

Alguns elementos para uma didática da tradução de conteúdos para a internet

Oscar Diaz Fouces¹

Resumo: O autor apresenta alguns conteúdos para “alimentar” a formação de tradutores/localizadores de páginas e *sites* da internet: língua-cultura, tecnologia e gestão de projetos.

Palavras-chave: localização, tradução, internet, *sites*, formação

Abstract: The author presents some content to feed the training of translators/localizers of webpages and websites: language-culture, technology and project management.

Keywords: localization; translation; Internet; sites; training.

“The medium is the message”
Marshall McLuhan

1. Introdução

A configuração acadêmica dos cursos relacionados com os Estudos sobre a Tradução experimentou, nos últimos anos do século XX e nestes primeiros anos do século XXI, uma extraordinária expansão. Podemos apontar dois responsáveis, com base em nossa experiência mais imediata. Diremos, por exemplo, que, na Espanha, antes contávamos com apenas dois centros que ofere-

¹ Doutor. Universidade de Vigo. fouces@uvigo.es

ciam cursos universitários de tradução do nível mais básico (*diplomaturas* de três anos) agora dispomos de mais de trinta centros de Ensino Superior com graduações (quatro anos), pós-graduações e mestrados (um ou dois anos) e doutoramentos nessa área. E diremos também que a base de dados BITRA da Universitat d'Alacant, compilou até 53 mil referências bibliográficas correspondentes a estudos no nosso campo – no âmbito internacional (http://aplicacionesua.cpd.ua.es/tra_int/usu/buscar.asp?idioma=en). São dados que demonstram e que podem ser um bom indício do grau de maturidade que a área vem atingindo.

Não é por acaso que essa expansão está relacionada com o processo de globalização a que tem sido submetida a economia nos últimos tempos e que trouxe consigo a necessidade premente de incorporar nos processos e nos produtos um autêntico “exército” de fornecedores de serviços linguísticos. Por outro lado, as circunstâncias tecnológicas com que esse contingente de profissionais deve lidar agora têm muito a ver com os novos suportes para a informação, em que os formatos digitais são, certamente, hegemônicos. A formação de tradutores não ficou à margem desta realidade, e passou a incorporar, não apenas nos cursos mais específicos de pós-graduação, como também nas próprias graduações, as habilitações correspondentes.

Do nosso ponto de vista, a tradução de conteúdos para a internet (páginas e *sites*) é um âmbito especialmente fascinante para trabalhar de modo simultâneo com aspetos formativos de natureza diferente, embora todos eles estejam relacionados com a prática profissional. Nos próximos parágrafos tentaremos apresentar alguns dos elementos basilares para “alimentar” com conteúdos um programa de formação de tradutores nesta área.

2. Língua(s) e cultura(s)

É habitual designarmos a prática profissional que envolve a tradução de conteúdos digitais – *sites* da internet, programas informáticos (*softwares*), *videogames*... – bem como a adaptação cultural deles com o nome de *Localização* (do inglês *Localization*, abreviado habitualmente como L10N). Para a *Localization Industry Standards Association* (LISA)², “*Localization involves taking a product and making it linguistically and culturally appropriate to the target locale (country/region and language) where it will be used and sold.*”

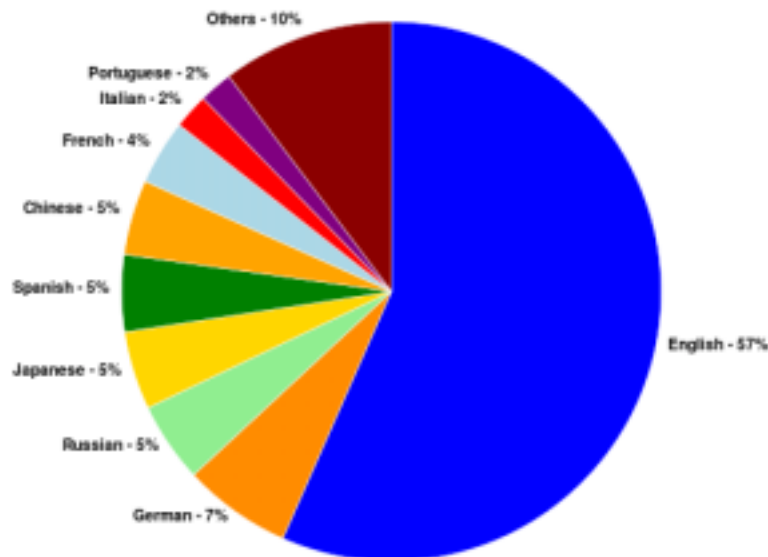
Mata Pastor (2005: 189) sintetiza com estas palavras os conteúdos que os processos de localização envolvem:

² Todas as citações de LISA foram retiradas de <http://www.aolti.com/helplocalization.asp>

(...) la traducción, entendida como trasvase interlingüístico e intercultural, constituye sólo un eslabón más de un proceso, el de la localización, extraordinariamente complejo, en el que se superponen a tareas de índole lingüística (traducción, gestión de terminología, elaboración de guías de estilo, etc.) otras propias de la programación y la ingeniería informática, el tratamiento y diseño gráfico o multimedia, la gestión de proyectos, el marketing o la venta, entre otras áreas.

Agora, diremos también que LISA definia a Internacionalização (sobre a qual depois voltaremos) deste modo: *“Internationalization is the process of generalizing a product so that it can handle multiple languages and cultural conventions without the need for re-design. Internationalization takes place at the level of program design and document development.”*

A dimensão mais óbvia da localização tem a ver com as línguas de trabalho envolvidas. Uma característica especialmente interessante da localização de conteúdos para a internet, que a distingue dos outros casos que antes citamos, é a dimensão multilíngue real que ela possui. Com efeito, enquanto a maior parte dos programas informáticos de propósito geral, bem como de *video games*, têm o inglês como língua de partida, a internet continua a ser linguisticamente plural, embora a presença do inglês seja também hegemônica. Segundo os dados para 2011 da *W3Techs.com* (http://w3techs.com/technologies/overview/content_language/all), a distribuição para 31 de maio de 2011 era esta:



Língua para os conteúdos da internet³

³ http://en.wikipedia.org/wiki/Languages_used_on_the_Internet#cite_note-UofCLBW-3

Por outro lado, como já indicamos em outro lugar (DIAZ FOUQUES, 2012; cf. também PIMIENTA; PRADO; BLANCO, 2009), será preciso levar em conta um fator essencial: o número de pessoas anglófonas com acesso à internet tem aumentado muito pouco nos últimos anos, mas o crescimento do volume das pessoas que falam chinês⁴, espanhol ou português, por exemplo, avançou significativamente, de modo que o número potencial de consumidores de conteúdos localizados não deixa de aumentar. A explicação justifica-se, nestes casos, pela incorporação progressiva ao *cibermundo* dos cidadãos da República Popular da China, bem como dos diferentes países de língua oficial espanhola e de língua oficial portuguesa da América e da África, à medida que as suas economias se desenvolvem. Portanto, a internet é, nestes momentos, um espaço multilíngue, e existem argumentos sólidos para prever que esta situação continuará nos próximos tempos, embora a hierarquia da demanda linguística seja essencialmente transitória. Em termos muito mais pragmáticos, podemos citar aqui os dados fornecidos por Michael Kingled, *Globalization Division Manager da Venturi's Globalization Division* (apud FOLARON 2006: 218), segundo os quais o negócio da localização de *sites* da internet passou de um mercado de 499 milhões de dólares USA em 2001, a 3,1 bilhões em 2007. É, com certeza, um volume mais do que interessante para sustentar um espaço profissional alician-te (note-se que o volume global para o mercado de tradução nesse último ano foi de 11,5 bilhões, segundo as mesmas fontes).

