

## **Website localization**

© Anthony Pym 2010

Revised January 28, 2010.

Pre-print text written for the *Oxford Companion to Translation Studies*. Our thanks to José Ramón Biau, Yoonji Choi and Henry Wang for their input.

*Abstract: The localization of websites involves translating and adapting content to specific local markets. The process typically involves language technologies, and the degree of adaptation varies according to marketing criteria. Straight translation is thus only part of localization, which creates problems for the role of translators in this high-growth sector.*

### **1. Introduction**

Cross-cultural communication via websites, and of the elements that comprise the individual screens of each site, is one of the fastest growing areas opened up by the era of electronic communications. The basic standard known as Hypertext Markup Language (HTML) was first proposed in 1991 and has been followed by similar technical conventions, notably Personal Homepage Tools (PHP) from 1995 and Extensible Hypertext Markup Language (XHTML) from 2005, with each technical innovation bringing greater complexity. The translation and localization of websites has thus become a lucrative, dynamic and inter-professional field, often involving marketing, design, and software engineering, as well as linguistic processes. At the same time, the development of the Internet as an interactive medium is giving rise to a series of creative non-professional translation practices. In theory, the translation part of work on websites need not be any different from any other kind of translation, since texts can be extracted from site, rendered in accordance with the required communicative purposes, then reinserted into the site. The *localization* of websites, however, involves more complex processes, in addition to the normal constraints and goals of translation.

### **2. Localization in addition to translation**

The localization of a website differs from non-hypertext translation with respect to the identification of translatable elements, the tools needed to render them, their non-linearity, the way in which the translation process is prepared and coordinated, and the extent of the changes that may be introduced. Some of these features are shared with current technical translations, which increasingly also use hypertext technologies. The resulting complexity can be explained in terms of a basic translation process that is modified by a series of factors relatively specific to communication via multilingual websites.

### *2.1. Peripheral text*

A webpage includes textual units in its title (which appears in the title bar, at the very top of web browsers), in descriptors and keywords (which appear in search engines, and are important for the website to be located), in menus and hyperlinks (which lead to other webpages), in descriptors of images, sound files and audiovisual files (which may be used as pop-ups or as references for the location of files), and in the graphics, sound files and audiovisual files themselves. All those things may be rendered into different languages, in addition to the main body text. To access and translate the written strings (sequences of words), one can use a simple text editor and take care not to change the surrounding technical code (HTML or later), since most changes to the code will affect the appearance and functionality of the webpage. However, since this is risk-laden and visually difficult to do, various electronic tools are generally used to separate the codes from the 'translatables', the strings to be translated.

### *2.2. Website localization tools*

Most translation memory suites can be used to extract translatables from code, which is then protected, that is, blocked, and sometimes even hidden from the translator's eyes. The translator can then work on the isolated and segmented natural-language strings, as with any other use of translation memories. Some web-editing tools are helpful for the management of the translated pages in a site, which need to have all their hyperlinks coordinated with each other. In complex business situations, use is also made of content management systems that keep track of the short texts ('chunks' or 'information objects'). Global management systems can also be used, joining translation-memories and/or machine translation with the modification of elements for communication on websites, other web-based communication, or in print media. These tools enable a translation of a short text or update to be distributed automatically across the various communication media. There are also tools designed to calculate the cost of localizing a website, accounting for the complexity of the site as well as the automatic word count. Quality-assessment tools then check whether links still work, or whether all the content has been translated. Many of these tools, especially those incorporated in translation memory suites, create situations in which the translator cannot easily see the actual webpage in its visual format, which may result in a loss of communicative context (Biau 2005). The translator may simply not be able to grasp the nature and purpose of the text to be translated. On the other hand, much translation work done in this mode is not on whole sites or pages or such.

### *2.3. Work on updates*

Since electronic texts are easily modified, particularly as compared with print technologies, websites are frequently changed and updated. Much translation work thus does not start from the whole site but is limited to the modifications or updates. Moreover, translation of the website often starts before the source version is complete and running on a server. Translation memories can be efficient at locating the new translatables, although translators may also receive the isolated segments in a simple spreadsheet or word-processing format. As with much of the technical linguistic work in the localization industry, the translating then takes place on decontextualized segments. The translator does not have easy access to information about the communication act, but the nature of the work does not always require such access.

