Evaluation of Virtual Keyboards for West-African Languages

Chantal Enguehard[1], Harouna Naroua[2]

[1] chantal.enguehard@univ-nantes.fr
Laboratoire d'informatique de Nantes Atlantique, Université de Nantes
2, rue de la Houssinière, BP 92208, 44322 Nantes Cedex 03, France

[2] hnaroua@yahoo.com
Université Abdou Moumouni, Département de Mathématiques et Informatique
BP.10662, Niamey, Niger

Abstract

West African languages are written with alphabets that comprize non classical Latin characters. It is possible to design virtual keyboards which allow the writing of such special characters with a combination of keys. During the last decade, many different virtual keyboards had been created, without any standardization to fix the correspondence between each character and the keys to press to obtain it.

We define a grid to evaluate such keyboards and apply it to five virtual keyboards in relation with the five main languages of Niger (Fulfulde, Hausa, Kanuri, Songhai-Zarma, Tamashek), Bambara and Soninke from Mali and Dyoula from Burkina Faso.

We conclude hat the African LLACAN keyboard should be recommended in Niger because it covers all the characters used in the alphabets of the main languages of this country, it produces valid Unicode codes and it minimizes the number of keys to be pressed.

1. Context

Localization of software has become an important activity since the Unicode standard appeared in 1992 which gives a unique code to each character (Andries, 2002).

Fifty years after the Bamako conference which started work on the definition of the official alphabets of many African languages, these alphabets appear now to be stable. Official decrees which fix these alphabets have been published for the last ten years in Niger and Mali.
These alphabets are Latin based but also include several characters which cannot be found amongst the 26 alphabetical letters required to write in English (Chanard, 2001), (Mbodj, 2002). This situation is not unusual. Several European languages, such as French, Spanish and Czech make use of additional characters or characters with diacritics.

Keyboards are the main devices used to enter text into computers. Different types of keyboards have been designed, initially for typewriters, and now for computers. Keyboards may differ according to which language will be used (David, 1998). This paper will deal with the historical development of keyboards (Lilen, 2003).
Producing a material keyboard should be a profitable business but this may not be possible in the case of many of the less widely spoken languages where the market is too small.

In such cases it is possible to design virtual keyboards: the Unicode code which is mapped to each key or combination of keys can be easily modified. This approach allows the addition or the substitution of signs to standard keyboards.

Because there was a strong need for such localized keyboards, the past decade has seen the definition of many keyboards for West African languages, without any standardization. This situation is quite confusing because the users of these keyboards are faced with different keyboards, and many of them may find the situation very uncomfortable.

2. Goal

The story of the Qwerty or Azerty keyboards (used to write in English and French, respectively) shows that the design quality of a keyboard (the way the keys are situated on the keyboard) is not the most important feature. These two keyboards, even if they are not properly designed, became de-facto standards because they are widespread and because it is extremely difficult for a user who has learned to type on one keyboard, to become accustomed to another.
This research does not intend to design keyboards but rather to evaluate different virtual keyboards by defining an evaluation grid and applying its criteria to the different virtual keyboards which we have decided to compare.

3. Evaluation grid

We defined several criteria to evaluate keyboards.

Firstly, it is important to determine whether the keyboard has the capacity to associate Unicode codes for each key or combination of keys which corresponds to a character,

1

and not ASCII codes.

Secondly, Western African countries generally have several national languages and one official language. Several of these countries, like Mali or Niger, started to include their national languages in education programmes and these languages are widely used by Non Governmental Organisations. These activities produce written documents which are produced in coordination with linguists and typed by the secretaries of these organisations. A secretary may frequently be required to type texts in different languages. It is then important to have a single keyboard which may be used to produce texts in as many languages of a country as possible, including the official language.

Thirdly, it is important that all the characters on the original keyboard continue to be accessible on the virtual one.

Fourthly, the number of keys that have to be pressed to produce a text must be kept as low as possible.

Fifthly, a user-friendly documentation must accompany the keyboard.

The three first criteria we chose are mandatory. The two last criterion can be used to decide between keyboards that fill the three mandatory criteria.

Even if these criteria are quantitative, the final evaluation should be qualitative, based on the five previously established measures.

We did not consider as a criterion the ergonomic key placement on the keyboard because history taught us that this feature is not crucial for the massive use of a specific keyboard, as we saw with the cases of the Qwerty and Azerty keyboards. For instance, the Dvorak Simplified Keyboard (DSK) that were developed by August Dvorak to be ergonomically adapted to English, never succeed to dethrone the Qwerty keyboard. The French version of this keyboard, ergonomically adapted to French failed also to become essential (Liebowilz, 1998).

4. Experiments

We have evaluated five virtual keyboards for the five main national languages of Niger which are used in educational programmes. In addition, we examined two languages of Mali and one of Burkina Faso.

The five virtual keyboards are: AfricainLLACAN, AfricaKeyboardDead, AfricaKeyboardUS, Pan-Sahelien, SIL_IPA_Unicode.

The five languages of Niger are: Fulfulde (ful), Hausa (hau), Kanuri (kau), Songhai-Zarma (son), Tamashek (tmh).
We also added Bambara (bam) and Soninke (snk) from Mali and Dyoula (dyu) from Burkina Faso.

The transcription of these languages encounters several signs which are not accessible on usual keyboards as shown in Table 1.