Que a atividade profissional que identificamos com a localização de *sites* envolve o trabalho com línguas é uma obviedade sobre a qual não vale a pena insistir. Ora, no texto de Mata que citamos anteriormente, o autor chama a atenção para esse fato, sobretudo, numa dimensão *intercultural*. Embora ela não deixe de estar presente, é claro, em todos os processos de mediação linguística, aqueles que têm a ver com diferentes dimensões semióticas resultam especialmente suscetíveis para esse *manuseamento intercultural*. A localização de páginas da internet envolve o trabalho com línguas, sim, mas também com as imagens fixas e em movimento, com os vídeos, com as cores, com a disposição dos elementos textuais para serem lidos (esquerda-direita ou direita-esquerda), com os tipos de letra, com a própria seleção de conteúdos e o modo como que eles são apresentados. Todos esses elementos, aparentemente “periféricos”, fazem parte da mensagem, informações, serviços ou produtos tangíveis que os *sites* da internet apresentam. No entanto, de que forma todos esses fatores “extralinguísticos” podem ser avaliados para depois ser tratados pelos localizadores profissionais?

⁴ É claro que a consideração de *uma* língua chinesa não deixa de ser uma abstração cientificamente insustentável. Segundo todos os indícios (<http://www.internetworldstats.com/stats17.htm>) as estatísticas apresentadas dizem respeito ao *mandarim*.

Nos últimos tempos, diversos investigadores já se debruçaram sobre esse tipo de assunto, de modo que contamos com um bom arsenal teórico-metodológico, que inclui pontos de vista mais genéricos (como WITTE, 2008, a propósito da dimensão cultural da tradução, *lato sensu*); outros diretamente relacionados com a criação e o *design* de *websites* otimizados para um público global (YUNKER, 2003) e que, portanto, combinam as categorias de adaptação cultural e aptidão para o uso (pelas palavras de BARBER; BADRE, [1998], a *culturabilidade* [*culturability*]); análises que partem de uma empresa-produto (WÜRTZ, 2005, sobre os *sites* da multinacional McDonalds); comparações da percepção das interfaces de usuário em diversos âmbitos culturais (CYR; TREVOR-SMITH, 2004); estudos tão específicos como o das traduções/localizações dos *sites* universitários na União Europeia que desenvolveu Fernández Costales (2010). Dadas as características deste trabalho, não vamos tentar compilar aqui uma bibliografia exaustiva, nem pretendemos mostrar todos os pormenores dessas investigações. Podemos sim apresentar aquilo que é, provavelmente, o núcleo teórico em que se alicerça boa parte delas, e que podemos sintetizar (embora a simplificação não faça justiça, claro é, à densidade da temática) nos modelos para a análise das diferenças culturais de Edward T. Hall e Geert Hofstede.

Devemos ao primeiro dos autores citados uma classificação de base antropológica, que parte das diversas formas de efetivarem-se as relações socio-profissionais. Hall (1976, HALL; HALL, 1990) salientava o valor nas interações comunicativas do *contexto*, entendido como a informação sobre um evento e que faz parte do significado dele. Na sua opinião, as diferentes culturas atribuem diferente valor à informação que as pessoas já possuem e àquela que está codificada linguisticamente. Assim, nas culturas de “contexto forte” (como no Japão, nos países árabes e nos países latinos), as pessoas já dispõem de muitas informações sobre a família, os colegas e os clientes, que é assumida como um *background* em que as informações explícitas — que podem ser mais “leves” — são processadas. No entanto, nas culturas de contexto fraco (Alemanha, países escandinavos, Estados Unidos da América) a transmissão das informações é muito explícita (ou, se quisermos utilizar a metáfora anterior, muito mais “pesada”).

Hall distinguia ainda dois tipos de comportamento diferentes, relativos à organização do tempo nas sociedades humanas, a que deu o nome de *policromismo* e *monocromismo*. Nas culturas *monocromáticas* as atividades são sequenciais (uma de cada vez), e essas circunstâncias fazem, por exemplo, com que exista um importante respeito pelos prazos estabelecidos nas agendas e muito pouca margem para para a distração externa aos caminhos operacionais que estão planificados (a cultura organizacional alemã seria o exemplo paradigmático). No caso das culturas *policromáticas*, contrariamente, existe

predisposição para simultanear as atividades, bem como uma grande abertura à flexibilidade (o exemplo seria neste caso, mais uma vez, o dos países latinos e árabes).

O segundo dos quadros teóricos a que nos referíamos é o que apresentou em 1991 Geert Hofstede⁵. Para este investigador social, que partiu de um trabalho de campo com operários da empresa IBM, as culturas podem ser classificadas a partir de quatro variáveis:

- **A Distância Hierárquica (DH) (Power Distance)**, que *“expresses the degree to which the less powerful members of a society accept and expect that power is distributed unequally. The fundamental issue here is how a society handles inequalities among people.”*
- **A dimensão Individualismo/Comunitarismo (IC) (Individualism versus Collectivism)**:
The high side of this dimension, called Individualism, can be defined as a preference for a loosely-knit social framework in which individuals are expected to take care of themselves and their immediate families only. Its opposite, Collectivism, represents a preference for a tightly-knit framework in society in which individuals can expect their relatives or members of a particular in-group to look after them in exchange for unquestioning loyalty.
- **A dimensão Masculino/Feminino (MF) (Masculinity versus Femininity)**:
The masculinity side of this dimension represents a preference in society for achievement, heroism, assertiveness and material reward for success. Society at large is more competitive. Its opposite, femininity, stands for a preference for cooperation, modesty, caring for the weak and quality of life. Society at large is more consensus-oriented.
- **A Aversão à Incerteza (AI) (Uncertainty Avoidance)**, que *“expresses the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity.”*

⁵ V. Hofstede (1991) e <<http://geert-hofstede.com/national-culture.html>>. Todas as citações que seguem são tiradas deste último site (acesso em: 1 jun. 2012).

Posteriormente, o modelo ganhou ainda uma quinta variável⁶,

- **A orientação curto prazo/médio prazo (CM) (Long-Term Orientation)**

The long-term orientation dimension can be interpreted as dealing with society's search for virtue. Societies with a short-term orientation generally have a strong concern with establishing the absolute Truth. They are normative in their thinking. They exhibit great respect for traditions, a relatively small propensity to save for the future, and a focus on achieving quick results. In societies with a long-term orientation, people believe that truth depends very much on situation, context and time. They show an ability to adapt traditions to changed conditions, a strong propensity to save and invest, thriftiness, and perseverance in achieving results.