#### *2.4. One-to-many production*

Print-media translation mostly takes place after completion of the source text. This relation cannot be assumed in the case of website localization. A site might be developed first in one language and then localized in others, but once those first sites have been created, the translation process involves successive rounds of modifications and updates. A site might have one central language in which most initial changes are made, and those changes are then translated into the parts of the site that are in all the other languages. This means that the translations into the various languages occur more or less simultaneously, thus ensuring coherence and enabling synergies in marketing and promotion. Translation teams will ideally be working in parallel, often on the basis of ongoing contracts to handle the maintenance of a site over a period of time. In such a frame, translation operates not on a text, nor at the level of a finite project, but as part of an indefinite maintenance program. Ideally, such work situations mean that translators gain comprehensive knowledge of the website and its associated products, and this knowledge may be used to counter the decontextualized nature of the translatable strings.

#### *2.5. Internationalization*

Given the dynamics of one-to-many production, cost efficiency in website localization is sought by preparing the basic central version in such a way that many translation problems are actually avoided before they occur. The preparation, dubbed ‘internationalization’ in the field of software localization, means ensuring that the general site has as few culture-specific features as possible, since those are the elements most likely to cause problems downstream. The internationalized site is thus supposed to be neutral, functional, and constructed in such a way that the later localization teams can add elements (colours, images, references) that will make the site attractive to users in particular cultural locales. In effect, the apparent removal of culture-specific elements from the internationalized version tends to leave that site within a functionalist technical culture, sometimes identifiable with the company culture or client concerned. The ideology of internationalization nevertheless creates the illusion of a culture-less technical world.

#### *2.6. Localization*

Given the reduction of cultural elements in the internationalization process, the versions going into specific target languages may then have to add many features considered specific to the target culture. This is the process technically known as ‘localization’, even though the same term is misleadingly used for the entire production process. In the narrow sense, localization means adapting features to suit a particular ‘locale’, which is in turn understood as a market segment defined by criteria including language, currency, and perhaps educational level or income bracket, depending on the nature of the communication. The technical elements localized include the following, listed in most of the textbooks (for example, Esselink 2000):

- Date and time formats, as well as calendar settings.
- Currency formats and other monetary-related information (taxes)
- Number formats (decimal separator, thousand separator)

- Address formats, such as postal codes, provinces, states.
- Name formats. For instance, in Spanish-speaking countries there are two surnames.
- Telephone number formats
- Units of measure
- Paper sizes for print-outs
- Colour conventions: red means danger in European cultures, but good fortune in Chinese cultures
- Iconic conventions: A door might mean ‘exit’ in European cultures, but is likely to be misunderstood in others. A pestle-and-mortar signifies good cooking in Catalan culture (aioli is made that way), but tends to be associated with a pharmacy in English-language culture (perhaps the result of eating too much aioli). More generally, images of people are often changed to look like the projected users of the site.
- Sound files: Songs and music might have to be adapted. The Japanese are ostensibly embarrassed if a beep indicates they have made a mistake. Other audiences are sensitive to linguistic varieties. For example, the website of Bob the Builder, designed for pre-school boys, has its theme song in six different languages, including British and American English as two separate versions. These issues are generally dealt with on the basis of global geo-linguistic regions: the Australian Bob speaks British English; Canadian Bob speaks American English (or Canadian French); Spanish Bob seems to solve the problem by not speaking.
- Legal conventions: Copyright and personal data protection differ from country to country.
- Content: Adding locale-specific content (for instance, news on the opening of new offices in Paris only for the French version of the site).
- Connection speeds: For users in many countries with slow connection speeds, a highly sophisticated website will be difficult to use. Localization might thus require the removal of animations, sound files and high-resolution images.

Some of these adaptations would be included in print-media conceptions of translation; a few more might come under the heading of ‘translation’ in a highly functionalist approach; but not many notions of translation would include all the technical and marketing decisions that are encompassed by the concept of localization.

One should nevertheless be careful not to confuse this sense of ‘localization’ as adaptation with the nature of communication in the localization industry as a whole. As noted above, the automatic extraction of translatables, together with the nature of ongoing maintenance work, means that much translation is performed on decontextualized fragments, where quite literalist equivalence strategies become far more common than anything approaching adaptation. Thanks to the technologies, the localization industry commonly requires its translators to work at the level of quite restrictive phrase-to-phrase equivalence, with constant respect for pre-established glossaries.

