

El Internet y los idiomas

EL INTERNET Y LOS IDIOMAS

[alrededor del año 2000]

MARIE LEBERT

NEF, Universidad de Toronto, 2009

Copyright © 2009 Marie Lebert. Todos derechos reservados.

ÍNDICE

Introducción

"Naciones de los idiomas" en línea

Hacia una web multilingüe

El inglés todavía preponderante

Codificación: del ASCII al Unicode

Primeros proyectos multilingües

Diccionarios de traducción en línea

Aprender los idiomas en la web

Los idiomas minoritarios

Localización e internacionalización

Traducción asistida por computador

Traducción automática

Cronología

Sitios web

INTRODUCCIÓN

Se dice a menudo que el internet elimina la barrera del tiempo, de las distancias y de las fronteras, pero ¿qué hacer con la barrera de los idiomas?

En 2000, la web es multilingüe, y la mitad de los usuarios de internet no es de habla inglesa, pero la barrera del idioma está lejos de haber desaparecido. Si bien todos los idiomas están presentes en la web, olvidamos a veces que muchos usuarios son monolingües o bilingües, y que incluso los políglotos no pueden conocer todos los idiomas. También es importante tener en cuenta cada lengua, y no sólo las lenguas dominantes. Quedan por construir puentes entre comunidades lingüísticas para favorecer el flujo de textos en otros idiomas, contando entre otras cosas con programas de traducción de mejor calidad.

"Naciones de los idiomas"

"Como el internet no tiene fronteras nacionales, los usuarios se organizan según otros criterios propios de este medio de comunicación. En términos de plurilingüismo, existen comunidades virtuales, por ejemplo lo que suelo llamar las 'Naciones de los idiomas', es decir todos los usuarios que comparten la misma lengua materna, cualquiera que sea su entorno geográfico. Así pues, la nación de habla hispana no abarca sólo a los usuarios de España y de Latinoamérica, sino también a todos los hispanohablantes que viven en los Estados Unidos y en otros lugares, en Marruecos por ejemplo." (Randy Hobler, consultor en marketing internet de productos y servicios de traducción, septiembre de 1998)

La "democracia lingüística"

"En un informe de la UNESCO a principios de los años 1950, la educación en lengua materna era considerada como un derecho fundamental para todos los niños en el mundo. En la Era de la Información, esto equivaldría a la oportunidad para ellos de navegar en la web en su lengua materna. Si el internet quiere convertirse realmente en la Red Global que se promueve, todos los usuarios deberían tener acceso a él sin el problema del idioma. Considerar el internet como exclusivo de la gente que, por accidente histórico, necesidad práctica o privilegio político, habla inglés, resulta injusto para la gente que no habla este idioma." (Brian King, director del Worldwide Language Institute, septiembre de 1998)

Un medio de comunicación para el mundo

"Es muy importante poder comunicar a través del internet en diferentes lenguas, es más bien obligatorio. Ya que la información la tenemos a nivel mundial, ¿por qué no podríamos tenerla en el idioma que hablamos o que deseamos? ¿Acaso no es contradictorio?" (María Victoria Marinetti, profesora de español en empresas y traductora, agosto de 1999)

Buenos softwares

"Cuando la calidad de los softwares sea suficiente para que la gente converse por escrito y verbalmente en la web en varios idiomas y en tiempo real, veremos un nuevo mundo abrirse ante nosotros. Científicos, políticos, empresarios y muchos otros grupos podrán comunicarse directamente entre ellos sin la intervención de mediadores o traductores." (Tim McKenna, escritor y filósofo, octubre de 2000)

En todos los idiomas

"La investigación sobre la traducción automática debería permitir una traducción automática a los idiomas deseados, pero ofreciendo aplicaciones para cada lengua y no sólo para las lenguas dominantes (por ejemplo la difusión de documentos en japonés si el emisor es de lengua japonesa, y la lectura en bretón si el receptor es de lengua bretona). Así pues queda mucho por hacer en el campo de la traducción automática y escrita de todas las lenguas." (Pierre-Noël Favennec, experto en la dirección científica de France Télécom R&D, febrero de 2001)

Las citas que aparecen en este libro son fragmentos de las Entrevistas del NEF <<http://www.etudes-francaises.net/entretiens/index.htm>>, salvo que se indique lo contrario. Muchas gracias a Anna Álvarez por su ayuda, y muchas gracias a todas las personas que han contestado preguntas sobre el multilingüismo en la web durante varios años y que están citadas aquí. La autora, cuya lengua materna es el francés, es responsable de los errores que podrían quedar en español.

Este libro también está disponible en francés y en inglés en el NEF, con textos diferentes <<http://www.etudes-francaises.net/entretiens/multi.htm>>.

Marie Lebert, investigadora y periodista, se interesa por las tecnologías para el libro, los medios de comunicación y los idiomas. Sus libros son publicados por el NEF (Net des Etudes Françaises / Net de los Estudios Franceses), Universidad de Toronto, y se pueden consultar libremente en el NEF <<http://www.etudes-francaises.net>>.

"NACIONES DE LOS IDIOMAS" EN LÍNEA

= [Cita]

Randy Hobler, consultor en marketing internet de productos y servicios de traducción, escribe en septiembre de 1998: "Como el internet no tiene fronteras nacionales, los usuarios se organizan según otros criterios propios de este medio de comunicación. En términos de plurilingüismo, existen comunidades virtuales,

por ejemplo lo que suelo llamar las 'Naciones de los idiomas', es a saber todos los usuarios que comparten la misma lengua materna, cualquiera que sea su entorno geográfico. Así pues la nación de habla hispana no abarca sólo a los usuarios de España y Latinoamérica, sino también a todos los hispanohablantes que viven en los Estados Unidos y en otros lugares, como Marruecos por ejemplo."