Tudo bem, mas afinal como é que estes diferentes padrões culturais se concretizam na estrutura e na organização dos conteúdos dos *sites*? Um dos estudos que parte da base teórica que apresentamos, o trabalho de Sing & Pereira (2005) atribui às diferentes marcas culturais diversas características que ficam refletidas no *design* dos *websites*:

- **Coletivismo:** presença de bate-papos, *newsletter*, utilização de temas familiares, existência de programas de fidelização dos clientes e usuários.
- **Individualismo:** prioridade à privacidade dos utilizadores, preferência pelos temas relacionados com a independência, apelo à exclusividade do produto ou serviço, dá hipótese a uma personalização (da estrutura dos conteúdos, da interface com o usuário).

⁶ De fato, até chegou a ser introduzida recentemente uma sexta variável, da qual não trataremos aqui. Também no próprio *site* do autor lemos um resumo bastante esclarecedor: "A fifth Dimension was added in 1991 based on research by Michael Bond who conducted an additional international study among students with a survey instrument that was developed together with Chinese employees and managers. That Dimension, based on Confucian dynamism, is Long-Term Orientation (LTO) and was applied to 23 countries. In 2010, research by Michael Minkov allowed to extend the number of country scores for this dimension to 93, using recent World Values Survey data from representative samples of national populations. In the 2010 edition of *Cultures and organizations*, a sixth dimension has been added, based on Michael Minkov's analysis of the World Values Survey data for 93 countries. This new dimension is called Indulgence versus Restraint."

- **Aversão à incerteza:** apresentam-se mecanismos de apoio ao cliente, listagem de perguntas mais frequentes (FAQ), são utilizados temas tradicionais, salienta-se a segurança (p.ex. nas transações econômicas, ou na gestão dos dados pessoais).
- **Distância hierárquica:** pode mostrar um mapa da organização, fotografias dos diretores (CEOs), referências sobre padrões internacionais e controle de qualidade (QA).
- **Masculinidade:** disponibiliza jogos, fornece *tips&ticks*, utiliza temas essencialmente realistas, informa sobre a eficácia dos produtos.
- **Curto prazo:** realiza promoções agressivas, oferece descontos, cupons e garantias de recompra, em termos de linguagem utiliza superlativos.
- **Contexto forte:** estilo indireto e delicado, linguagem requintada, destaque para o elemento emocional e afetivo dos produtos, e utiliza temas lúdicos para promovê-los.
- **Contexto fraco:** promoções agressivas, descontos e cupons, são colocadas em evidência as vantagens dos produtos com respeito aos da concorrência (com comparações diretas), uso de superlativos e expressões hiperbólicas (“o número um”, “o máximo”...), destaque para a importância da empresa, garantias de recompra e de todas as condições que envolvem a aquisição.

A partir dos dados compilados no trabalho de Hofstede, pode estabelecer-se, ainda, um “catálogo” de países-cultura, como este que reproduzimos, que sintetiza os valores para quatro dos indicadores propostos:

	DH	IC	MF	AI
Brasil	69	38	49	76
China	80	15	55	40
Mundo Árabe	80	38	52	68
Espanha	57	51	42	86
Estados Unidos	40	91	62	46
Japão	54	46	95	92
Israel	13	54	47	81
Argentina	49	46	56	86

É claro que este tipo de generalizações/compartimentações devem ser tomadas sempre com todas as prevenções: não existem, por exemplo, culturas “masculinas puras”, “individualistas puras” ou “curtopracistas puras”. Contudo, as *tendências* observáveis são sempre um indicador muito útil. Por outro lado, falar em países-cultura envolve também, uma simplificação absurda. Num exemplo muito evidente, o dos Estados Unidos da América, a existência de duas comunidades linguístico-culturais e, portanto sociais, que apresentam perfis claramente diferentes, levou a IKEA, uma multinacional sueca do âmbito da mobília e decoração, a utilizar páginas principais diferentes no seu *site*, como se vê a seguir (todos os *screenshots* são de julho de 2011):



A primeira evidência será, claro, o uso de uma versão em língua diferente (espanhol para o coletivo hispânico e inglês para o anglo-saxônico). Vale a pena salientar também o emprego de algumas marcas textuais e pragmáticas diferentes. No texto inglês, o submenu da parte superior-esquerda da página abre-se a partir de indicações como “My account”, “My shopping cart” e “My list”, mas os “equivalentes” para o público hispânico não são simplesmente “Mi cuenta”, “Mi carro de compras” ou “Mi lista”, senão “Encuentra tu tienda IKEA”, “Únete a nuestra lista”, ou “Visita nuestro sitio móvil”. No texto inglês, o *protagonista* é o cliente. No texto espanhol, é o *site* que “dirige-se” a ele.

Por outro lado, o elemento afetivo mais óbvio, a imagem familiar, a que é dado destaque na página em espanhol, aparece “minimizada” na página em inglês (e ocupa o seu lugar uma cama vazia, com um cobertor cuja cor condiz com o quadro de texto à esquerda, que não é vermelha, mas sim cor-de-rosa). Certamente, as diferenças entre o perfil individualista/coletivista das duas comunidades terão alguma coisa a ver com isso. A dimensão forte/fraco resulta menos transparente numa sociedade como a americana. Em qualquer caso, será bastante óbvio que nem todo cidadão dos Estados Unidos da América responde a um mesmo padrão. Existem, aliás, alguns outros indicadores para além das classificações anteriores, que também o demonstram, e que entram nas dinâ-

micas mais puras do *marketing*. Assim, os elementos salientados na página “inglesa” têm a ver com o estilo próprio do cliente, e na página “hispanica” assume-se um perfil de cliente com família numerosa (“Comprar sin niños no tiene precio”) e com menor poder aquisitivo (a possibilidade de poupança, que no *site* em inglês aparece dissimulada, neste caso está claramente sobredimensionada: “Apúrate en llegar para ahorrar hasta un 20%.”)

Podemos ver um segundo exemplo para ilustrar, de um modo um pouco mais “extremo”, a dimensão cultural da localização. Neste caso, tirado da multinacional Kodak nos seus *sites* na China e no Brasil.



O recurso às imagens familiares é um lugar-comum nos *sites* orientados ao público chinês (como corresponde a uma cultura-alvo que é exemplo paradigmático da dimensão coletivista). Neste caso concreto, a estrutura geral da página (molduras e cores de fundo) segue o mesmo padrão para quase todos os *sites* da Kodak. A exceção mais evidente é a do *site* para o Brasil, em que o fundo não é preto, e sim branco, e portanto extremamente luminoso, deixando já de lado outras evidências como a composição multiétnica da imagem, ou a combinação de cores, a lembrar claramente a bandeira do país. Os dois *screenshots* foram tomados em julho de 2011. Nestes momentos (junho de 2012), os temas de todos os *sites* da Kodak já mudaram, mas o *background* do *site* do Brasil continua a ser significativamente branco (cf. www.kodak.com).

Permita-nos ainda um último exemplo, para ilustrar o grau de pormenor a que deve chegar a localização de um *site* para atingir de modo correto o público alvo visado.