### **3. Degrees of localization**

There are clearly different degrees to which a website can be localized. Schewe (2001) proposes a basic distinction between monolingual, bilingual and multilingual sites. He points out that the choice between these options depends on the language policy or

marketing strategy of the organization communicating through the site. Localization, however, is not limited to language issues alone. An English-language site may be localized for the different markets where English is used (Bob the Builder changes accent as he crosses the Atlantic), just as a multilingual site might choose to keep the same format and content across all its language versions, without any of the modifications mentioned above (major banks tend to adopt this option, in the interests of branding their stability and reliability). Degrees of localization thus concern the cultural implications of marketing strategies, as well as the existence of many languages.

Singh & Pereira (2005) recognize five degrees of localization: 'standardized' (one website for all countries), 'semi-localized' (one site gives information on many countries), 'localized' (a whole translated site for each country), 'highly localized' (translations plus country-specific adaptations), and 'culturally customized' (a new site completely immersed in the target culture). Only the 'localized' and 'highly localized' options involve any degree of translation in the traditional sense. In their survey of 307 US multinationals, Singh & Pereira found that most companies have one of the three degrees of 'localized' sites, with 17% using 'standardized' sites (i.e. no translation) and none rating as 'culturally customized' (i.e. complete re-generation).

There are many intermediary stages between these types. One might find, for example, that general information is translated but specific technical information, intended for specialized users, remains untranslated. That solution might be called a mode of 'standardization'.

It is also possible to adopt dynamically hybrid localization strategies. An example might be the basic Google homepage, which looks the same in all language versions and would thus seem to be a case of extreme standardization, maintaining the company image in the interests of branding. In Japan and Korea, however, that homepage is considered empty or incomplete, since users are accustomed to webpages that are crowded with many invitations to do exciting things. In those countries, Yahoo! or local search engines are more popular than Google. In this case, Google's answer has been not to change its standardization strategy for those markets, but to offer users the possibility to build their own personal Google homepages, adding in as many things as they want. All users can thus have their own crowded or uncrowded pages. The resulting localization is thus both 'standardization' and 'cultural customization' at the same time. Electronic communication means that different communication strategies need not be mutually exclusive.

Hybrid strategies are further enhanced by the use of hyperlinks. For example, a print-media translator might be faced with the dilemma of how to explain cultural realities. Faced with something like 'Australian-rules football', do you add a few phrases to note that it is not like soccer and not like American football? Or a footnote? Or just leave it as such? A website translator, however, could theoretically link the term to as much information as the user could possibly want, perhaps in one of the language versions of Wikipedia, in effect allowing the user to determine the extent of explanation. Some of the classical translation dilemmas might thus be resolved quite simply. In common practice, however, few translators are allowed responsibility for such things as adding hyperlinks.

Hybrid strategies move some communicative decision-making to the user's side, thus constituting one of the truly liberating and democratic aspects of electronic communication. However, the reverse tendency can be found in cases where companies seek to control and manage the range of mono-strategies. A logical consequence of the latter approach is the profiling of users. For example, the writer of these lines has a computer that uses Iberian Spanish as the locale for its operating system, so he is

automatically directed to Iberian Spanish-language versions of the major multinational websites, whether he wants to go there or not. Curiously, the same user, using the same computer, is automatically taken to the ‘Australia and NZ’ locale of Yahoo!, perhaps due to some dark secret held in a database somewhere. In a world of travelling users and complex cultural identities, such profiling can be annoying, and it can be quite difficult (although not impossible) to undo. Website systems that identify the user’s locale include, but are not limited to, identification of the locale of the operating system, regional origin of the IP number (where is the person connecting from?) and cookies, text files stored in your computer which remember your decision and reproduce it whenever you visit the same site again (which is probably how the ‘Australia and NZ’ got there). At the same time, profiling marks the extent to which the logic of nations and national languages still informs the era of electronic communication. A good deal of the tasks assigned to translators result from that blanket logic: all users in a particular country will need and want their information in the national language(s) of that country, despite the fact that the vast majority of language users in the world are polyglot. Profiling tends to force translation on many who do not always need it.