= [Texto]

Randy da el ejemplo de la comunidad de habla hispana repartida entre continentes, pero lo mismo vale para la francofonía, una comunidad de lengua francesa repartida en cinco continentes, o para el criollo, que se habla no sólo en el Caribe sino también en París, Montreal y Nueva York.

En sus inicios, el internet es casi al 100% en inglés, a consecuencia de haber empezado en los Estados Unidos como una red establecida en 1969 por el Pentágono, antes de extenderse a los organismos públicos y las universidades tras la creación del protocolo TCP/IP (Transmission Control Protocol/Internet Protocol) en 1974 por Vinton Cerf y Bob Kahn. Después de la invención de la web en 1990 por Tim Berners-Lee en el Centro Europeo para la Investigación Nuclear (CERN: Centre Européen pour la Recherche Nucléaire) en Ginebra (Suiza) y del lanzamiento en noviembre de 1993 del primer navegador Mosaic, el predecesor de Netscape, el internet se desarrolla rápidamente, primero en los Estados Unidos gracias a las inversiones muy importantes del gobierno, luego en los países de habla inglesa y finalmente en el mundo entero.

En 1998, el internet sigue dominado por el inglés al 80%, un porcentaje alto que se explica por tres factores: (a) el uso extensivo del inglés en el comercio internacional, (b) la creación de muchos sitios web por empresas e individuos que viven en los Estados Unidos, en Canadá o en el Reino Unido, (c) un gran número de usuarios en Norteamérica, gracias a computadoras más baratas que en otros países, y una conexión al internet barata con un paquete mensual.

En varios países europeos, el precio de la conexión al internet depende de la duración, con una tasa de día cara y una tasa de noche más barata. Los usuarios pasan mucho menos tiempo en el internet de lo que les gustaría, y navegan de noche para evitar facturas demasiado altas. A finales de 1998 y a principios de 1999, se inician huelgas en Francia, Italia y Alemania para exigir de los proveedores de servicios internet que reduzcan sus precios y ofrezcan paquetes mensuales. Los usuarios obtienen satisfacción en los siguientes meses.

En 1997, Babel, una iniciativa conjunta de Alis Technologies y de la Internet Society, realiza el primer estudio sobre la distribución de las lenguas en la web. En junio de 1997, la página "Palmarés de las Lenguas en la Web" da los porcentajes de 82,3% para el inglés, 4% para el alemán, 1,6% para el japonés, 1,5% para el francés, 1,1% para el español, 1,1% para el sueco y 1% para el italiano.

En un artículo publicado el 21 junio de 1998 por ZDNN (ZD Network News), la periodista Martha Stone escribe: "Este año, es probable que el número de nuevos sitios en otros idiomas que el inglés supere el número de nuevos sitios en inglés, y el cybermundo realmente se está convirtiendo en una red mundial.

(...) Según Global Reach [una empresa que promueve la localización de los sitios web], los grupos lingüísticos que están desarrollándose más rápido no son los grupos que hablan inglés: hay un aumento del 22,4% para los sitios web en español, del 12,3% para los sitios web en japonés, del 14% para los sitios web en alemán y del 10% para los sitios web en francés. Se estima que 55,7 millones de personas que no son de habla inglesa tienen acceso a la web. (...) Si bien solamente el 6% de la población mundial es de lengua materna inglesa (y el 16% de lengua materna española), el 80% de los sitios web está todavía en inglés." Según Global Reach, sólo el 15% de los 500 millones de europeos es de lengua materna inglesa, el 28% tiene un buen nivel en inglés, y el 32% consulta la web en inglés.

Brian King, director del World Wide Language Institute (WWLI), desarrolla el concepto de "democracia lingüística" en septiembre de 1998: "En un informe de la UNESCO a principios de los 1950, la educación en lengua materna era considerada como un derecho fundamental para todos los niños en el mundo. En la Era de la Información, esto equivaldría a la oportunidad para ellos de navegar en la web en su lengua materna. Si el internet quiere convertirse realmente en la Red Global que se promueve, todos los usuarios deberían tener acceso a él sin el problema del idioma. Considerar el internet como exclusivo de la gente que, por accidente histórico, necesidad práctica o privilegio político, habla inglés, es injusto para la gente que no habla este idioma."

Autor de "Chroniques de Cybérie" (Crónicas de Cyberia), una crónica semanal electrónica sobre las noticias del internet, el periodista quebequés Jean-Pierre Cloutier escribe en agosto de 1999: "Este verano se tomó un giro. Ahora más del 50% de los usuarios de la red viven fuera de los Estados Unidos. El año próximo, más del 50% de los usuarios no serán de habla inglesa. Hace solamente cinco años representaban un 5%. Estupendo, ¿verdad?"

En efecto el número de usuarios que no son de habla inglesa alcanza un 50% durante el verano de 2000. Según Global Reach, este porcentaje se eleva al 52,5% en el verano de 2001, el 57% en diciembre de 2001, el 59,8% en abril de 2002, el 64,4% en septiembre de 2003 (de los cuales un 34,9% de europeos de habla no inglesa y un 29,4% de asiáticos) y el 64,2% en marzo de 2004 (de los cuales un 37,9% de europeos de habla no inglesa y un 33% de asiáticos).

Muchas comunidades son bilingües en la vida cotidiana, por ejemplo Ginebra (Suiza) con el francés y el alemán, Toronto (Canadá) con el inglés y el francés, o San Francisco (Estados Unidos) con el inglés y el español, por citar sólo tres ejemplos. El caso extremo es el de la Unión Europea con sus 11 lenguas oficiales en 2003 y sus 24 idiomas oficiales en 2007 después de su expansión hacia Europa del Este.