A imagem da esquerda, na realidade, uma parte de uma pequena animação *Flash* (capturada em 08.06.2012) corresponde à campanha de Verão da rede de *shoppings* “El Corte Inglés”, de origem espanhola, que nos últimos anos abriu várias lojas em Portugal. Embora a imagem de fundo pareça idêntica, há um detalhe sutil que a torna diferente. Na União Europeia, as placas dos carros seguem um padrão semelhante: um retângulo branco com a numeração e um pequeno retângulo na parte esquerda de cor azul com o distintivo da UE (doze estrelas brancas em círculo) e a letra inicial do nome do país. No caso de Portugal, é acrescentado ainda um outro pequeno retângulo amarelo do lado direito, do mesmo tamanho que o azul da esquerda, com outras informações sobre a data de registro, que não existe no caso espanhol. Repare que os *designers* optaram por algumas estratégias de internacionalização interessantes: colocar o corpo da modelo cobrindo o lado esquerdo da matrícula (tapando um “eventual quadro amarelo”), utilizando um mapa de estradas da Espanha e de Portugal (na realidade “Espanña y Portugal”, o que só é possível observar se ampliarmos a imagem da versão espanhola). Também há uma alteração mínima para a localização: substituir o “E” de Espanha e por um “P” de Portugal. Já agora, as cores do cachecol que aparece pendurado do lado esquerdo do carro, por acaso os da bandeira da Espanha (vermelho-amarelo), perdem “protagonismo” com uma maior presença do verde na imagem “portuguesa” (as cores básicas da bandeira de Portugal são vermelho, verde e bastante amarelo).

3. Tecnologia, muita tecnologia

Este último exemplo permite intuir, de um modo bastante claro, que, para além dessa dimensão linguístico-cultural que acabamos de apresentar, a

localização de páginas e *sites* da Rede supõe ter de lidar com aspectos práticos, de base tecnológica. Vamos apresentar aqui, de um modo necessariamente introdutório (e, portanto, bastante “leve”) alguns desses aspectos. Partiremos, para isso, de duas perguntas fundamentais: quem é que decide como as páginas e os *sites* da internet “funcionam”? E de que matéria eles são “feitos”?

Muitas “regras do jogo” na Rede são estabelecidas pelo *World Wide Web Consortium* (W3C), organização internacional dedicada a desenvolver padrões de bom funcionamento (*standards*) para a internet. Dirigido por Timothy Berners-Lee, o “pai” do URL e das especificações HTTP, foi criado em 1994, a partir da *European Organization for Nuclear Research*, no *Massachusetts Institute of Technology* (com apoio da *Defense Advanced Research Projects Agency* e da Comissão Europeia). Conta com 369 membros (<http://www.w3.org/Consortium/Member/List>. Acesso em: 08 jun. 2012) e o seu protocolo de trabalho para elaborar Recomendações segue um rigoroso processo de cinco fases: *Working Draft*, *Last Call Working Draft*, *Call for implementation*, *Call for Review of a Proposed Recommendation*, *W3C Recommendation*.

Quanto aos materiais, diremos antes de tudo, que a internet que nós conhecemos se alicerça em três elementos fundamentais: um protocolo para implementar os saltos hipertextuais (o *HyperText Transfer Protocol*, HTTP); uma linguagem para representar os hipertextos, a estrutura deles, a formatação, os *hyperlinks* (o *HyperText Markup Language*, HTML); e diferentes aplicações-cliente para todas as plataformas, para acessarmos todas as informações armazenadas (dados, imagens, som, vídeo...), a partir de diversos protocolos (FTP, Gopher, o próprio HTTP, WAIS...). As versões 1 e 2 dos “tijolos” da Rede, o código HTML, foram desenvolvidas pela *Internet Engineering Task Force* (IETF), que depois se integraria ao W3C, inicialmente com um caráter estático (os utilizadores não podiam interagir com as páginas). As recomendações do HTML 4.0 (1997) introduziram as novidades que irão dar lugar ao DHTML (*Dynamic HTML*), que permite utilizar novas marcas e atributos que dão suporte às folhas de estilo em cascata (CSS), às linguagens de *script*, à possibilidade de incluir efeitos multimídia, etc. Os elementos HTML são tratados como objetos, de modo que o seu comportamento pode ser definido e programado, inclusive durante o acesso dos usuários às páginas (são objetos dinâmicos). Pela permissividade de HTML, juntamente com a rigidez das marcas, o W3C desenvolveu o XML (*Extensible Markup Language*). As primeiras recomendações são de 1997 e a primeira especificação é de 1998. Não pretende substituir o HTML, antes facilitar a interação das aplicações e os documentos. XML descreve a estrutura e o conteúdo dos documentos, e deixa a formatação para as folhas de estilo (em arquivos CSS). Como resultado das especificações XML 1.0 e HTML 4.1, o W3C lançou, em 2000, a recomendação XHTML (*Extensive HyperText Markup Language*) 1.0 (ou HTML extensível) que é, na realidade, uma reescrita de HTML como uma aplicação XML. De fato, XHTML não é mais do que HTML escrito com o rigor

sintático que exige XML. Atualmente (junho de 2012) existe um *working draft* do HTML 5, em elaboração pelo W3C (<http://www.w3.org/TR/html5/>).

Em termos mais práticos, e de modo muito sintético, diremos que um documento HTML não é mais do que um documento de texto simples, alojado num servidor com acesso à internet, que quando é interpretado por um *browser* exhibe as informações que pretendemos mostrar. Acrescentaremos ainda que os elementos que compõem um documento HTML representam-se mediante marcas (em inglês *tags*), com esta estrutura geral:

```
<nome_da_marca> TEXTO </nome_da_marca>
```

A estrutura básica de um documento HTML incluirá os seguintes elementos: HTML, HEAD e BODY.

- HTML: está representado pelas marcas <html> e </html>, que indicam, respetivamente, o início e o fim do documento. Estabelece que o documento está baseado nesta linguagem.
- HEAD: trata-se do cabeçalho do documento, que inclui informações ou metadados complementares (por exemplo, as palavras-chave que utilizam os motores de busca – como o Google – para indexar a nossa página). Representa-se pelas marcas <head> e </head>. Necessariamente (para estar bem formado) deve incluir o título do documento entre as marcas <title> e </title>.
- BODY: contém o corpo do documento. O seu conteúdo é o próprio documento (texto e marcas para formatar o texto), delimitado pelas marcas <body> e </body>.

A estrutura básica de um documento HTML fica então assim:

```
<html>
  <head>
    <title>Título do documento</title>
  </head>
  <body>
    Conteúdos do corpo do documento
  </body>
</html>
```


Um texto bem formatado deve incluir também o prólogo do documento, que indica a versão de HTML requerida, a partir de uma definição de tipo de documento (DTD) como, por exemplo, estas:

<!DOCTYPE PUBLIC "-//W3C/DTD HTML 4.01//EN"> – A versão mais rigorosa (*strict*). Estilos só em ficheiros de folha de estilo.

<!DOCTYPE PUBLIC "-//W3C/DTD HTML 4.01 Transitional//EN"> – Mais permissiva (*loose*) e mais utilizada, permite incluir elementos das versões anteriores

<!DOCTYPE PUBLIC "-//W3C/DTD HTML 4.01 Frameset//EN"> – Idem, com molduras.

