

Public Policies and Multilingualism in HCI

Loïc Martínez Normand, Universidad Politécnica de Madrid, Fabio Paternò, CNR-ISTI, Marco Winckler, Université Paul Sabatier

The design principle “speak the user’s language” underpins the universal idea that interactive systems should be able to process and display information not only in the user’s native language but also in an understandable and meaningful way, and, if possible, should also accommodate the user’s cultural background. This is not simple in a world that counts more than 7,000 known living languages [1]. On top of this is the debate among scholars over the distinction between languages and dialects. Living languages evolve over time, and new minority languages are discovered and lost virtually in the time it takes to read this paragraph [2].

Regardless of the exact number of languages in the world, it is interesting to note that, as far as history can tell, mankind has always been multilingual. Some ancient knowledge has reached us only through the availability of multilingual documents (like the Rosetta stone—c. 196 B.C.—and the Arabic texts of the Greek Galen of Pergamon—c. 200 A.D.). Currently we are more likely to use a digital medium to store knowledge, but the questions about the choice of the language remain.

The quip “a language is a dialect with an army and navy” reported by Max Weinreich (1894–1969) illustrates the weight countries and governments place on the use of dominant languages and the

preservation of minority ones. Our concern here is not to discuss how languages are important in the construction of national identities but rather how public policies influence the use of language in the design of interactive systems. What is the impact of such policies when questions about language use affect the design of interactive systems that should be made available to citizens? Why should we be concerned about multilingual issues in a globalized world where English has become the lingua franca for most scientific exchanges, at least in computer science?

Indeed, adaptation of a system for a specific region involves various language-related aspects, including the availability of translations and support for different characters and scripts, as well as cultural aspects, such as units, currencies, time and date formats, meaning of colors and symbols, suitability of images, and legal requirements.

Hereafter we illustrate how some countries address these issues and

Insights

Multilingualism and public policies on language use are important in HCI. Different countries address multilingual issues in different ways, from laws to design guidelines. The application domain has impact on the choice of languages for the user interface. Language use is also related to accessibility for persons with disabilities.

pinpoint the challenges for creating public policies around them.

PUBLIC POLICIES FOR USE OF LANGUAGES

Currently most countries worldwide have recommendations, laws, and/or regulations about the use of the languages in their territories. The issues raised often concern general questions related to the use of one or more official languages and the preservation of minority languages and/or dialects. Such initiatives regulate the communication between citizens and administrations, which ultimately have a huge impact on digital communication and, in particular, administrative websites.

Multilingual countries and official regional languages. Some countries, such as Belgium, Canada, India, and Switzerland, are well known for being multilingual (i.e., having more than one official language). Nonetheless, these are not isolated examples, and countries such as Spain have as many as four official languages. The Spanish constitution recognizes Castilian (Spanish) as the common official language and then recognizes that the autonomous communities (regions) can define their own languages, which become co-official in those regions. Today those co-official languages are Catalan (Catalonia, Balearic Islands, and Valencia), Galician (Galicia), and Euskera (Basque Country and Navarra). In contrast, other countries such as the U.S. do not have an official language at the federal level but have to deal with states with

two official languages (e.g., Hawaii with English and Hawaiian) and de facto bilingual states (e.g., Texas and New Mexico, where both English and Spanish are widely used).

In multilingual countries there are usually public policies that mandate the use of those co-official languages by the public administration. These policies tend to be associated with citizens' rights to use their own language when communicating with the public administration. The mandatory use of co-official languages includes ICT-related cases such as public websites or software applications

support the co-official languages, to fully translate the homepage of public websites to all the co-official languages, and to fully translate all forms and standardized document models related to administrative processes. In addition, the guide has many recommendations. Some are relevant to HCI, such as “to incorporate in the homepage links to change to any of the languages offered” and “to use the word welcome in the original language to indicate language switch”—*Bienvenido* (Spanish), *Benvinguts* (Catalan), *Benvidos* (Galician), and *Ongietorri* (Euskera).

a means to implicitly expose users to the official language of Quebec, so users must explicitly select an alternative language.

Nonetheless, not all countries have specific guidelines for the use of languages in their territory. This is particularly tricky with languages not officially recognized by law.