HACIA UNA WEB MULTILINGÜE

= [Cita]

Maria Victoria Marinetti es profesora de español en empresas y traductora. Escribe en agosto de 1999: "Es muy importante poder comunicarse a través del internet en diferentes lenguas, es más bien obligatorio. Ya que la información la tenemos a nivel mundial, ¿por qué no podríamos tenerla en el idioma que hablamos o que deseamos? ¿Acaso no es contradictorio?"

= [Texto]

En California, dos estudiantes de la Universidad de Stanford, Jerry Lang y David Filo, lanzan en enero de 1994 Yahoo!, un directorio web en el cual hacen un inventario de los sitios web y los clasifican por temas. El directorio propone una clasificación más precisa que otros buscadores como AltaVista que sólo proponen listas completamente automatizadas. Tres años después, Yahoo! se divide en 63 grandes categorías, con una interfaz en varios idiomas: inglés, alemán, coreano, francés, japonés, noruego y sueco. Yahoo! trabaja en colaboración con AltaVista. Cuando una búsqueda no produce resultados en Yahoo!, es dirigida automáticamente hacia AltaVista, y viceversa.

En diciembre de 1997, AltaVista es el primer buscador en lanzar un servicio gratuito de traducción automática del inglés a cinco otros idiomas (alemán, español, francés, italiano y portugués) y viceversa, un sistema en el cual la página original y la traducción aparecen una al lado de la otra en la pantalla. AltaVista Translation, también llamado Babel Fish, es la obra de SYSTRAN, una empresa pionera en programas de traducción automática. Babel Fish es alimentado por diccionarios multilingües de 2,5 millones de términos. Aunque tenga sus límites, con una capacidad de traducción que no puede exceder tres páginas y traducciones muy aproximadas, este servicio se hace de inmediato muy popular ante los doce millones de usuarios de AltaVista, entre los cuales se observa un creciente número de usuarios que no son de habla inglesa. Babel Fish también abre la vía a otros servicios del mismo tipo - desarrollados por ejemplo por Alis Technologies, Lernout & Hauspie, Globalink y Softissimo - y contribuye al plurilingüismo de la red.

Otra iniciativa es OneLook Dictionaries (Diccionarios en una Mirada), un sitio web lanzado en abril de 1996 por el informático Robert Ware. Este sitio ofrece un punto común para una búsqueda rápida en cientos de diccionarios generales y diccionarios especializados en negocios, informática, medicina, religión, ciencia y tecnología, deportes y argot.

Robert Ware explica en septiembre de 1998: "Personalmente, casi estoy únicamente en contacto con personas que sólo hablan el inglés y que no tienen mucho incentivo para desarrollar sus habilidades lingüísticas. El hecho de que yo esté en contacto con el mundo entero cambia el enfoque, ¡y lo cambia para mejor! (...) Yo tardé un poco en incluir diccionarios de lengua no inglesa - en parte porque soy monolingüe - pero ahora se puede encontrar algunos." OneLook Dictionaries cataloga 2 millones de términos a partir de 425 diccionarios en 1998, 2,5 millones de términos a partir de 530 diccionarios en 2000, y 5 millones de términos a partir de 910 diccionarios en 2003.

Robert cuenta en la misma entrevista: "Una cosa interesante que ocurrió ha sido muy instructiva para mí. En 1994, trabajé para una escuela y trataba de instalar un programa en determinado tipo de computadora.

Conocí a una persona que tenía el mismo problema que yo, y empezamos a intercambiar correos electrónicos. De súbito yo me quedé impresionado... El programa había sido escrito a cuarenta kilómetros de distancia, pero quien me estaba ayudando era una persona encontrada en otra parte del mundo. ¡Las distancias geográficas ya no importaban! Bueno, esto está genial, pero ¿cómo hacer con eso? Yo puedo comunicarme sólo en inglés pero, afortunadamente, mi correspondiente podía escribir tanto en inglés como en alemán, que era su lengua materna. El internet ha eliminado una barrera, la de la distancia, pero sigue existiendo una barrera real, la del idioma.

Me parece que el internet impulsa a la gente simultáneamente hacia direcciones diferentes. El internet, de habla inglesa al principio, conecta a la gente en el mundo entero. De esta manera, promueve un idioma común para comunicarse. Pero también suscita contactos entre personas de varios idiomas y puede desarrollar

uninterés por el multilingüismo. Un idioma común es importante, pero de ninguna manera puede reemplazar la necesidad de muchos idiomas. Así el internet promueve tanto un idioma común como el multilingüismo, y eso es un factor que contribuye a encontrar soluciones. El creciente interés por los idiomas y "la necesidad de utilizarlos" estimulan en el mundo la creación de cursos de idiomas y de herramientas lingüísticas, y el internet ofrece la oportunidad de ponerlos a disposición de todos en la web de forma rápida y barata." Si el inglés sigue siendo predominante a finales de los años 1990, los sitios web bilingües o multilingües son cada vez más numerosos, por razones tanto comerciales como culturales que toman en cuenta el hecho de que todo el mundo no entienda el inglés.

Brian King, director del Worldwide Language Institute (WWLI), escribe en septiembre de 1998: "Así como los usuarios que no son de habla inglesa pueden ahora tener acceso a la tecnología en su propio idioma, el impacto del comercio electrónico puede ser una fuerza importante para convertir el multilingüismo en una vía natural hacia el ciberespacio. Los vendedores de productos y servicios en el mercado mundial virtual que ha llegado a ser el internet deben estar preparados a servir un mundo virtual tan multilingüe como el mundo físico. Si quieren tener éxito, ¡deben asegurarse de que están hablando el idioma de sus clientes!" Con cientos de socios académicos y comerciales, la red ELSNET (European Network in Language and Speech - Red Europea para la Lengua que Habla) tiene el objetivo de desarrollar sistemas multilingües para la lengua hablada y escrita.