Portanto, o código-fonte da nossa página poderia ficar assim:

```
<!DOCTYPE PUBLIC "-//W3C/DTD HTML 4.01 Transitional//EN">
<html>
  <head>
    <title>Título do documento</title>
  </head>
  <body>
    Conteúdos do corpo do documento
  </body>
</html>
```

Se escrevermos o texto anterior num editor simples (por exemplo o NotePad, no MSWindows, o Editor de Texto no MacOSX, ou Gedit em GNU/Linux), salvarmos com o nome “pagina.html” e depois abrirmos com um navegador, o resultado será semelhante a este:



Em termos muito básicos, localizar uma página da internet consistirá em substituir os trechos traduzíveis (quer dizer aqueles que ficam entre as marcas HTML) da língua origem pelos equivalentes numa (ou em várias) língua alvo. Assim, uma versão em espanhol do documento anterior ficaria deste modo:

```
<!DOCTYPE PUBLIC "-//W3C/DTD HTML 4.01 Transitional//EN">
<html>
  <head>
    <title>Título del documento</title>
  </head>
  <body>
    Contenidos del cuerpo del documento
  </body>
</html>
```

É claro que não existem muitas páginas reais com uma estrutura tão simples. O cabeçalho da página inicial da Universidade Federal do Rio de Janeiro em 11.06.2012 é, por exemplo, este que segue (incluímos a numeração para facilitar alguns comentários posteriores). Já agora, se o leitor ainda não souber, esta será uma boa altura para lembrar que podemos observar o código fonte de qualquer *site* simplesmente com o nosso *browser* (clicando sobre a opção Exibir código-fonte do botão Exibir da Barra de menus ou do menu que aparece depois de apertar o botão direito do *mouse* em cima da página que estamos navegando).

1. <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
2. <html xmlns="http://www.w3.org/1999/xhtml">
3. <head>
4. <title>Universidade Federal do Rio de Janeiro</title>
5. <meta http-equiv="content-type" content="text/html; charset=UTF-8">
6. <!-- SEO -->
7. <meta name="description" content="Portal da instituição UFRJ, Informações institucionais, Acesso, Graduação, Pós Graduação, Rio de Janeiro, RJ.">
8. <meta Name="robots" content="index, follow">
9. <meta name="keywords" content="UFRJ, Portal, Graduacao, Pos, Pesquisa, Jornal, Olhar, Virtual, Vital, Webtv, Plano, Diretor, Imagem,">
10. <meta name="alexaVerifyID" content="ci57ndzOmKTgwiugChKwRoUEUPs" />
11. <meta name="Author" CONTENT="Thiago Caldeira de Lima, thiagolima@reitoria.ufrj.br">
12. <!-- /SEO -->
13. <script language="JavaScript" src="js/dhtml.js" type="text/JavaScript"></script>

```

14.     <script language="JavaScript" src="js/dhtml2.js" type="text/JavaScript"></script>
15.     <script language="JavaScript" src="js/jquery.js" type="text/JavaScript"></script>
16.     <script language="javascript" type="text/javascript" src="libraries/jquery.js"></script>
17.     <script language="JavaScript" src="js/linkVeiculos.js" type="text/JavaScript"></script>
18.
19.     <meta name="google-site-verification"
content="NrtTMdFz26bIOOfcbKYsD6r7opsoTC4yGcB7ejaM8DI" />
20.     <meta name="google-site-verification"
content="PwjYJa6JlCS3IRGqQxXk83sBUwz5XoSqxL7yYDD6Sos" />
21.     <link href="css/estilo.css" rel="stylesheet" type="text/css" />
22.
23.     <!-- TIC esteve aqui! -->
24.     <link href="inc/barraGoverno/barraGoverno.css" rel="stylesheet" type="text/css" />
25.
26.     <script type="text/javascript">
27.
28.         var _gaq = _gaq || [];
29.         _gaq.push(['_setAccount', 'UA-1327593-1']);
30.         _gaq.push(['_trackPageview']);
31.
32.         (function() {
33.             var ga = document.createElement('script'); ga.type = 'text/javascript'; ga.async = true;
34.             ga.src = ('https:' == document.location.protocol ? 'https://ssl' : 'http://www') + '.google-
analytics.com/ga.js';
35.             var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(ga, s);
36.            })();
37.
38.         var _gaq = _gaq || [];
39.         _gaq.push(['_setAccount', 'UA-18182733-1']);
40.         _gaq.push(['_setDomainName', '.ufrj.br']);
41.         _gaq.push(['_trackPageview']);
42.
43.         (function() {
44.             var ga = document.createElement('script'); ga.type =
45. 'text/javascript'; ga.async = true;
46.             ga.src = ('https:' == document.location.protocol ? 'https://ssl' :
47. 'http://www') + '.google-analytics.com/ga.js';
48.             var s = document.getElementsByTagName('script')[0];
49.             s.parentNode.insertBefore(ga, s);
50.            })();
51.
52.     </script>
53. </head>

```

Ainda nem entramos no corpo do documento (onde estão os conteúdos “sérios” do *site*) e já nos deparamos com um monte de situações novas. Vamos comentar algumas delas, como exemplo:

- Na linha 1 vemos que a DTD faz pensar que estamos perante um *site* dinâmico, e não estático (utiliza XHTML), em que as informações provavelmente são geradas a partir de um CMS (*Content Manager System*). Para além de outros problemas específicos, esta situação nos leva a

intuir, por exemplo, que o *site* tem atualizações bastante frequentes. Muito provavelmente só uma parte do *site* seria eventualmente localizada (por exemplo, informações genéricas em inglês, para estudantes estrangeiros).

- Nas linhas 7, 8 e 9 aparecem diversas marcas “META”. A primeira e a terceira delas incluem elementos que serão utilizados pelos *robôs* (ou também *spiders*), programas que utilizam os motores de busca da internet, como o Google, o Yahoo ou o Bing para indexar as páginas. Em termos de localização é habitual manter o texto original e acrescentar-lhe a versão traduzida (e não substituí-lo por ela), dado que a empresa ou instituição proprietária do *site* pretenderá, com toda certeza, ser encontrada nas buscas realizadas nas duas (ou mais) línguas.
- Os elementos que aparecem nas linhas 6 e 23 são comentários. O *designer* que criou o *site* deixou-os aí para simplificar o trabalho posterior, bem como para “documentar” a sua própria atividade. Não são visíveis para o público (os *browsers* nunca os exibem).
- O código entre as linhas 26 e 52 é um pequeno *script* criado na linguagem Javascript.

Um *script* é um pequeno programa desenvolvido numa linguagem de *scripting* (ou *de extensão*), como o próprio Javascript, PHP ou (em alguma medida), Python, que realiza algumas funções específicas (validar formulários, abrir janelas, fazer contagens de visitas, até pequenos jogos...) As mensagens que exibem esses *scripts* também devem ser traduzidas, quando existem, de modo que as operações de tradução não fiquem limitadas, como vemos, a simples código HTML (embora este material represente sempre o volume mais significativo).