Full translation of documents into all living languages seems utopian, but even on a smaller scale, multilingualism is problematic. For example, the European Union recognizes 24 “official and working” languages: Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish, and Swedish. There are two main entitlements for languages with “official and working” status: documents may be sent to EU institutions and a reply received in any of these languages, and EU regulations and other legislative documents are published in the official and working languages, as is the Official Journal.

The implication for HCI of recognizing these 24 languages in Europe is that official websites of the EU must be multilingual and need to be translated into these 24 languages. The EU has an inter-institutional style guide that, among other things, provides information about how to officially identify these languages, in long and abbreviated forms. The long form is the name of the language written in that language (e.g., English, Français, Italiano, Español), as shown on the website in Figure 1. The short form is a two-letter code taken from the ISO codes 639-1 in their alpha-2 format. The style guide also mandates alphabetizing the formal titles in their original written forms.

Choices of the languages depending on the application domain. Nowadays, English is widely regarded as being the global language. Indeed, English is the de facto official language of science and not only is used in international forums but also has been adopted in



PHOTOGRAPH BY TOM FASSBENDER

provided by the government. This obligation has a significant impact on the design of those systems, which must be multilingual.

For example, in Spain there is a legal obligation to use co-official languages in the ICT of the national public administration. Act 30/1992 [3] mandated that the national public administration must allow citizens to use co-official languages if they live in any of the autonomous communities that have co-official languages. Spain has published a guide for digital communications [4] with several obligations: to

Specific guidelines for the use of language in public websites can also be found in Canada, where in 2006 the Quebec government published the “Standard for Multilingual Web sites/*Standard sur les sites Web multilingues* (SGQRI 011),” which regulates the use of the languages to be used on government websites. The SGQRI 011 standard defines French as the default language, even for multilingual websites. It is interesting to note that the guidelines recommend disabling the automatic verification of the preferred user language as



Figure 1. Languages on the EU Web portal.

national scientific forums (e.g., the Brazilian HCI conference accepts contributions in English in addition to Portuguese; the official language of the Italian HCI conference is English). Research institutions such as CNR-ISTI in Italy have websites only in English because they are used to working in international research contexts. Sometimes this raises discussions when people complain about Italian public research organizations not using the national language for their websites.

However, if we look at the websites of the Italian ministries, there are some (such as agriculture and education) only in Italian; some in Italian, English, and French (such as defense); and some in Italian, English, and even Arabic—this is the case with the foreign affairs ministry, which must handle requests from many immigrants who speak this language. Since Arabic is read from right to left, this has an impact on the layout of Web pages.

Some public policies concern the use of languages in electronic commerce. A good example of this is the Italian regulation on the matter of remote contracts (Decretolegislativo 22 maggio

1999, n. 185, [5]), which recognizes the user's/client's right to see the contents of the contract (e.g., information regarding the seller, the good or service bought, mode of delivery, etc.) in his or her mother tongue.

In Spain, the guide for digital communications also provides counsel for the use of foreign languages [4]. It recommends deciding which parts of a website are to be translated and how, depending on the expected audience—for example, tourists or foreign persons living in Spain. The guide has several explicit recommendations about designing for a specific audience, need of official translations, how to design the language-selection area (e.g., by using the word *welcome* as a link to the English version of a site), and so on.

In France, the Toubon law (No.

94-665 of August 4, 1994) forbids the sale of goods and services in France in any language other than French unless accompanied by a French translation. This law raised several important questions regarding the Internet and has heavily influenced the labor code in France, which has extended that obligation to the use of French in job offers, contracts, conventions, agreements, internal rules, and any other document “required by the employee to perform his tasks.”

This will ultimately affect all operational procedures in software, including screen displays and audio messages used as instructions for employees. In 1996 a lawsuit based on the Toubon law was filed against a Georgia Tech University campus located in the city of Metz, France, whose website provided information only in English. The court dismissed that lawsuit against Georgia Tech in 2007, based on technicalities of procedural grounds (Georgia Tech provided versions of the website in French, German, and English afterward). Nevertheless, two important lawsuits based on the Toubon law followed: General Electric Medical Systems (GEMS) [6] and Europ Assistance provided their French-speaking employees with information systems only in English. Both companies were condemned in different instances to indemnify employees and translate those systems into French. Interestingly, potential safety problems related to the use of the English language by non-native speakers created a judicial precedent for the need for software translation.