Steven Krauwer, coordinador de ELSNET, explica en septiembre de 1998: "Como ciudadano europeo, me parece que el multilingüismo en la web es absolutamente esencial. En mi opinión, no es una situación sana a largo plazo que sólo la gente que entiende el inglés pueda aprovechar plenamente los beneficios de la web. Como científico (que se especializa en la traducción automática), veo el multilingüismo como un gran reto: garantizar que la información en la web esté accesible a todos, independientemente de las diferencias lingüísticas. (...) Pienso dedicar el resto de mi vida profesional a utilizar las tecnologías de la información para eliminar o al menos reducir la barrera del idioma."

Steven añade en agosto de 1999: "Cada vez estoy más convencido de que debemos tener cuidado con no abordar el problema del multilingüismo aislando del resto. Regresé de unas vacaciones en Francia, donde yo había pasado un muy buen verano. Aunque mi conocimiento del francés es superficial (esto es lo menos que se pueda decir), me sorprende que ya puedo comunicar sin problema al combinar este francés superficial con gestos, expresiones faciales, señales visuales, esquemas, etc. Creo que la web (a diferencia del sistema anticuo del correo electrónico textual) puede combinar con éxito la transmisión de la información a través de varios canales (o varios medios), mientras que este proceso sólo sería parcialmente satisfactorio si cada medio se tomara de forma aislada."

Para un verdadero multilingüismo en la web, sugiere varias soluciones prácticas: "(a) en lo que se refiere a los autores: una mejor formación de los autores de páginas web para explotar las posibles combinaciones que permitan mejorar la comunicación superando la barrera del idioma; (b) en lo que se refiere a los usuarios: programas de traducción de tipo Alta Vista Translation, cuya calidad no es excepcional, pero que tiene el mérito de existir; (c) en lo que se refiere a los programas de navegación: programas de traducción integrados, especialmente para los idiomas no dominantes, y diccionarios integrados más rápidos de consultar."

El multilingüismo es asunto de todos, como lo destaca la petición del CERCLE (Comité Européen pour le Respect des Cultures et des Langues Europe - Comité Europeo Para el Respeto de las Culturas y las Lenguas en Europa). La petición se emitió en 1998 en las once lenguas oficiales de la Unión Europea (alemán, danés, español, finlandés, francés, griego, holandés, inglés, italiano, portugués y sueco) para defender "una Europa humanista, multilingüe y rica de su diversidad cultural". El CERCLE propone a los editores del Tratado de la Unión Europea doce enmiendas para tener en cuenta el respeto de las culturas y lenguas en lugar de favorecer la lengua inglesa. La petición menciona entre otras cosas que "la diversidad y el pluralismo lingüístico no son un obstáculo a la circulación de personas, ideas y productos o servicios, como lo quisieran sugerir algunos aliados objetivos, conscientes o no, de la cultura y del idioma dominante. La normalización y la hegemonía son las que representan un obstáculo al desarrollo libre de los individuos, de las empresas y de la economía de lo inmaterial, principales fuentes de empleos en el futuro. El respeto de las lenguas, por el contrario, es la última oportunidad para acercar Europa a sus ciudadanos, un objetivo siempre anunciado pero casi nunca puesto en práctica. La Unión debe renunciar a dar prioridad a la lengua de un grupo único."

Bruno Didier, webmaster de la biblioteca del Instituto Pasteur en París (Francia), escribe en agosto de 1999: "El internet no es una propiedad ni nacional ni lingüística. Es un vector de cultura, y el primer soporte de la cultura es la lengua. Cuantas más lenguas estén representadas en toda su diversidad, más culturas

estarán representadas en el internet. No pienso que tengamos que ceder a la tentación sistemática de traducir las páginas a una lengua más o menos universal. Los intercambios culturales suponen una voluntad deponerse al alcance de la persona a quien queremos encontrar. Y este esfuerzo pasa por la comprensión de su lengua. Por supuesto, mis palabras son muy utópicas. Porque concretamente, mi actividad de vigilancia en la red me incita más bien a echar pestes contra algunos sitios noruegos o brasileños que no dan ninguna información en inglés, por más mínima que sea."

EL INGLÉS TODAVÍA PREPONDERANTE

= [Cita]

Marcel Grangier, director de la sección francesa de los servicios lingüísticos centrales de la Administración Federal Suiza, destaca en enero de 1999: "La supremacía del inglés no es nefasta en sí, en la medida en que resulta de realidades esencialmente estadísticas (más PC por habitante, más locutores de este idioma, etc.). La respuesta adecuada no es 'luchar contra el inglés' y tampoco basta con lamentarse, sino que hace falta multiplicar los sitios web en otros idiomas. Cabe añadir que en calidad de servicio de traducción, también preconizamos el plurilingüismo en los sitios web mismos. La multiplicación de los idiomas presentes en el internet es inevitable, y sólo puede hacerse en beneficio de los intercambios multiculturales."

= [Texto]

El inglés sigue siendo preponderante y las cosas no están por cambiar muy pronto. Con o sin razón, mucha gente deplora la hegemonía del inglés en el internet. Al principio era imprescindible, ya que el internet nació en los Estados Unidos y se desarrolló primero en Norteamérica antes de extenderse al mundo entero. En 1997, ya se observa que muchos idiomas están representados. Su presencia depende del dinamismo de cada comunidad lingüística.