Com esses pequenos exemplos (e sem sairmos do cabeçalho!) já dá para perceber que trabalhar com o código de uma página apresenta bastante complexidade. Porém, todo o código, mesmo que inclua *scripts*, continua a ser um simples documento de texto. Isso quer dizer que, pelo menos na teoria, seria possível resolver uma encomenda de localização de páginas da internet com um simples editor de textos. Claro que, profissionalmente, essa estratégia seria inviável. Um *site* moderadamente complexo pode conter centenas de páginas, com materiais em pastas diferentes. Por outro lado, essas mesmas páginas vão conter, para além do código traduzível, matérias a que antes fazíamos referência, como imagens ou vídeos digitais. Felizmente, existem ferramentas criadas *ad hoc* para resolver esse tipo de tarefa, ao exemplo do Aquino WebBudget (www.aquino.net). O WebBudget é um *software* criado para traduzir *sites* da internet, isolando o texto das marcas. Permite importar a estrutura completa de um *site*, fazer contagens de palavras (também de imagens) para orçamentos

e faturas, lidar com diversas linguagens de *script* ou até externalizar o texto traduzível. Interessa-nos salientar que esse *software* trabalha a partir de *memórias de tradução*. Para os leitores que ainda não estiverem familiarizados com esse tipo de ferramenta, diremos que se trata de repositórios de segmentos equivalentes em língua origem e língua alvo que são recuperados quando, no texto que se está traduzindo, aparecem segmentos semelhantes (*fuzzy matching*) ou idênticos (*full matching*), de modo a otimizar a atuação dos profissionais. Podemos nos referir ainda a uma outra ferramenta para observar objetivos semelhantes, embora muito mais modesta, mas por um preço muito mais reduzido: o *Catscradle* (<http://www.stormdance.net/software/catscradle/overview.htm>). Embora sua funcionalidade seja muito mais limitada, dá para traduzir razoavelmente bem pequenos projetos em modo *wysigyg* (*what you see is what you get*, quer dizer, vendo os resultados diretamente como eles vão aparecer quando forem exibidos num *browser*). Na página do desenvolvedor existem algumas outras aplicações interessantes, como o *Caterpillar*, que permite isolar o texto traduzível das páginas.

Para além destes *softwares* específicos, também é possível utilizar para este tipo de tarefa algumas outras ferramentas de tradução assistida, como por exemplo o SDL Trados, DejaVuX, StarTransit ou, se se tratar unicamente de trabalhar com linguagens de marcas e *sites* de uma complexidade muito pequena, o gestor de memórias de tradução livre OmegaT⁷.

O tratamento dos materiais “extralinguísticos” (imagens estáticas, imagens dinâmicas, áudio e vídeo digital) envolve uma complexidade (e uma diversidade de casos) que aqui nem podemos começar a tratar. Existe, por exemplo uma especificação baseada em XML, chamada SVG (*Scalable Vector Graphics*) que permite um tratamento tão simples como qualquer linguagem de marcas. Assim, o código a seguir

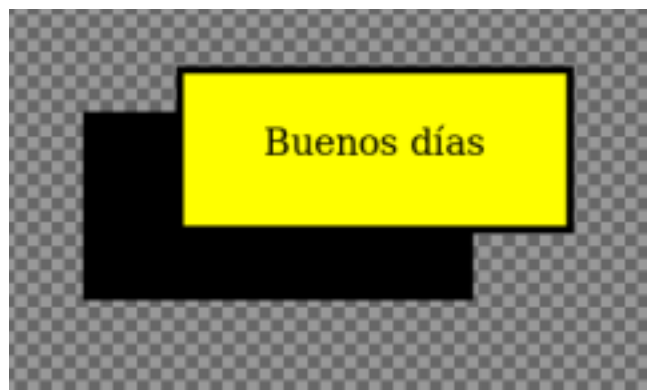
⁷ Há alguns anos um grupo de professores da Universidade de Vigo começamos a desenvolver um projeto de I+D para promover o uso do *software* livre entre os estudantes, professores e profissionais do campo da tradução, o projeto GETLT (<http://webs.uvigo.es/getlt>) Nesse âmbito, preparamos a *MinTrad*, uma distribuição GNU/Linux em forma de liveDVD (um DVD autoexecutável, que trabalha diretamente na RAM do computador), que inclui diversos *softwares* livres que podem ser utilizados na formação de tradutores. O endereço para fazer *download* é este: <ftp://ftp.uvigo.es/pub/asignaturas/GETLT/mintrad.iso>. Atualmente estamos trabalhando numa segunda versão atualizada. As pessoas mais interessadas em conhecer mais pormenores podem escrever para o autor deste trabalho.

```

<?xml version="1.0"?>
<svg width="640" height="480" xmlns="http://www.w3.org/2000/svg">
  <!-- Created with SVG-edit - http://svg-edit.googlecode.com/ -->
  <g>
    <title>Layer 1</title>
    <rect id="svg_4" height="119" width="251" y="178" x="135" stroke-width="5"
stroke="#000000" fill="#000000"/>
    <rect id="svg_5" height="105" width="256" y="148" x="196" stroke-width="5"
stroke="#000000" fill="#ffff00"/>
    <text xml:space="preserve" text-anchor="middle" font-family="serif" font-size="24"
id="svg_6" y="204" x="323" stroke-width="0" stroke="#000000" fill="#000000">Buenos días</
text>
  </g>
</svg>

```

consegue gerar esta imagem, gravada num arquivo de texto simples com a extensão SVG, que pode ser interpretado por um *browser*:



Para localizar esta imagem (em termos estritamente linguísticos, é claro) seria suficiente alterarmos no código o texto que nós destacamos em negrito (e escrever “Bom dia”, por exemplo). Um problema diferente seria tratar das formas, das combinações de cores, etc., mas isto também não representaria um problema importante. Existem bastantes *softwares* que conseguem lidar bem com o formato SVG. Nós queremos salientar aqui os extraordinários Gimp e o Inkscape, ambos livres. Para este exemplo utilizamos um editor SVG online, o SVG-Edit (<http://code.google.com/p/svg-edit/>).

Infelizmente, a situação anterior é pouco frequente. A rede está cheia de imagens em outros formatos, como JPG, TIFF, PNG, GIF (às vezes também animados), além de diversos formatos de vídeo (AVI, MOV...), entre os quais resul-

ta especialmente importante o *Macromedia Flash*. Em geral, a localização de imagens resulta relativamente simples quando o *designer* trabalhou com uma boa estratégia de localização, tendo, por exemplo, criado uma camada “transparente” com os textos, de modo a fazer com que seja possível isolá-los facilmente para tratar deles. Já na área das animações, conhecemos apenas um *software* específico para localizar aquelas que foram criadas em Flash, o Avral Tramigo (infelizmente desaparecido, embora ainda possa ser localizada alguma versão antiga na internet). Os leitores interessados nesse assunto vão encontrar um muito bom compêndio de informações nos trabalhos de Mata (2009a e 2009b).

À vista dos nossos comentários anteriores, resultará evidente que a localização de grandes *sites* não é habitualmente o resultado do trabalho de uma única pessoa. Quase sempre, a tradução/localização deste tipo de produtos envolve a participação de vários profissionais, para tratarem do texto, das imagens, dos vídeos. A localização de *sites* efetiva-se a partir de *projetos*, que devem ser geridos de um modo eficiente para levar até ao fim as solicitações do cliente.

