Multilingual Internet: Designing Internet Sites in a Variety of Languages

William L. Boletta
School of Letters
Senshu University

Abstract*

This paper offers some suggestions and solutions for designing and planning multilingual web sites. The Internet had its origins in America, but it now extends over the entire world. While English has become the lingua franca of the Internet, there is often a need for sites that employ more than one language. Such multilingual sites present design challenges that do not arise in monolingual site planning. Typography and navigation often present special problems for the multilingual site designer. The level of complexity added by having several language alternatives to choose from has the potential to be confusing to users. We examine some of these general problems and offer some concrete solutions for designing a coherent interface that is both transparent to the individual user yet serves the needs of multilingual audiences. Keywords: Internet, World Wide Web, Site Design, HTML.

The Internet: An English-Only Medium?

The purpose of this short paper is to look at some of the possibilities for designing materials for the Internet, specifically the World Wide Web, which employ more than one language. The rapid rise in importance of the Internet and the World Wide Web as a means of communication between nations has been one of the most remarkable developments of the twentieth century. ARPANET, the progenitor of the Internet, was a project of the United States Defense Department, and

*The author wishes to express his gratitude to the Institute of Information Science, Senshu University, for a generous research grant which funded research for this article.
had goals that were primarily strategic and intimately connected with security considerations within the United States. Cold War era fears of imminent nuclear attack by a hostile power (namely, the Soviet Union) gave rise to the idea that vulnerable computer installations containing data vital to national security interests should be decentralized in such a way as to spread their vulnerability out, so to speak, and lessen the risk of complete data shutdown in the event of an attack by hostile missiles or other such destructive weapons. In order to accomplish this goal, computer nexuses were set up around the United States in such a way that, should one be destroyed, others could equally well communicate crucial data. It is this very decentralization that gave rise to the networking capabilities that are the most characteristic feature of today’s Internet. Subsequently, of course, the academic establishment in the United States, primarily scientific researchers, took advantage of the networking setup, and eventually access to the Internet was extended to other academics, and ultimately to the world.

It has become a virtual truism that the language of the Internet is English. This may be partially due to the American background and origins of the Internet, but a further reason is the role of English as the lingua franca of international communication. I do not wish to enter here into a discussion of the merits and demerits of English and the “world language,” nor do I intend to discuss the politics of English as a worldwide second language. I do wish, however, to question the assumption that the Internet equals English. This is demonstrably untrue, as any brief excursion to Internet sites originating in countries where English is not an official language will show. While English may be the default language of the Internet, so to speak, there are other languages in the world, and the speakers of those languages wish to communicate with each other in them. Also, while English is the dominant second language today, this was not so in the recent past. Latin was the lingua franca of the educated person in the Middle Ages; French was the primary language of diplomacy until very recently; and German was a language widely familiar
among academics and researchers. Many people still know these languages as second languages, but more to the point is the reality that communication within a country is usually carried on in the native language. India, the Philippines, Singapore, and certain African countries, because of their linguistic plurality, do use English widely as a lingua franca, but widely spoken native languages often fill this role as well (e.g., Hindi in India).

**Some General Guidelines for Multilingual Sites**

While multilingual sites share the common characteristics of all Internet page design, they have certain peculiarities which one doesn’t encounter in sites that have only one language. We can group these special considerations roughly into two groups: typography and navigation. Typography here refers primarily to the character sets used in the various languages. While virtually all publications in industrialized countries are now generated by means of computer typesetting, the typesetting capabilities of the web page designer are usually limited by the kinds of software available on desktop computing systems. Even the more advanced systems fully outfitted with all of the bells and whistles of the “power user,” are not always as extensive in their functional application as proprietary typesetting systems used in professional printeries. As more and more typesetting is done on desktop computers, however, the limitations of particular standard configurations are becoming progressively more uniform.

Navigation refers to the architecture of a site and the various links between pages and groups of pages on the site. One of the chief problems of multilingual site design is forging a path, as it were, for the user of a specific language to follow when visiting various areas of the site which are language-specific while steering clear of other areas which are in languages the user does not wish to read. One of the prime strategic locations in a multilingual site is the main menu page, where the user makes a key decision about which language she will use.
Typographical Considerations

The writing systems employed in some of the world’s languages have always presented huge challenges to students unfamiliar with them. Alphabetic languages such as Russian, Hebrew, and Arabic, while perhaps difficult to master for beginners who are learning to read and write them, nonetheless have a certain comfortable finitude about them. Phrased differently, in digital jargon, they have specific and limited character sets, running at most into the low hundreds. For the manual student they are limited in the demands they make. Nonalphabetic languages, such as Japanese and Chinese, are quite a different matter. Japanese, particularly, presents problems for the student and for the computer. Japanese uses not only logographic characters adapted from Chinese, but also employs two extensive syllabaries as well as the Roman alphabet, which is widely seen in contemporary Japanese advertising and popular culture. Chinese and Japanese employ relatively large character sets (into the thousands), and both require special versions of the operating system, in addition, of course, to language-specific fonts to render the language typographically (for Japanese, see Lunde for a definitive guide).