Name	Sign	Languages	Unicode
Latin letter e with tilde	ẽ	son	U+0065 U+0303
Latin letter i with tilde	ĩ	son	U+0069 U+0303
Latin letter o with tilde	õ	son	U+006F U+0303
Latin letter r short stroke overlay	r̵	kau	U+0072 U+0335
Latin letter u with tilde	ũ	son	U+0075 U+0303
Latin letter a with tilde	ã	son	U+00E3 or U+61 U+303
Latin letter a with breve	ă	tmh	U+0103 or U+61 U+306
Latin letter ENG	ŋ	bam, ful, son	U+014B
Latin letter s with caron	š	tmh	U+0161 or U+73 U+30C
Latin letter k with hook	ƙ	hau	U+0199
Latin letter y with hook	ƴ	ful, hau	U+01B4
Latin letter turned e	ə	kau, tmh	U+01DD or U+259
Latin letter g with caron	ǧ	tmh	U+01E7 or U+67 U+30C
Latin letter j with caron	ǰ	tmh	U+01F0 or U+6A U+30C
Latin letter b with hook	ɓ	ful, hau	U+0253
Latin letter open o	ɔ	bam	U+0254
Latin letter d with hook	ɗ	ful, hau	U+0257

2

Name	Sign	Languages	Unicode
Latin letter gamma	ɣ	tmh	U+0263
Latin letter epsilon	ɛ	bam	U+025B
Latin letter n with retroflex hook	ɲ	bam, son	U+0272
Latin letter d with dot below	ḍ	tmh	U+1E0D or U+64 U+323
Latin letter l with dot below	ḷ	tmh	U+1E37 or U+6C U+323
Latin letter s with dot below	ṣ	tmh	U+1E63 or U+73 U+323
Latin letter t with dot below	ṭ	tmh	U+1E6D or U+74 U+323
Latin letter z with dot below	ẓ	tmh	U+1E93 or U+7A U+323

Table 1: special characters used in the alphabets of the studied languages

5. Results

The evaluation of the number of keys that have to be pressed to produce a text has been determined using corpora of the different languages.

Virtual keyboard	Fulfulde	Hausa	Kanuri	Songhai-Zarma	Tamashek	Total
AfricainLLACAN	87,066	163,225	109,542	82,234	213,230	655,297
AfricaKeyboardDead	87,871	165,551	109,593	85,606	217,512	666,133
AfricaKeyboardUS	87,871	165,551	109,593	85,606	217,512	666,133
Pan-Sahelien	91,329	168,623	109,703	87,989	221,903	679,547
SIL_IPA_Unicode	89,048	165,001	108,175	82,866	213,850	658,940

Table 2: number of keys that have to be pressed for languages from Niger

Virtual keyboard	Unicode	Fulfulde	Hausa	Kanuri	Songhai-Zarma	Tamashek
AfricainLLACAN	Yes	Total	Total	Total	Total	Total
AfricaKeyboardDead	Yes	Total	Total	Total	Partial no Ɲ (13/14)	Partial (14 / 24)
AfricaKeyboardUS	Yes	Total	Total	Total	Partial no Ɲ (13/14)	Partial (14 / 24)
Pan-Sahelien	No	None	None	Partial (1 / 4) (ə only)	None	Partial (16 / 24)
SIL_IPA_Unicode	Yes	Partial (4 / 8)	Partial (3 / 8)	Partial (1 / 4) (ə only)	Partial (4 / 14)	Partial (6 / 24)

Table 3: code types and coverage of special characters of languages from Niger

6. Results

The Pan-Sahelien keyboard must be disqualified because it does not produce Unicode codage.
The AfricaKeyboardDead, AfricaKeyboardUS and SIL_IPA_Unicode keyboards must be also disqualified because they don't cover all the characters that are used in the five languages we consider.

Thus, The African LLACAN keyboard is the only one to meet the three mandatory criteria we defined. In addition, it is provided with a documentation. written in French.

This documentation should be translated in other languages (of course in English, but also in the West-African languages that can be easily written with this keyboard).

The African LLACAN keyboard can be downloaded on the Tavulsoft website[1]: choose Niger, then any of the languages that has been studied in this paper

---

1http://www.tavultesoft.com/keyman/downloads/keyboards

3

7. Conclusion

The African LLACAN keyboard should be recommended in Niger because it covers all the characters used in the alphabets of the different languages of this country and produces valid Unicode codes. In addition, this keyboard minimizes the number of keys to be pressed and covers also all the alphabetic characters of Bambara, Soninke and Dyoula.

It will soon be fully evaluated for the main languages of Mali.

References

Andries, P. (2002). Introduction à Unicode et à l'ISO 10646. Document numérique, vol.6, n°3-4, pp.51-88.

Chanard, C. and Popescu-Belis A. (2001). Encodage informatique multilingue : application au contexte du Niger. Les cahiers du RIFAL, n°22, pp.33-45.

David, P. A. (1998). Comprendre les aspects économiques de qwerty la contrainte de l'histoire. Réseaux, French Journal of Communication, No 87 CNET.

Liebowilz S. J. and Margolis Stephen E. (1998). La fable des touches. Réseaux, French Journal of Communication, No 87 CNET.

Lilen, H. (2003). La saga du micro-ordinateur. Vuibert, Paris.

Mbodj, C. (2002). Orthographe commune et législations nationales. Writing African – The Harmonisation of Orthographic Conventions in African Languages, ed. Kwesi Kwaa Prah, pp. 55-64.