4. Um bocado de gestão de projetos

Em termos práticos, embora muito sintéticos, o fluxo de trabalho nos processos de localização de *sites* da internet organiza-se de um modo semelhante a este que a seguir descrevemos. É habitual designarmos com o nome de gestor de projetos (*project manager*) a pessoa responsável por este tipo de tarefas:

1. Recepção do projeto.
2. Isolamento dos diferentes elementos (texto, imagens localizáveis...)
3. Geração de um orçamento, a partir de uma contagem prévia dos diferentes elementos e a cotação econômica deles. Já agora, resultará evidente que um orçamento para trabalhar com imagens fixas ou em movimento, por exemplo, não poderá responder aos mesmos padrões de contagem que a tradução de texto simples. Localizar uma imagem como acontecia no exemplo anterior (do “El Corte Inglés”) envolve um trabalho que ultrapassa claramente a simples substituição de uma letra por uma outra. Por outro lado, localizar texto numa imagem SVG bem internacionalizada não criará muitos obstáculos. É por isso que resulta difícil reduzir todos os casos a um único padrão de atuação. Haverá algumas ocasiões em que será absolutamente legítimo cobrar por ho-

ras de trabalho, por exemplo, e alguns outros em que a simplicidade do projeto permitirá uma redução de custos. É claro que neste tipo de decisão importam também outros fatores, como o nosso desejo de fidelizarmos os clientes, aplicando preços mais reduzidos (ou não).

4. Depois de ser aprovado o orçamento, distribuição entre os membros da equipe dos respectivos trabalhos e prazos dos mesmos. Eventualmente, nos casos de agências de tradução que não dispõem de pessoal suficiente para atender todas as solicitações de um projeto (quer seja pelo seu volume, quer pela especialização – por exemplo, pela existência de material gráfico), externalização (de parte) das tarefas. A distribuição do trabalho envolve, via de regra, algum tratamento prévio do material original: para além de uma “simples” extração de cadeias de texto, imagens, etc. Não raro os gestores devem, por exemplo, obter ou gerar guias de estilo ou glossários terminológicos para que os tradutores utilizem, ou transformar o material para formatos *standard* (v. embaixo).
5. Acompanhamento do projeto.
6. Compilação do material traduzido/localizado e controle de qualidade.
7. Entrega e acompanhamento de supervisão do cliente.
8. Faturação.
9. Compilação do material traduzido/localizado suscetível de reaproveitamento futuro e gestão de clientes (arquivo documental de dados, para eventuais projetos futuros). Gestão das memórias de tradução geradas, das bases de dados terminológicas, etc.

Esta última ação (o tratamento do material que resulta do processo, para o eventual reaproveitamento dele) pressupõe o uso por parte dos participantes do projeto de ferramentas de tradução assistida por computador, às que já antes nos referíamos. Contudo, o que acontece quando nem todos os participantes utilizam a mesma ferramenta? Essa pergunta está na base dos intuitos de diferentes agentes da indústria da localização, para gerar padrões de trabalho, *standards*, nomeadamente para os formatos de intercâmbio das memórias de tradução. Entre 1990 e 2011 a referência no setor foi a *Localization Industry Standards Association* (LISA), uma organização não governamental no campo da globalização e das indústrias associadas, cujo objetivo foi manter, desenvolver e certificar as normas relacionadas com a localização. Teve mais de quinhentos membros. Entre eles estavam grandes empresas na área de Tecnologia da Informação e Comunicações, IBM, HP, Nokia, Adobe, Novell, SDL, entre outras. Em fevereiro de 2011, a LISA entrou em falência e encerrou as atividades,

libertando os padrões desenvolvidos sob uma licença *Creative Commons*. Hoje, eles estão aos cuidados da GALA (*Globalization & Localization Association*, <http://www.gala-global.org/lisa-oscar-standards>). A LISA mantinha um boletim eletrônico, *The Globalization Insider*, que ainda é acessível em <http://www.lisa.org/globalizationinsider>. O *Open Standards for Container/Content Allowing Re-use* (OSCAR), um dos grupos de trabalho da LISA, tratou da criação de padrões livres e abertos para as indústrias da tradução e localização. Entre outros, desenvolveu as seguintes normas:

- TMX (Translation Memory eXchange) é um padrão aberto e independente, baseado no XML, criado em 1983 pelo OSCAR para facilitar o intercâmbio de memórias de tradução entre os diferentes fornecedores de serviços linguísticos e entre as diferentes ferramentas de tradução assistida.
- SRX (Segmentation Rules eXchange) é um segundo padrão, que visa estabelecer critérios para a segmentação dos trechos de que se “alimentam” as memórias de tradução. A versão 2.0 do SRX foi adotada como *standard* pelo OSCAR em 2008.
- O Term Base eXchange (TBX) é um formato para o intercâmbio de dados terminológicos, baseado em XML e gerado pela LISA em 2002, revisado e publicado como a norma ISO 30042:2008
- *XML Localization Interchange File Format* (XLIFF) é um padrão baseado em XML, para armazenagem dos dados linguísticos no processo de localização. XLIFF distingue as cadeias de texto traduzível do formato delas e do lugar que elas ocupam no texto original. Este segundo tipo de informação fica reservado num *skeleton*, de modo que depois seja possível reconstruir o documento alvo com a mesma estrutura do documento original. Algumas ferramentas de tradução assistida por computador como o *SwordFish* são especialmente aptas para trabalhar com o XLIFF.

Os *project managers* devem possuir um excelente conhecimento deste tipo de tecnologias que aqui apenas citamos. Também, claro, das técnicas de gestão de projetos num sentido mais alargado. Existem propostas interessantes de aplicação da norma de qualidade ISO 9001 na indústria da localização (v. DUNNE, 2006), sobre as quais não pudemos nos deter aqui, mas que apontam já num caminho certo para o futuro. Algumas dessas técnicas dizem respeito à gestão do tempo e dos recursos, materiais e humanos. Alguns pacotes informáticos, como o StarTransit NXT, dispõem de funcionalidades próprias dos gestores de projetos. E até existem *softwares* específicos nessa área, como o *Transla-*

tion Office 3000 (http://www.translation3000.com/Translators_Software/Accounting_Software.html) ou o TOM (*Translator's Office Manager*, <http://www.jovo-soft.de/>), que permitem manter também um controle eficaz de clientes e projetos. É claro que a própria decisão da escolha do *software* que deverá ser utilizado, bem como a gestão da comunicação, a gestão do risco, ou até o modelo global para o processo, fazem parte de um novo perfil profissional complexo, o do *project manager* das indústrias da localização. Numa monografia recente, cuja leitura recomendamos, a editada por Dunne & Dunne (2011), os leitores poderão encontrar bastantes referências e informações detalhadas sobre estes assuntos que aqui apenas apontamos.

5. E algumas sugestões didáticas

Os materiais que acabamos de apresentar podem lançar alguns alicerces para a criação de cursos de formação de localizadores de páginas e *sites* da internet. As três áreas que nós consideramos (língua-cultura, tecnologia e gestão) podem ser expressas em termos de diferentes competências que devem ser atingidas, os *outcomes* pretendidos para os processos formativos, como sugere Folaron (2006: 213-216). Seguimos a proposta desta autora para uma síntese final:

Competência 1: Gestão. Inclui, entre outros itens,

- a compreensão dos processos GILT (globalização, internacionalização, localização, tradução);
- a compreensão dos elementos e fases dos projetos de localização e de seus diferentes tipos;
- a capacidade de valorizar o grau de “localizabilidade” de um projeto;
- a capacidade de avaliar os resultados em termos de conteúdos e de funcionalidades;
- a capacidade de analisar, avaliar e priorizar os diferentes níveis de informação;
- a identificação das diferentes tarefas e dos problemas potenciais.