Designers of multilingual Internet sites need to take into consideration a few important points regarding typography (Holzschlag, Weinman and Weinman, and Pirouz offer general design guides). The first consideration is the audience. Who is the prospective audience for the site, and what is the assumed operating system they will be using? In the case of Roman alphabet languages, this is no great problem, although Polish and certain other Eastern European languages do use different character sets from the standard ISO-8859-1 that is usually used for English. The main consideration here is for languages which employ other alphabets or nonalphabetic languages, such as Japanese or Chinese. If one is targeting, say a Saudi Arabian audience or a Japanese audience, it is probably safe to assume that users in these countries will have the appropriate operating systems for viewing
their respective languages on the computer. Users who do not have the language-specific operating systems necessary will not, of course, be able to view these languages. They will see a hodgepodge of random Roman characters (which are in fact the various combinations used to generate the language-specific characters). Such an incomprehensible screen display is usually quite noneuphemistically termed “garbage.”

The rule of thumb here is that if you want to display any language in a way that it is readable to all operating systems, make a graphic of the text (Weinman is a comprehensive guide to Internet graphics). This is quite easy to do in any of the standard image editing programs, such as Adobe Photoshop or Paint Shop Pro. Of course, in order to make such graphics, the creator must have the requisite operating system and fonts on hand. Even if a user is not interested in some of the languages available on your multilingual site, she will not be confronted by incomprehensible gibberish in the form of screen garbage, which might very well cause some users to think that something is wrong with the site. They would, of course, not be inaccurate, since the key guideline is never to expose users to a language they do not wish to read unless it is in an initial menu which allows a choice of languages. We discuss this opening menu below in the section on navigation.

A further typographical nuance that lends unity to a multilingual web site is achieved by preserving some semblance of visual/stylistic unity across the various language pages. This can be problematical, but an easy way to get around glaring dissimilarities is to use graphics and type to advantage. If buttons or other graphics are part of your design, make them identical across the language pages and, where possible, use fonts which either resemble each other slylistically or express the same feeling typographically. As a simple example of this, I offer the language menus for my own home page, which in many respects functions as an academic résumé, with sections including biodata, publications, and summaries of courses taught (See Figure 1 on the following page).
Figure 1: English, Japanese, and German Versions of the Same Page
Navigation Considerations

Perhaps the most important consideration when designing multilingual web sites involves deciding how to guide users to the language that is appropriate for them. Needless to say, a knowledge of HTML is essential in any good web designer’s armamentarium of skills and techniques (Aronson, Castro, Musciano and Kennedy offer ample direction in this area). One of the key principles in this regard is that users should have a choice very early in their visit to the site about which language they will proceed in. For this reason, it is highly useful to have an opening page with clear choices available, usually in the form of a menu that permits an immediate language choice. On this menu, of course, all languages, especially those non-Roman writing systems, such as Japanese, Chinese, Korean, Hebrew, Russian, Arabic, and the various languages of the Indian subcontinent, should appear as graphics, since one can never know what the underlying operating system will be. In other words, buttons or other menu items should permit users to choose the language they wish to use.

In my own courses in computer literacy, I have had even the beginning students create simple bilingual home pages in English and Japanese. Since these students are English majors and by and large native speakers of Japanese, it only makes sense that they would have a home page in both languages. One of the key features of their home pages is the linkability of both pages to each other and to the index and main pages in both English and Japanese (Boletta). The Internet has the potential for being the most democratic medium in the history of communication, and it should be available to all. It is not the elite province of a technical priesthood of cognoscenti, as a very iconoclastic, irreverent, and entertaining book by a member of that very priesthood, Professor Philip Greenspun of the Massachusetts Institute of Technology makes clear.
The initial page which confronts the user when she visits your site should be as easy to navigate as possible. The first question the user of a multilingual site must answer, the primary decision she must make, is to decide which language to proceed in. A clear choice at this juncture is crucial. Conversely, a failure to provide clarity on this entry page could be disastrous.

Below (Figure 2) is a simple example which I created for my quadrilingual home page allowing users to choose whether to proceed in English, Japanese, German, or French. The individual buttons for each language are clickable and will take the reader to the main menu page of that language (the pages shown above in Figure 1). In designing the interface for this page, my goal was to communicate graphically (one would hope without embracing too many typographical stereotypes) a feeling for each language. While this is surely an idiosyncratic area, and I would hazard to lay down rules for anyone, part of the richness of a multilingual site can derive from taking advantage of just such characteristics which add variety within a context of overall consistency. Readers will have to judge the ultimate success of the endeavor. Whatever typefaces and graphic designs one might choose, however, some sort of initial language choice is desirable.

![Figure 2](image-url)

Opening Main Menu for a Multilingual Web Site
Conclusion

The suggestions offered in this short paper are largely practical and grow out of experience rather than springing forth from any theoretical soil. There is a vast wealth (some might even say surfeit) of books, manuals, magazines, and online tutorials to help anyone who is interested in HTML or in web page design. Some of these are mentioned in the References at the end of this article. However, there are very few materials in any language which deal with the specific problems that confront the designer of multilingual sites. I discovered this myself when I set about preparing such sites. Although this brief look at some of the considerations that confront web designers in the areas of typography and navigation only covers a limited territory, I would hope that planners and designers of sites for information or pleasure might find it useful. There are, of course, many other areas which we have not touched on, but perhaps this modest beginning will encourage others to continue the exploration we have begun here.
References