En diciembre de 1997, Tim Berners-Lee, inventor de la web, declara a Pierre Ruetschi, periodista de La Tribune de Genève, un diario suizo: "¿Por qué las personas de habla francesa no ponen más información en la web? ¿Acaso piensan que nadie quiere consultarla en línea, que su cultura no tiene nada que ofrecer? Es una locura: por supuesto que la oferta es inmensa." Este reto se lleva a cabo en los años siguientes.

Profesor en tecnologías de la comunicación en la Webster University de Ginebra (Suiza), Henk Slettenhaar insiste en la necesidad de sitios web bilingües, en su idioma original y en inglés. Escribe en diciembre de 1998: "Las comunidades locales presentes en la web deberían utilizar antes de todo su propio idioma para difundir información. Si desean presentar esta información a la comunidad mundial, esta información deberá ser disponible también en inglés. Pienso que los sitios web bilingües son verdaderamente necesarios. (...) Pero estoy encantado de que ahora existan tantos documentos disponibles en su lengua original. Prefiero – y con mucho – leer el original con cierta dificultad antes que una traducción mediocre."

Henk añade en agosto de 1999: "A mi parecer, existen dos tipos de búsquedas en la web. En el primer caso, se trata de una búsqueda global en el campo de los negocios y de la información. Para llevarla a cabo, el idioma es primero el inglés, con versiones locales si es necesario. El segundo tipo de búsqueda concierne información local de toda clase en los lugares más remotos. Si la información se dirige a una etnia o a un grupo lingüístico, debe aparecer primero en el idioma de dicha etnia o de dicho grupo, con quizás un resumen en inglés."

Alain Bron, consultor en sistemas de información y escritor, explica en noviembre de 1999: "Se seguirán usando lenguas diferentes, y eso por mucho tiempo, tanto mejor para el derecho a la diferencia. El riesgo es por supuesto la invasión de una lengua en perjuicio de otras, y por lo tanto la nivelación cultural en detrimento de las otras. Pienso que poco a poco van a crearse servicios en línea para paliar esta dificultad. Al principio, algunos traductores podrán traducir y comentar textos a petición de los usuarios, y sobre todos los sitios más frecuentados invertirán en versiones en lenguas diversas, como lo hace la industria audiovisual." Philippe Loubière, traductor literario francés, denuncia por su parte la preponderancia del inglés que se ha apoderado de la red. Explica en marzo de 2001: "Todo lo que puede contribuir a la diversidad lingüística, tanto en la web como en otros ámbitos, es indispensable para la supervivencia de la libertad de pensamiento. No exagero en absoluto: el hombre moderno se juega en ello la supervivencia. Ahora bien, yo soy muy pesimista respecto a esta evolución. Los anglosajones suelen escribir a todo el mundo en inglés sin ningún escrúpulo. La inmensa mayoría de los franceses comprueba con una indiferencia total cómo el inglés macarrónico de los comerciales y de los publicistas va sustituyendo poco a poco su propio idioma, y el resto del mundo ha dado por perfectamente admitida la hegemonía lingüística de los anglosajones porque no tiene más perspectivas sino servir a esos ricos y poderosos dueños. La única solución consistiría en recurrir a legislaciones internacionales lo suficientemente restrictivas como para obligar a los gobiernos nacionales a respetar y a hacer respetar la lengua nacional en sus propios países (el francés en Francia, el rumano en Rumanía, etc.), y eso en todos los campos, y no sólo en internet. Pero no hay que hacerse ilusiones..."

Richard Chotin, profesor en la Escuela Superior de Negocios (ESA: École Supérieure des Affaires) de Lille (Francia), recuerda con razón que la supremacía del inglés sustituyó a la del francés. "El problema es político e ideológico, explica en septiembre de 2000. Es el problema del 'imperialismo' de la lengua inglesa que emana del imperialismo norteamericano. De hecho basta con recordar el 'imperialismo' de la lengua francesa a lo largo de los siglos 18 y 19 para comprender por qué se les da tan mal a los estudiantes franceses aprender idiomas. Al no tener que esforzarse para hacerse entender, uno no se esmera, y son los demás los que tienen que hacer esfuerzos."

Bakayoko Bourahima, bibliotecario en la Escuela Nacional Superior de Estadística y de Economía Aplicada (ENSEA: Ecole Nationale Supérieure de Statistique et d'Économie Appliquée) de Abiyán (Costa de Marfil), escribe en julio de 2000: "Para nosotros, africanos de habla francesa, la imposición del inglés en el internet representa para las masas una doble desventaja para acceder a los recursos de la red. Primero por culpa del problema de alfabetización que aún dista mucho de estar resuelto y que el internet enfatizará mucho más; luego se plantea el problema del dominio de una segunda lengua extranjera y de su adecuación al entorno cultural. En suma, si no se hace multilingüe, el internet nos impondrá una segunda colonización lingüística con todas las limitaciones que esto supone. Y esto nos deja de ser un gran problema porque a nuestros sistemas educativos ya les cuesta mucho trabajo optimizar sus prestaciones debido al uso del francés como lengua de instrucción básica, según dicen algunos peritos. Por lo tanto cada vez se evoca más la posibilidad de recurrir a las lenguas vernáculas para la instrucción básica, para 'desenclavar' el colegio en África e implicarlo lo mejor posible en la valorización de los recursos humanos. ¿Cómo proceder? En mi opinión no cabe la posibilidad para nosotros de hacer prevalecer cualquier clase de excepción cultural en la red, pues esto sería una reacción absolutamente gregaria. Entonces es preciso que los diferentes bloques lingüísticos se impliquen mucho para promover su propio acceso a la red, sin olvidar sus especificidades internas."