#### **4. Usability**

Studies on the reception of websites mostly refer to ‘users’ rather than ‘readers’, and there is indeed some doubt about the extent to which the latter term is appropriate. Some now dated research suggests that reading from a computer screen is about 25% slower than reading from paper (Nielson 1999: 101) and that this is only partly due to problems of screen resolution. Websites are felt to be a non-linear means of communication, where the receiver determines the rhythm of the communication act (as opposed to reading a book or watching a film in the cinema). Receivers tend to look over a webpage quickly, only focusing on isolated items of interest. In a study on the use of English-language newspaper websites, Nielson (2008) finds that highly educated users may read only 20-28% of the total information per visit, and that 17% stay on a single page for less than ten seconds.

What this means for translation is not always clear. If content is not going to be read with any significant attention, should it be translated with any degree of care? In the most ideal case, the translation process should involve some distinction between high-risk text requiring careful renditions (perhaps double reviewing and a user test), and low-risk text that is unlikely to be read (perhaps suitable for machine translation output with perfunctory reviewing). In practice, however, all strings tend to be treated equally, since the distribution of risk requires attention to context, and the electronic tools are designed precisely to separate text from context. Since translators mostly cannot see what the communication is about, they cannot assess where their best efforts should go.

Perhaps the most significant consequence of use patterns is that the design of the webpage is at least as important as its linguistic content. Nielson’s studies with eye-tracking (2006) suggest that English-language users start at the top-left and look across the page horizontally in one or two sweeps, and then skim down the page vertically, giving an F-shaped pattern. Key content should thus be located where the user is likely to look. More importantly, content should be arranged in such a way as to accommodate skimming, with many headers and with text in short paragraphs. A webpage cannot be designed or written in the same way as a printed page. It should be built for use, not just for reading.

For example, some online newspapers use content-management systems that move the position of a piece of news according to the number of time it is visited: the more visitors it has, the more visible it becomes, creating what may turn into a snowball effect.

The usability of websites is increasingly seen as the measure of their quality. Nielsen (1994: 26) suggests a framework for usability in five different categories:

- Learnability: How easy is it for users to accomplish basic tasks the first time they encounter the design?
- Efficiency: Once users have learned the design, how quickly can they perform tasks?
- Memorability: When users return to the design after a period of not using it, how easily can they re-establish proficiency?
- Errors: How many errors do users make, how severe are these errors, and how easily can they recover from the errors?
- Satisfaction: How pleasant is it to use the design?

Note that these criteria are not set patterns for all cultures; they are questions that should be answered for each particular locale. A design that has high usability in Korea may have very low usability in Mexico (cf. Choi 2008). One should also stress that the importance of visual elements and layout depends to a great degree on the overall purpose of the website. A fashion site like Dolce & Gabbana, for example, has rather low-quality translations as its linguistic content but the site itself is overwhelmingly based on visual design, to the extent that the linguistic text is small and hard to read—the site is meant to be looked at, not read. In other cases, such as newspaper sites, the linguistic content must be high-quality, since that is what the user is ultimately looking for (cf. Neilson and Loranger 2006).

Research on these issues sometimes makes reference to parameters that are presumed to typify entire cultures, much of it going back to Geert Hofstede's huge survey of IBM employees between 1967 and 1973, originally allowing numerical data to be synthesized for 40 countries (see Hofstede 1980, Hofstede and Hofstede 2005). Hofstede presents tables like the Uncertainty Avoidance Index or the Individualism Index, where different countries occupy different positions: the United States rates high on the individualism scale, for example, whereas Asian countries are at the bottom of that particular list. The logic of functionalist adaptation might suggest that a site going from American English into Korean should therefore have all individualistic elements removed, to give users what they are used to. Such shifts are rarely found, however (recall that Singh & Pereira found no 'culturally customized' sites in their survey). This may be because shifts on the macro level require too much effort of the translator, in tune with the 'law of interference' (Toury 1995), which predicts that translators will adapt the small units but not the big ones. However, it may also be because websites can be used to convey alterity, the new, the foreign as excitement or titillation, such that users actively go looking for ways to change their position on a Hofstede-type cultural table. Whatever the causes, the evidence suggests that the linguistically translated parts of websites tend not to display major cultural adaptations.

