



# Requirement and Functionality of Multilingual Extensions to European (Latin) Keyboard Layouts

Lunchtime Conference

31 January 2008

Erkki I. Kolehmainen, [eik@iki.fi](mailto:eik@iki.fi)

# Case: Europe

- Large number of different languages
- The same language (or variants thereof) may be spoken in several countries with different status
- Multiple scripts, even for one language; focus here on the Latin script
- Very large number of letters with diacritics, some of which are perceived as basic letters by the local population, and also collated separately (e.g. å, ä and ö after z in Sweden and Finland)

# Case: European Union

- Multiple official EU languages (plus some for limited EU usage only)
- Common market with free movement of goods and, more recently, also people
- Names of people, companies, products, and places to be written correctly for legal accuracy and politeness (in original form/transcription)
- Lists of ingredients and other consumer material must be available in designated languages in each country; often prepared for distribution in multiple countries

# International conventions

- UNESCO Convention on the Protection and Promotion of the Diversity of Cultural Expressions
- European Charter for Regional or Minority Languages
- Convention relative à l'indication des noms et prénoms dans les registres de l'État Civil (CIEC 14)
- Bilateral, e.g., German-Polish Treaty

# UNESCO Convention...

## UNESCO Convention on the Protection and Promotion of the Diversity of Cultural Expressions

- Approved in October 2005 in the UNESCO general assembly by an overwhelming majority (148 to 2)
- Entered into force in March 2007

<http://ec.europa.eu/culture/portal/action/diversity/pdf/st08668.EN06.pdf>

# UNESCO Convention: Objectives (1)

- to protect and promote the diversity of cultural expressions;
- to create the conditions for cultures to flourish and to freely interact in a mutually beneficial manner;
- to encourage dialogue among cultures with a view to ensuring wider and balanced cultural exchanges [...]
- to foster interculturality [...]
- to promote respect for the diversity of cultural expressions and raise awareness of its value at the local, national and international levels;

# UNESCO Convention: Objectives (2)

- to reaffirm the importance of the link between culture and development for all countries [...]
- to give recognition to the distinctive nature of cultural activities, goods and services as vehicles of identity, values and meaning;
- to reaffirm the sovereign rights of States to maintain, adopt and implement policies and measures that they deem appropriate for the protection and promotion of the diversity of cultural expressions on their territory;
- to strengthen international cooperation and solidarity in a spirit of partnership [...]

# European Charter...

## European Charter for Regional or Minority Languages

- Signed in 1992 and in force since 1996
- Concentrates on languages traditionally used in a state by a minority and not an official language of that state
- Ratification of this Council of Europe treaty for regional and minority languages (by the national parliaments) entails certain responsibilities, details of which are specified upon ratification

<http://conventions.coe.int/Treaty/en/Treaties/Html/148.htm>

# European Charter: Key Requirements

- the facilitation and/or encouragement of the use of regional or minority languages, in speech and writing, in public and private life (§7.1.d)
- the provision of appropriate forms and means for the teaching and study of regional or minority languages at all appropriate stages (§7.1.f)
- to ensure that users of regional or minority languages may submit oral or written applications [to public authorities] and receive a reply in these languages (§10.1.a.iii)
- to ensure that users of regional or minority languages may validly submit a document in these languages (§10.1.a.v)
- to allow the administrative authorities to draft documents in a regional or minority language (§10.1.c)

# Convention CIEC 14

Convention relative à l'indication des noms et prénoms dans les registres de l'État Civil (CIEC 14)

- Signed in 1973 in the context of the Commission Internationale de l'État Civil (CIEC)
- Ratified by Austria, Belgium, France, Germany, Greece, Italy, Luxemburg, the Netherlands, Switzerland and Turkey
- The official French text and the unofficial English translation are available at:

<http://www.ciec1.org/ListeConventions.htm>

# CIEC 14: Rendering of Names

- Where a record is to be made in a civil register by an authority of a Contracting State and there is produced for that purpose a copy of or extract from a civil status record or some other document that shows the surnames and forenames in the same characters as those used in the language in which the record is to be made, those surnames and forenames shall be reproduced literally without alteration or translation.
- Any diacritic marks forming part of such surnames and forenames shall also be reproduced, even if such marks do not exist in the language in which the record is to be made

Article 2 of CIEC 14 (unofficial English translation)

# CIEC 14: Use of Transliteration

- Where a record is to be made in a civil status register by an authority of a Contracting State and there is produced for that purpose a copy of or extract from a civil status record or some other document that shows the surnames and fornames in characters other than those used in the language in which the record is to be made, those surnames and forenames shall be reproduced as far as possible by transliteration, without being translated.
- If there are standards recommended by the International Organisation for Standardisation (ISO), they shall be applied.