Competência 2: Tecnologia. Diz respeito

- à compreensão do conceito de “dado”, ao modo em que são criados, estruturados, organizados, armazenados e recuperados, bem como aos agrupamentos deles (em bases de dados);

- ao conceito de “documento” e o de “conteúdo”;
- às tecnologias utilizadas para criar conteúdos originais e às técnicas utilizadas para isolar os conteúdos dos elementos não localizáveis;
- à gestão das estruturas e aos fluxos de informação; às tecnologias que permitem analisar e avaliar os conteúdos originais;
- às tecnologias alternativas para suprir funcionalidades inexistentes nas ferramentas de tradução assistida disponíveis;
- às dinâmicas da interação homem-máquina e à ergonomia;
- às tecnologias utilizadas para localizar diferentes conteúdos para diferentes dispositivos;
- à criação e ao manuseamento de conteúdos baseados em *standards*, bem como ao trabalho com diferentes codificações;
- à capacidade de tomar em consideração variáveis como os formatos de data e hora, moedas, convenções locais e qualquer tipo de marca cultural, tanto do ponto de vista do trabalho prévio de internacionalização como da tradução;
- às tecnologias utilizadas na criação de *sites* da internet, em particular;
- às tecnologias utilizadas na tradução assistida, em geral.

Competência 3: Língua-Cultura. Tem a ver com

- os conhecimentos relativos à história da localização em termos do desenvolvimento da programação, bem como das especificidades do código para cada família linguística;
- a visão das línguas em termos de cultura e o contato com diversos grupos etnolinguísticos;
- a compreensão das dinâmicas da globalização, dos fluxos de informação que ela traz consigo e dos papéis que correspondem às diferentes instituições;
- as estruturas e as culturas baseadas em redes de comunicação (e a web 2.0);
- as linguagens *controladas* (linguagens criadas *ad hoc* para serem facilmente transferidas);
- a consultoria nas áreas da localização e da internacionalização;
- a implementação de diversos níveis de adaptação cultural.

Se calhar, uma das características mais óbvias da internet é o seu dinamismo. Um *site* criado em 1995 tem muito pouco em comum com um *site* criado em 2010. É previsível que as mudanças continuem, mesmo a curto prazo. Na prática, isso quer dizer que a formação de tradutores de conteúdos para a internet apresenta um perfil bastante diferente de outras especialidades mais suscetíveis de “estabilizarem-se”, em que as estratégias de trabalho mudam muito mais devagar, como pode ser o caso da tradução literária (mas também o da interpretação de conferência, apesar de possuir alguma base tecnológica). Os formadores deverão estar atentos a essa circunstância, como também devem estar sempre os profissionais. Portanto, e para sermos honestos, devemos advertir que os conteúdos deste texto (pelo menos uma parte deles), provavelmente vão deixar de ser úteis brevemente (se é que eles são úteis agora). Por outro lado, essa “instabilidade” do espaço profissional e, correlativamente, do espaço pedagógico, fazem com que esta especialidade resulte especialmente fascinante para algumas pessoas. Para o autor deste texto o é, com certeza.

Referências bibliográficas

BARBER, W.; BADRE, A. Culturability, the merging of culture and usability. In: PROCEEDINGS of the 4th Conference on Human Factors and the Web. NJ: Basking Ridge, 1998. Disponível em: <<http://research.microsoft.com/en-us/um/people/marycz/hfweb98/barber/>>. Acesso em: 1 jun. 2012.

CYR, D.; TREVOR-SMITH, H. Localization of Web Design: An Empirical Comparison of German, Japanese, and U.S. Website Characteristics. *Journal of the American Society for Information Science and Technology*, 2004, p. 1199-1298 Disponível em: <www.dianneecyr.com/docs/localization_of_webdesign.pdf>.

DIAZ FOUQUES, O. A localização de páginas da Internet na formação de tradutores. *Confluências. Revista de Tradução Científica e Técnica* 1:, 2004, p. 16-52

_____. Algumas considerações sobre a combinação linguística Português-Espanhol. In: _____. *Olhares & Miradas. Reflexiones sobre la traducción portugués-español y su didáctica*. Granada: Atrio, 2012, p. 119-135.

DUNNE, K. Putting the cart behind the horse. Rethinking localization quality management. In: _____. (ed.). *Perspectives on Localization*. Amsterdam/Philadelphia: John Benjamins, 2006, p. 95-117.

_____.; DUNNE, E.S. *Translation and Localization Project Management*. Amsterdam/Philadelphia: John Benjamins, 2011.

FERNÁNDEZ COSTALES, A. *Traducción, localización e internacionalización. El caso de las páginas web universitarias*. Tese de doutoramento: Universidad de Oviedo, 2010.

FOLARON, D. A discipline coming of age in the digital age. In: DUNNE, K. (ed.). *Perspectives on Localization*. Amsterdam/Philadelphia: John Benjamins, 2006, p. 195-219.

HALL, E.T. *Beyond Culture*. Garden City: Doubleday, 1976.

_____.; HALL, M.R. *Understanding Cultural Differences*. Yarmouth: Intercultural Press, 1990.

HOFSTEDE, G. *Cultures and Organizations: Software of the Mind*. London: McGraw-Hill, 1991.

MATA PASTOR, M. Localización y traducción de contenido web. In: REINEKE, D. (ed). *Traducción y localización. Mercado, gestión y tecnologías*. Las Palmas: Anroart, 2005, p. 187-252.

_____. Algunas pautas para el tratamiento de imágenes y contenido gráfico en proyectos de localización (I e II). *Entreculturas 1*, 2009. Disponible em: <<http://www.entreculturas.uma.es/n1pdf/articulo26.pdf>> e <<http://www.entreculturas.uma.es/n1pdf/articulo27.pdf>>.

PIMIENTA, D.; PRADO, D.; BLANCO, A. *Twelve years of measuring linguistic diversity in the Internet: balance and perspectives*. UNESCO Publications for the World Summit on the Information Society. Communication and Information Sector, 2009. Disponible em: <<http://unesdoc.unesco.org/images/0018/001870/187016e.pdf>>.

PYM, A. *Localization, Training, and the Threat of Fragmentation*. Intercultural Studies Group, Universitat Rovira i Virgili. Tarragona, Spain. 2006. Disponible em: <http://isg.urv.es/publicity/isg/publications/technology_2006/index.htm>

SINGH, N.; A. PEREIRA. *The Culturally Customized Web Site*. Oxford: Elsevier, 2005.

WITTE, H. *Traducción y percepción cultural*. Granada: Comares, 2008.

WÜRTZ, E. A Cross-Cultural Analysis of Web sites from High-Context Cultures and Low-Context Cultures. *Journal of Computer-Mediated Communication* 11(1), 2005. Disponible em: <<http://jcmc.indiana.edu/vol11/issue1/wuertz.html>>.

YUNKER, J. *Beyond Borders: Web globalization Strategies*. Indianapolis: New Riders, 2003.