## 5. Who localizes, who translates?

There are several ways of relating the groups working on the different language versions of a website. At one extreme, everything may be controlled at the one central location, with the one general image or promotional campaign serving for all countries. This is the mode of organization that Lockwood (2000) dubs 'monarchist', well suited to a communication strategy that maintains the global image or brand. At the other, promotion and marketing may be undertaken by local experts in each case, in accordance with a model that Lockwood labels 'anarchist', at which point there need be no translational relationship between the various sites in different languages. Between monarchy and anarchy, there is a series of 'federalist' or 'subsidiary' approaches, where some content is generated centrally and translated for global use, other content is produced regionally and is translated for regional use (for example, the West European market), and still other content is produced locally, without recourse to translation. Thanks to such mixed strategies, most organizations attempt to preserve their global image while at the same time incorporating elements likely to enable regional synergies and appeal to local users.

The relations between these strategies may map onto different ways of organizing the teams engaged in the various localization processes.

### 5.1 *In-house localization*

Governmental and inter-governmental websites tend to localize by law or policy, usually into the languages considered official. Given the official status of the sites, there is still a preference to employ in-house translation teams, especially when security issues are involved. External language-service providers or certified freelancers may also be used for non-sensitive sites. For many bureaucratic services, sites of this nature carry heavy information loads with little attention to design. The translation process thus proceeds as if print-media were involved, or indeed as a part of print-media distribution. The result tends to be relatively uninviting sites with low usability levels. They are difficult to navigate around and first-time users struggle to locate the information they are looking for. Virtually any site of the European Commission could be taken as an example, in keeping with a political entity with a strong translation policy (information is available in official languages) but a weak communication policy (the information is poorly adapted to website usability).

Some good counter-examples can be found in sites of the Canadian government, for instance, and of various city sites where serious thought has been given to accessibility and different user needs. Such questions concern not just the translation strategies considered appropriate (literalist and exact in the case of information, adaptive in the case of motivating elements), but more importantly the design of the webpage, the ease with which hyperlinks can be located, and the information load of the natural-language strings. A traditional printed page such as this one, when put on a website, will not be considered accessible – its chances of being read are slight, no matter how accurate the translations.

### 5.2 *Localization companies and language-service vendors*

Given the special requirements of website communication, translation services are increasingly outsourced to specialized communication companies, which sometimes also provide website development services such as the treatment of graphic and audio

material or the adaptation of campaigns, in addition to straight translation services. The current trends, however, seem to be moving against the integration of translation into general communication services. Since various electronic tools make it fairly simple to extract natural-language strings, as noted above, those strings are sent to language-service vendors, often in a decontextualized format. The language-service providers then coordinate the translations. A global multilanguage vendor might take the contract for as many as twenty or so languages, then sell the work to a regional vendor for Asian languages, for example, who might in turn sub-contract to local single-language vendors, often located in countries where the translation rates are lowest. From there, the files are usually sent to freelancers, mostly in formats requiring free or cheap versions of translation-memory suites and probably accompanied by no special information on the nature of the website concerned. This structure means that the client who owns the website is paying up to three times what the end-translator is paid. Such is the hypertext mark-up, only some of which can be justified in terms of revisions, preparation and cleaning of translation memories, or integrated multimedia language services. In this respect, website translation is fundamentally no different from the practices of the localization industry as a whole.

The tendency to separate translation services from the general development and maintenance of websites further contributes to the perception of translation as a burdensome cost, a problem to be solved, rather than a creative process on a par with the other semiotic levels of a website.

The work-flow separation of translation from more general localization can also have repercussions for the training of translators. Few trainee translators acquire the multimedia and interactive skills most in demand at the more creative points of the industry, and those who do acquire such skills are unlikely to seek long-term employment doing nothing but decontextualized string-replacement. Good communicators might thus be forced out of the translation market or up in the decision-making scale, but then they stop translating.

### *5.3 Web-based machine translation*

Website communication also partakes of the vibrant democracy of Internet communication, where receivers can potentially be senders, users ostensibly drive growth, and technical skills are ideally more important than external qualification or personal financial reward. In most other modes of communication, translation tends to be an expensive strategy for cross-cultural translation (as opposed to various combinations of code-switching and language learning). On the Internet, however, translation can become a cheap, user-driven solution, where levels of quality are adjusted to the means available. This gives rise to a range of non-professional modes of translation, many of which are enhanced by technology.