Article 2 of CIEC 14 (unofficial English translation)

# German-Polish Treaty ...

Treaty on good neighbourly relations and friendly cooperation of 17 June 1991

**Art. 20 § 3** ... shall in particular have the right, individually or in community with other members of their group:

- To use their mother tongue freely, both privately and in public, to disseminate and exchange information in it and to have access to such information; ...
- To express their given and family names in the mother tongue form; ...

§ 4 ... membership in the groups referred to in § 1 ... may not place him at any disadvantage.

[http://untreaty.un.org/unts/60001\\_120000/28/1/00054050.pdf](http://untreaty.un.org/unts/60001_120000/28/1/00054050.pdf)

# CEN/ISSS and CDFG

- CEN/ISSS (European Information Society Standardization System) has produced several CEN Workshop Agreements
  - Multilingual European Subsets (MES) of the UCS (MES-2 as collection 282 in 10646 A.4.2)
- CDFG (Cultural Diversity Focus Group) is the driving force in CEN for cultural diversity issues
- CEN/ISSS CDFG is a Unicode Liaison member, like many National Standards Bodies

# Case: European keyboards

- Own keys for local non-a-to-z letters
- Other letters with diacritics are usually formed with dead-letter keys (ever since mechanical typewriters), except in the English keyboards
- Current layouts support only a limited repertoire – design based on 8-bit encoding schemes
- Although it is feasible (and relatively easy) to switch between keyboard drivers for different scripts, users would not adapt to switching for different languages within the same script

# The “Kotoistus” Initiative

- Aimed to support cultural requirements in ICT in Finland
- Kicked-off with special funding by the Ministry of Education at the Research Institute for the Languages of Finland in the fall of 2004 (as the result of a guest editorial in Helsingin Sanomat in May), and transferred in the fall of 2006 to CSC, the Center for Scientific Computing, and included in their regular operating plan
- Environmental background:
  - Unicode had assumed the responsibility for the development and maintenance of the Common Locale Data Repository (CLDR) project in April 2004
  - ICANN had allowed for internationalized domain names, IDNs
- The Working Group and its Steering Group consist of well over 40 individuals, mostly representing various organizations

# Finnish/Swedish Keyboards

- Finland and Sweden have traditionally had a common de facto keyboard layout and hardware
- Sámi Parliamentary Council (Sámediggi) decided on keyboard layouts to support Sámi (also for speakers of the majority languages!)
- Swedish Gov't decided on the Sámediggi layout as the base to facilitate the extension of the population registry (and created a Sw. standard)
- Finnish WG on cultural diversity issues in ICT decided to rather aim for use by general public (while hoping to maintain the economies of common keyboard engravings with Sweden)

# Finnish Multilingual Keyboard

- Refined support for the Finnish orthography, including proper punctuation
- Capability to enter all names (individuals, places, and products) and quotations in major European languages (in Latin script)
- Support for all Nordic languages, incl. regional and minority languages
- Support for a choice of punctuation marks, facilitating proper writing of text in e.g. English, French, German, and Spanish (by the speakers of Finnish or Swedish)

# Finnish Multilingual Keyboard

- Wide coverage, yet intuitively comprehensible and thus easy to learn-and-use approach
- Extensive use of dead-letter combinations for letters with diacritics (or stroke), even if the target character is not decomposable
- Specific key allocations for only those “foreign” letters that cannot be combined on the keyboard
- All “required” letters (with one diacritic) have precomposed encodings, in order to provide for easier processing in legacy-rooted systems

# Development and Review

- Development done in a KB Working Group within the fully open Kotoistus framework
- Approach presented and debated in a BOF (bird-of-a-feather) session at IUC (Internationalization and Unicode Conference) 29 on 6 March 2006 in San Francisco, and discussed further with the industry on several occasions
- The current design has been greatly influenced by the comments from industry and users alike
- The evaluation versions (Windows by Microsoft, Linux) have received highly positive publicity by the press and similar feedback from the users; a national standard is being developed
- The Finnish experience will be available to WS/MEEK

# New Finnish Keyboard Layout

2005-12-14

## Appendix 2

Allocations (dead keys in grey)

	00	01	02	03	04	05	06	07	08	09	10	11	12
E	½ §	! i 1	" " 2 @	# » 3 £	« 4 \$	% “ 5 ‰	& ” 6 ,	/ {	( [	) ]	= ° 0 }	? ¿ + \	ò ø ó õ
D		Q q	W w	E € e €	R r	T Þ t þ	Y y	U u	I i	O Æ o æ	P ö p ö	Å å å å	Ô õ ö ö
C		A ə a ə	S s β	D Ð d ð	F f	G g	H h	J j	K k κ	L l	Ö Ø ö ø	Ä Æ ä æ	* ö ' õ
B	> <	Z 3 z 3	X · x x	C c	V v	B b	N Ŋ n ŋ	M – m μ	; ‘ , ’	: ö . ø	– ö - -		
	<div style="border: 1px solid black; padding: 5px; display: inline-block;">NB SP</div>												