Guy Antoine, creador de Windows on Haiti, un sitio web de referencias sobre la cultura haitiana, cree en la necesidad del inglés como lengua común. Escribe en noviembre de 1999: "Por razones prácticas, el inglés seguirá dominando la red. No pienso que sea una cosa mala, a pesar de los sentimientos regionales que se oponen a eso, porque necesitamos una lengua común que permita favorecer la comunicación a nivel internacional. Dicho esto, no comparto la idea pesimista según la cual las otras lenguas solamente tendrían que someterse a la lengua dominante. Al contrario. Primero porque el internet puede reunir información útil para las lenguas minoritarias, que de otra manera correrían el riesgo de desaparecer sin dejar rastro. Además, en mi opinión, el internet incita a la gente a aprender las lenguas asociadas a las culturas que les interesan. Estas personas se dan cuenta rápidamente que la lengua de un pueblo constituye un elemento fundamental de su cultura."

Por este hecho, no les tengo mucha confianza a los instrumentos de traducción automática que, aunque traduzcan las palabras y expresiones, no pueden traducir el alma de un pueblo. ¿Qué son los haitianos, por ejemplo, sin el 'Kreyòl' (el criollo, para los no iniciados), una lengua que se desarrolló y que permitió unir entre ellas a varias tribus africanas desplazadas a Haití durante el periodo de la esclavitud? Esta lengua representa de la manera más palpable la unión de nuestro pueblo. Sin embargo es sobre todo una lengua hablada y no escrita. En mi opinión la situación irá cambiando gracias a la red, más que con cualquier otro medio tradicional de difusión de una lengua.

En el sitio web Windows on Haiti, la lengua principal es el inglés, pero también hay un foro de discusión animado en 'Kreyòl'. Existen también documentos sobre Haití en francés y en criollo antiguo colonial, y estoy dispuesto a publicar otros documentos en español y en otras lenguas. No propongo traducciones, pero el plurilingüismo es palpable en mi sitio, y me parece que esto no tardará en convertirse cada vez más en una norma en la red."

Tarde o temprano, ¿acabará correspondiendo la proporción de idiomas en la red con su repartición en el planeta? Nada parece menos cierto a la hora de la fractura digital entre ricos y pobres, entre zonas rurales y zonas urbanas, entre regiones favorecidas y regiones desfavorecidas, entre el hemisferio norte y el hemisferio sur, entre países desarrollados y países en desarrollo.

Según Zina Tucsnak, ingeniera de estudios en el Laboratorio ATILF (Analyse et Traitement Informatique de la Langue Française - Análisis y Procesamiento Informático de la Lengua francesa), entrevistada en octubre de 2000, "la mejor opción sería la aplicación de una ley por la cual se atribuyera una 'cuota' a cada idioma. Pero ¿no será acaso una utopía pedir la aplicación de una ley como ésta en una sociedad de consumo como la nuestra?"

En la misma fecha, Emmanuel Barthe, documentalista jurídico, expresa una opinión contraria: "Señales recientes dejan pensar que basta con dejar las lenguas tal y como están actualmente en la

web. Efectivamente, los idiomas – aparte del inglés – se desarrollan gracias a la multiplicación de sitios web nacionales que se dirigen específicamente a públicos nacionales, a fin de atraerlos hacia el internet. Basta con observar cuánto se han multiplicado las lenguas disponibles en las interfaces de los buscadores generalistas."

Olivier Gainon, creador de CyLibris, una editorial literaria electrónica, escribe en diciembre de 2000: "Me parece muy importante que los futuros protocolos permitan transmitir perfectamente muchos idiomas (...). Si se consigue esto, lo demás se derivará de ello: la representación de los idiomas se hará en función del número de personas conectadas, y al fin y al cabo habrá que contemplar la idea de usar buscadores multilingües."

Pierre Magnenat, responsable del servicio "gestión y previsión" del centro de informática en la Universidad de Lausana (Suiza), escribe en octubre de 2000: "La única solución que veo sería un esfuerzo mundial importante para desarrollar traductores automáticos. Creo que ningún incentivo o cuota bastaría para impedir la dominación total del inglés. Este esfuerzo podría - y debería - estar iniciado al nivel de los estados, y tener recursos suficientes para llegar a un buen resultado."

CODIFICACIÓN: DEL ASCII AL UNICODE

= [Cita]

Olivier Gainon, creador de CyLibris, una editorial literaria electrónica, escribe en diciembre de 2000: "Me parece muy importante que los futuros protocolos permitan transmitir perfectamente muchos idiomas – lo cual podría resultar bastante complicado (en las evoluciones futuras del lenguaje HTML o de los protocolos IP, etc.). Por lo tanto, es necesario que cada uno pueda sentirse a gusto con el internet y no sólo los individuos que dominen (más o menos) el inglés. No parece normal que hoy en día la transmisión de los acentos planteen problemas en los correos electrónicos."

= ASCII

El primer sistema de codificación informática es el ASCII (American Standard Code for Information Interchange – Código Estadounidense Estándar Para el Intercambio de Información). Publicado en 1968 por el ANSI (American National Standards Institute - Instituto Estadounidense Nacional de Normas) y actualizado en 1977 y 1986, el ASCII es un código de 128 caracteres convertidos en lenguaje binario en siete bits (A se convierte en "1000001", B se convierte en "1000010", etc.). Los 128 caracteres incluyen 33 caracteres de control (que no representan símbolos escritos) y 95 caracteres imprimibles: las 26 letras sin acento en mayúsculas (A-Z) y en minúsculas (a-z), las cifras, los signos de puntuación y algunos símbolos, que corresponden en su conjunto a las teclas del teclado estadounidense o del teclado inglés en Europa y en otros países.

