: 8-11. http://www.ogmios.org/ogmios/Ogmios_027.pdf

Ziegler, Johannes C., Ludovic Ferrand & Marie Montant. 2004. Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition* 32.5: 732-41.

Ziegler, Johannes C. & Mathilde Muneaux. 2007. Orthographic facilitation and phonological inhibition in spoken word recognition: A developmental study. *Psychonomic Bulletin & Review* 14.1: 75-80.