In recent years there have been no new big lawsuits based on the Toubon law, but the law came up recently (May 2013) in the French Assembly during discussions about regulating teaching in foreign languages in French universities. Up until now, French universities were supposed to teach only in French, except for foreign language classes. However, amendments to the Toubon law have been proposed to authorize regular classes being taught in foreign languages as well, which actually had already been

occurring (often in business schools and institutes such as SciencePO and the Instituts d'études politiques - IEP). This may be an ongoing policy debate in the near future.

Standardization issues with multilingual applications. To provide more general support for addressing problems related to the standardization of languages used on the Web, the W3C created the W3C Internationalization (I18n) Activity and Multilingual Web initiative. The goal of the I18n Activity is to ensure that the W3C's formats and protocols are open to all of the world's languages, writing systems, character codes, and local conventions. I18n advises W3C Working Groups and coordinates with the Unicode Technical Committee, the IETF, ISO committees, and the localization industry. I18n provides input on a wide range of topics, including Unicode character normalization, international typographic requirements, script issues in text-to-speech implementations, internationalization and localization requirements for schemas, usage scenarios and requirements for the internationalization of Web services, implementation of international resource identifiers, and many more. (For the curious, "I18n" is shorthand for the first, last, and 18 middle characters in the word *internationalization*.)

Languages and accessibility. Last but not least, we have to consider guidelines for language-related accessibility for persons with disabilities. This mainly affects persons who are deaf, who face two problems. First, they may have difficulties with reading and understanding written language. Due to the lack of hearing, born-deaf persons learn languages in a different way and may lack knowledge of the syntactic and semantic structure of the written language. This implies limited vocabulary, difficulty in reading comprehension, and sometimes difficulty in using search engines. Thus, guidelines for writing text that is suitable for deaf persons indicate aspects such as using simple words, always using the same word

to refer to a single thing, including only text that is clear and brief, placing a glossary of terms used most frequently on the site, and explaining in a simple way how to formulate a request to be entered into a search engine. In fact, these guidelines are strongly related to the concept of easy-to-read text ("plain language"), which benefits everyone, including persons with cognitive disabilities. We can even say that improving the accessibility of written text for the deaf would increase the usability of sites of public utility for all users!

The second language-related problem for deaf persons is the fact that many of them use sign languages as their mother tongue. This implies that to provide good accessibility for them it would be necessary to include content in sign language. Unfortunately, in many countries this need of sign language is not covered by public policies, such as the U.S. Americans with Disabilities Act, which usually require captioning for videos but not sign languages. Sign languages are not written languages and can be included in the user interface only as videos of signing persons or as animations of signing avatars. So the knowledge for multilingual user interfaces cannot be applied, and new approaches are needed.

CONCLUSION

Language is an essential aspect of human communication and interaction, not only between humans but also between humans and interactive systems. Indeed, the use of language is a keystone for the design of interactive systems. Inappropriate use of language might limit the access of users to information and induce users to make mistakes or to lose control of interactive systems. Despite the existence of public policies, the diversity of languages poses serious problems when considering full-fledged/seamless support for all existing languages. And yet standardization processes have successfully defined mechanisms for ensuring cross-platform compatibility between languages, at

least at the level of format and the set of characters that can be used.

ENDNOTES

1. Ethnologue Languages of the world; <http://www.ethnologue.com/>
2. Anderson, S.R. How many languages are there in the world? Linguistic Society of America. 2010; <http://www.linguisticsociety.org/content/how-many-languages-are-there-world>
3. Spanish Government. Ley 30/1992, de 26 de noviembre, de Régimen Jurídico de las Administraciones Públicas y del Procedimiento Administrativo Común (in Spanish). 1992; <http://www.boe.es/boe/dias/1992/11/27/pdfs/A40300-40319.pdf>
4. Spanish Government Guía de comunicación digital para la administración general del estado (in Spanish). 2013; <http://bit.ly/1jZV5fe>
5. Italian Government (1999). Decreto Legislativo 22 maggio 1999, n. 185. Attuazione della direttiva 97/7/CE relativa alla protezione dei consumatori in materia di contratti a distanza (in Italian). 1999; <http://www.parlamento.it/parlam/leggi/deleghe/99185dl.htm>
6. Cours d'Appel Versailles. N°05-1244, 1ère chambre, 1ère section, Sté GE Medical Systems c/ Comité d'entreprise GE Medical Systems SCS et al. March 2, 2006; <http://bit.ly/1pOobNU>