El ASCII no permite más que la lectura del inglés y del latín. No permite tomar en cuenta las letras acentuadas presentes en gran parte de las lenguas europeas y aún menos los idiomas no alfabéticos (chino, japonés, coreano, etc.). Esto no planteaba ningún problema de importancia en los primeros años, cuando el intercambio de archivos electrónicos se limitaba esencialmente a Norteamérica. Pero el plurilingüismo pronto se convierte en una necesidad vital. Variantes del ASCII (norma ISO-8859 o ISO-Latin) toman en cuenta los caracteres acentuados de idiomas europeos. La norma ISO 8859-1 (Latin-1) es la norma utilizada en la mayoría de los países de Europa occidental, incluido para el español.

Yoshi Mikami es informático en la sociedad Asia Info Network en Fujisawa (Japón). En diciembre de 1995, lanza la página web bilingüe (inglés-japonés) "The Languages of the World by Computers and the Internet" (Los Idiomas del Mundo con Computadoras y el Internet), conocida comúnmente como Logos Home Page o Kotoba Home Page. Su página presenta un breve historial de cada idioma con sus características, su sistema de escritura, su mapa de caracteres y la configuración del teclado en dicho idioma. Yoshi Mikami también es el coautor (junto con Kenji Sekine y Nobutoshi Kohara) de "The Multilingual Web Guide" (Guía por una web multilingüe), publicado en japonés por las ediciones O'Reilly en agosto de 1997, y traducido al inglés, al alemán y al francés en 1998.

Yoshi explica en diciembre de 1998: "Mi lengua materna es el japonés. Al haber realizado mis estudios de tercer ciclo en los Estados Unidos y trabajado en informática, he llegado a ser bilingüe japonés-inglés americano. Siempre me he interesado por otras lenguas y culturas, lo que me llevó a aprender ruso, francés y chino sobre la marcha. A finales de 1995, he creado en la web el sitio 'The Languages of the World by Computers and the Internet' y he intentado proponer – en inglés y en japonés – un breve historial de todos estos idiomas, así como las características propias de cada lengua y de su fonética. Avalado por la experiencia adquirida, he instigado a mis dos asociados a escribir un libro sobre la concepción, la creación y la presentación de sitios web multilingües. Este libro se publicó en agosto de 1997 bajo el título 'The Multilingual Web Guide' y fue el primer libro sobre dicho tema."

¿Y cómo se imagina la evolución hacia una red multilingüe? "Miles de años atrás, en Egipto, en China y en otros lugares, la gente era más sensible a la necesidad de comunicar sus leyes y reflexiones en varios idiomas y no en uno solo. En nuestra sociedad moderna, cada Estado ha adoptado más o menos una sola lengua de comunicación. A mi parecer, el internet permitirá un uso más amplio de varios idiomas y de páginas multilingües - y no sólo una gravitación alrededor del inglés americano - y un uso más creativo de la traducción informática multilingüe. ¡99% de los sitios web creados en Japón están redactados en japonés!"
= Unicode

Con el desarrollo de la web, el intercambio de datos se internacionaliza más y más. Comunicar en muchos idiomas requiere un sistema de codificación que tome en cuenta muchos alfabetos e ideogramas. Además el paso del ASCII original a sus diversas extensiones no tarda en convertirse en un verdadero rompecabezas, incluso en la Unión Europea, donde se plantean problemas como la multiplicación de las variantes, la corrupción de los datos durante los intercambios informáticos, la incompatibilidad de los sistemas o la visualización de las páginas web en un solo idioma a la vez.

Publicado por primera vez en enero de 1991, el Unicode es un sistema de codificación de caracteres "universal" en 16 bits que asigna un número único a cada carácter. Este número es legible desde cualquier plataforma y con cualquier programa. El Unicode puede tomar en cuenta 65.000 caracteres únicos, es decir todos los sistemas de escritura del planeta. Para gran satisfacción de los lingüistas, progresivamente el Unicode va sustituyendo el ASCII, así como sus variantes UTF-8, UTF-16 y UTF-32 (UTF: Unicode Transformation Format). Se convierte en uno de los componentes de las especificaciones del World Wide Web Consortium (W3C), el organismo internacional encargado del desarrollo de la web.

Patrick Rebollar, profesor de literatura francesa en Japón y moderador de la lista de difusión LITOR (Literatura e Informática), destaca en enero de 2000: "El primer problema es un problema de software. Como se ve con Netscape o Internet Explorer, es posible fijar múltiples idiomas en la web. Pero la compatibilidad entre esos navegadores y otros programas (el Office de Microsoft, por ejemplo) aún no es sistemático. La adopción del Unicode debería resolver muchos problemas, pero esto supone volver a escribir la mayor parte de los programas, lo que los productores de programas se muestran reacios a hacer debido a los gastos, pues el rendimiento no está garantizado ya que los programas totalmente multilingües tienen menos interés para sus clientes que los programas de navegación."

La utilización del Unicode se generaliza alrededor del año 2000, por ejemplo para los archivos de texto bajo plataforma Windows (Windows NT, Windows 2000, Windows XP y siguientes versiones), que hasta entonces estaban en ASCII.

Pero el Unicode no puede solucionar todos los problemas, como lo subraya en junio de 2000 Luc Dall'Armellina, co-autor y webmaster de Vosite, un espacio de escritura hipermedia: "Los sistemas operativos se van dotando de fuentes Unicode capaces de representar todos los idiomas del mundo. Ojalá sigan el mismo rumbo todas las aplicaciones, desde el procesamiento de texto hasta el navegador web. Las dificultades son inmensas: nuestro teclado, con sus ± 250 teclas, deja ver sus insuficiencias siempre que es necesario digitalizar Katakana o Hiragana japoneses, y aún peor con el chino. La gran variedad de los sistemas de escritura por el mundo y el número de caracteres que abarcan constituyen un freno potente. Sin embargo los obstáculos culturales no son menos importantes, pues están vinculados con los códigos y modalidades de representación propios de cada cultura o etnia."