Web-based machine translation has been a reality for some time. The transfer-based Systran system has been available as Babel Fish since 1997 (operated by AltaVista, now by Yahoo!), currently offering free automatic translations in nineteen pairs of languages. The translations produced by this system are far from perfect, but they do enable users to understand what is happening in a foreign website, thus enabling them to locate passages or pages that can be sent for human translation. Systran's publicity claims that their technology is used to translate more than four million webpages a day.

For most language pairs, better quality is given by data-based systems such as Google Translate, which currently caters for more than 50 languages. When users

operate through the free web-base translation-memory Google Translator Toolkit, released in June 2009, their modifications of the automatic output feed back into the database by default, thus improving future automatic output. This combination should eventually offer reasonably acceptable machine translations free online. This should in turn change the nature of professional translation services, with many of today's translators becoming tomorrow's technical writers (pre-editors) or revisers of machine translations (post-editors).

#### *5.4. Non-professional translations*

Since Internet communication can bring many people together around shared concerns, across languages and continents, the sociocultural groups thus formed can themselves take on many of the translation tasks. There are numerous terms for this phenomenon. In the development of open software it is sometimes called 'crowd sourcing' (as a variant of 'outsourcing'); the TAUS network prefers 'community translation', which generally refers to cases such as the social-networking site Facebook, where users perform translations and judge other users' translations (Facebook's translation tool currently operates in sixty-three languages); the expression 'citizen translations' colours the practice as collective political action; 'fan translation' is the term associated with translations done by fans of a specific product (video-games, cartoons, films, where subtitles thus become 'fan-sub'), increasingly distributed through websites.

All these alternatives to professional translation services, coupled with the availability of free online machine translation, greatly enhance the extent and creativity of cross-cultural communication. They potentially enable small, linguistically isolated entities to speak globally; they move the translator from silent mediator to active participant in a multilingual community. In doing so, non-professional translation practices violate most copyright agreements and codes of ethics, especially with respect to faithfulness to the source. Indeed, they question the profound political restrictiveness that has accompanied most ideologies of translation as a profession.

### **6. Directionality**

Websites were once predominantly in English, and website localization was thus assumed to go from English to the languages of the world's major markets. So dominant was this directionality that Schäler (2006) coined the term 'reverse localization' for movements in the opposite direction. The simultaneous translation of a multinational's website into twenty or forty languages is obviously quite a different operation from the English versions of each provincial bank, university or city in the world. What can be a sophisticated marketing operation in the first case tends to become a traditional textual translation in the second.

Limon (2008) compares websites from a small locale (the Slovene language) and those written directly in English for a global market. He finds that the global websites tend to be focused more on the user, incorporating the second person, attempting to cater to user needs and presenting themselves with a 'human face', whereas the sites from Slovenia tend to focus on the companies' achievements and products, using more technical registers and insisting on the modernity of their technology. This may reflect deep cultural differences, but it could also be due to the employment of professional marketing copywriters for the more global companies. That is, the cultural difference may go no deeper than the spreading professional culture of marketing.

The translation problems thus posed are nevertheless of some importance. Since web-based text genres are of recent, centralized creation, the translator working from a central language will tend to impose the genres on the peripheral target languages. Again in terms of Toury's proposed law of interference (1995: 275ff.), the translators will work at sentence-level but will tend not to alter the macrostructural features of the new genres. Similarly, when working from peripherally developed websites into English, the genre conventions tend not to be suited to electronic communication but are nevertheless reproduced in translation. The result of this kind of 'reverse localization' is a series of minor sites in major languages, mostly English, that fundamentally function as symbolic branding of the peripheral entities. The important thing is to 'have a website in English', since that in itself has a value within the peripheral culture.

Cursory analysis of the wider phenomenon nevertheless suggests that the world of websites is not so neatly divided between centre and periphery. Major languages such as Japanese, Korean and increasingly Chinese are part of web-based technological cultures that are developing their own new genres and styles of electronic interaction. Those websites are then localized into English and other major languages, sometimes with a surprising degree of adaptation and creativity, and often maintaining their own new genre conventions. We are aware of no solid research on this fast-moving phenomenon.

## **7. Catering to users**

A website may be Internet or Intranet. In the case of publicly available Internet sites, one of the frequently mentioned translation problems is the apparent lack of any specific target user, since the site may in principle be accessed by anyone at all. In some cases, translators into a major language such as English must simply suppose that the site is for 'anyone who knows English' (cf. Limon 2008), which would mean translating for a lowest common denominator. In practice, however, most sites conform to readily identifiable genres such as 'company promotion', 'product selection', 'games for kids', 'adolescent social networking', 'Internet pornography' or whatever, and the modes of likely use tend to be fairly evident. Further, website logs can give information about actual use of the site: what countries visitors are from, how long they stay on the site, and where they take their browser after leaving the site. The more interactive sites increasingly involve user feedback in the form of discussion lists or annotations (which may include anything from challenges on Wikipedia to video annotations on YouTube). In many cases, this empirical information enables fairly close tracking of usability, and can be used to modify the site accordingly. This is a further reason why website translations tend to be updates or re-translations rather than work from scratch. It is also a reason why translators should be prepared to find themselves in the middle of cross-lingual dialogues rather than simply reproducing a fixed source text.

## **8. Ethical issues in website localization**

When deciding how much of a website actually needs to be localized, companies are theoretically guided by calculations of return-on-investment. If a target language has enough speakers with enough economic or cultural capital to constitute a locale of interest, then localization will be invested in. The investment in high-quality or low-quality translations will then follow suit.

When markets grow beyond the central languages, simple business logic means that website communication will become increasingly multilingual, with strong growth

in the demand for translation services. The ethical problems facing website communication are thus not so much the once-dominant role of 'netspeak' English (Crystal 2001) or the imperialist imposition of centralized text genres (Limon 2008). The development of websites as places for interactive multimedia communication has promoted a more active multilingualism that should be seen as a sign of vibrant democratic interaction.

The ethical problems facing website translation these days have far more to do with democratic accessibility. This partly concerns disabled persons, where technical advances such as speech recognition, eye tracking and tactile screens will create a range of new modes of translation. However, accessibility also concerns the range of language varieties used and the design of highly usable interfaces: democratic participation means that translations must be able to speak to the old as well as the young, and that actual user-interaction, rather than linguistic accuracy, should be the measure of communicative success. Finally, accessibility has to do with opening the web to far more of the world's 6,000 or so languages, most of which do not have electronic forms. As websites increasingly use spoken and visual communication, more and more translations may be into voice files. Smaller languages may thus leap-frog the processes of graphic representation and finding spaces on Unicode, as new forms of translation bring participative democracy to a wider world.

## **9. Further reading and relevant sources**

Website communication is such a fast-moving phenomenon that this chapter will be out of date before it is printed. For research and updates, the best place to search is the web itself. The few textbooks that address website localization tend to be full of success stories (e.g. Esselink, Sprung, Singh & Pereira), plus a few horror stories to indicate why companies should spend money on professional services. The actual development of web-based translation practices would seem to be somewhere in between, in the more interesting social developments of non-professional and database-enhanced translation. The books that relate website communication to the theories of cross-cultural marketing (variously from Hofstede and De Mooij) tend to emphasize the need for cultural adaptation and the concerns of big business. The tendencies found in websites, however, have more to do with the creation of new cultural communities and with translation services that are non-professional, high-tech, low-cost and more creative than what has been said about them.

Biau Gil, José Ramón (2005), *Flying blind. Translation interfaces and non-verbal information in hypermedia texts*. Unpublished minor dissertation. Tarragona: Universitat Rovira i Virgili. [http://isg.urv.es/library/papers/flying\\_blind\\_biau.pdf](http://isg.urv.es/library/papers/flying_blind_biau.pdf). Accessed February 2009.

Cheng, Susan (2000), 'Globalizing an e-Commerce Web Site' in Robert Sprung (ed.), *Translating Into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam - Philadelphia: Benjamins: 29-42.

- Choi, Yoonji (2008), *Reading Patterns and their Impact on the Translation and Localization of Web Applications*. Unpublished minor dissertation. Tarragona: Universitat Rovira i Virgili.
- Corte, Noelia. 2002. 'Localización e internacionalización de sitios web'. *Revista Tradumatica* 1. Available online at: <http://www.fti.uab.es/tradumatica/revista/articles/ncorte/art.htm>. Visited February 2009.
- Crystal, David (2001), *Language and the Internet* (Cambridge: Cambridge: University Press).
- De Mooij, M (2004), *Consumer Behaviour and Culture: Consequences for Global Marketing and Advertising*. London: Sage.
- Esselink, Bert (2000), *A Practical Guide to Localisation* (Amsterdam, Philadelphia, John Benjamins).
- Gouadec, Daniel (2003), 'Le bagage spécifique du localiseur/localisateur. Le vrai "nouveau profil" requis'. *Meta* 48(4): 526-545.
- Hofstede, Geert. (1980). *Culture's Consequences: International differences in work related values*. Beverly Hill, CA, Sage.
- Hofstede, Geert and Gert Jan Hofstede (2005), *Cultures and Organizations. Software of the Mind*. 2nd edition. New York: McGraw-Hill.
- Limon, David (2008), 'Company websites, genre conventions, and the role of the translator', *Cultus* 1(1): 56-69.
- Mooij, Marieke de (2004), *Consumer Behavior and Culture. Consequences for Global Marketing and Advertising* (London: Sage Publications).
- Nielson, Jakob (1994), *Usability Engineering*. San Francisco: Morgan Kaufmann.
- Nielson, Jakob (1999), *Designing Web Usability*. Indianapolis: New Riders Publishing.
- Nielson, Jakob (2006), *F-Shaped Pattern for Reading Web Content*. Available online at: [http://www.useit.com/alertbox/reading\\_pattern.html](http://www.useit.com/alertbox/reading_pattern.html). Visited February 2009.
- Nielson, Jakob (2008), *How little do users read?* Available online at: <http://www.useit.com/alertbox/percent-text-read.html>. Visited February 2009.
- O'Hagan, Minako & David Ashworth (2002), *Translation-mediated Communication in a Digital World. Facing the Challenges of Globalization and Localization* (Clevedon: Multilingual Matters).
- O'Hagan, Minako & David Ashworth (2002), *Translation-Mediated Communication in a Digital World. Facing the Challenges of Globalization and Localization* (Clevedon: Multilingual Matters).
- Pym, Anthony (2004), *The Moving Text. Localization, Translation, and Distribution*

(Amsterdam, Philadelphia: John Benjamins).

Schäler, Reinhard (2006), 'Reverse Localization' in *MultiLingual Magazine* 17(3): 82.

Schewe, Theo (2001), 'Multilingual Communication in the Global Network Economy' in Eschenbach, J. & Theo Schewe (eds) *Über Grenzen gehen – Kommunikation zwischen Kulturen und Unternehmen* (Halden: Hogskolen i Ostfold). 195-209.

Singh, Nitish & Arun Pereira (2005), *The Culturally Customized Web Site: Customizing Web Sites for the Global Marketplace* (Burlington, US, Oxford UK: Elsevier Butterworth, Heinemann.)

Singh, Nitish & Hisako Matsuo (2004), 'Measuring cultural adaptation on the Web: a content analytic study of U.S. and Japanese Web sites'. *Journal of Business Research* 57, 8: 864-872.

Singh, Nitish, Olivier Furrer & Massimiliano Ostinelli (2004), 'To localize or to standardize on the Web: empirical evidence from Italy, India, Netherlands, Spain and Switzerland'. *Multinational Business Review* 12. 1. 69-87.

Singh, Nitish, Vikas Kumar & Daniel Baack (2005), 'Adaptation of cultural content: evidence from B2C e-commerce firms'. *European Journal of Marketing* 39(1/2): 71-86.

Sprung, Robert C. (ed.). (2000), *Translating Into Success. Cutting-edge strategies for going multilingual in a global age* (Amsterdam, Philadelphia: Benjamins).

Toury, Gideon (1995), *Descriptive Translation Studies and beyond* (Amsterdam, Philadelphia: Benjamins).

Tsikriktsis, Nikos (2002) 'Does culture influence Web site quality expectations? An empirical study'. In *Journal of Service Research* Vol. 5, no. 2. 101-112.

Vengadasamy, Ravichandran, Jaludin Azhar & Hamat Afendi (2004), 'Characteristics of Written Text in E-Commerce Websites' in *Internet Journal of e-Language Learning & Teaching* 1(2). 15-32.

Yunker, John (2003), *Beyond Borders. Web Globalization Strategies* (Indianapolis: New Riders Publishing).