= ASCII y/o Unicode

El Proyecto Gutenberg es fundado en 1971 por Michael Hart con el objetivo de difundir bajo forma electrónica las obras literarias para ponerlas gratuitamente a la disposición de todos. Tanto los libros que fueron digitalizados hace treinta años como los de ahora se digitalizan en modo texto, mediante uso del ASCII original de siete bits. De esta manera los textos pueden ser leídos sin problema con cualquier computadora, plataforma y programa. Para los idiomas con acentos como el español, los libros tienen dos versiones: una versión en ASCII de ocho bits para tomar en cuenta los caracteres acentuados, y una versión en ASCII de siete bits, sin acentos. A excepción, claro, de las lenguas no codificables en ASCII, como el chino, que está codificado en Big-5.

El Proyecto Gutenberg ofrece libros en otros formatos, incluso en formatos populares tales como HTML, XML, RTF y Unicode. Además, cualquier formato propuesto por voluntarios es generalmente aceptado - PDF, LIT, TeX y muchos otros - siempre y cuando un archivo ASCII también esté presente.

En cuanto a idiomas, el Proyecto Gutenberg es en su mayoría de lengua inglesa, ya que está instalado en los Estados Unidos y que sirve principalmente a la comunidad de habla inglesa a nivel nacional e internacional. En octubre de 1997, Michael Hart anuncia su intención de intensificar la producción de libros en otros idiomas. A principios de 1998, el catálogo incluye algunas obras en alemán, español, francés, italiano y latín.

Michael escribe en julio de 1999: "Lanzamos ahora una nueva lengua al mes, y voy a seguir con esta política durante tanto tiempo como sea posible."

A la misión original se añade un papel de puente entre idiomas y culturas diversas. El multilingüismo se convierte poco a poco en una de las prioridades del Proyecto Gutenberg, como la internacionalización, con el Proyecto Gutenberg Australia (creado en agosto de 2001), el Proyecto Gutenberg Europa (creado en enero de 2004), el Proyecto Gutenberg Canadá (creado en julio de 2007) y otros Proyectos Gutenberg por venir en varios países. El Proyecto Gutenberg original (el de los Estados Unidos) tiene libros en 25 idiomas en enero de 2004 y libros en 42 idiomas en julio de 2005.

Desde su creación en enero de 2004, el Proyecto Gutenberg Europa (PGEuropa) y Distributed Proofreaders Europa (DP Europa) son sitios web multilingües que tienen en cuenta las principales lenguas europeas. El sitio de DP Europa está inspirado del sitio original de Distributed Proofreaders (el de los Estados Unidos) para gestionar la corrección de los libros digitales entre los voluntarios, que comparan las versiones digitales con las páginas escaneadas de las versiones impresas. En abril de 2004, gracias a traductores voluntarios, el sitio de DP Europa está disponible en doce idiomas. El objetivo a medio plazo es un sitio en sesenta idiomas que tenga en cuenta todas las lenguas europeas. DP Europa utiliza Unicode en lugar del ASCII, para procesar libros en muchos idiomas.

PRIMEROS PROYECTOS MULTILINGÜES

= [Cita]

Tyler Chambers, creador de dos proyectos en la web - la Human-Languages Page (Página de las Lenguas Humanas) y el Internet Dictionary Project (Proyecto de Diccionarios en el Internet) - cuenta en septiembre de 1998: "Mi actividad en línea consiste en favorecer el acceso de más gente a datos lingüísticos, a través de dos proyectos en la web. Aunque yo no sea polígloto, ni siquiera bilingüe, estoy consciente de que muy pocas áreas tienen una importancia comparable al área de las lenguas y del plurilingüismo (...) En general, creo que la web es importante para la sensibilización a las lenguas y a los temas culturales. ¿En qué otro lugar se puede navegar al azar durante veinte minutos y encontrar información útil en tres lenguas o más?"

= Travlang

Travlang, un sitio web dedicado a la vez a las lenguas y al viaje, fue creado en 1994 por Michael C. Martin, un estudiante en física. Michael siguió administrando este sitio muy popular cuando empezó a trabajar como investigador en el Lawrence Berkeley National Laboratory en California.

La sección Foreign Languages for Travelers (Idiomas Extranjeros para Viajeros) da la oportunidad de aprender en la web los fundamentos de sesenta idiomas. La sección Translating Dictionaries (Diccionarios de Traducción) proporciona acceso a diccionarios gratuitos en varios idiomas (afrikáans, alemán, checo, danés, español, esperanto, finlandés, francés, frisón, húngaro, italiano, latín, noruego y portugués). Estos diccionarios son a menudo breves y de calidad desigual. El sitio también ofrece muchos enlaces a servicios de traducción, escuelas de idiomas, librerías multilingües, etc.

Michael Martin escribe en agosto de 1998: "Pienso que la web es un lugar ideal para acercar las culturas y las personas, y eso incluye que la web sea multilingüe. Nuestro sitio Travlang es muy popular por esta razón. A la gente le encanta el contacto con otras partes del mundo. (...) De verdad el internet es un instrumento importante para comunicar con personas con las cuales no tendríamos la oportunidad de interactuar de otro modo. Agradezco la cooperación que ha hecho posible las páginas de Foreign Languages for Travelers. ¿Cómo ve Michael el futuro? "Creo que las traducciones integrales informatizadas se convertirán en algo común para comunicarse directamente con más gente. Esto también ayudará a que el internet alcance más a la gente que no habla inglés."

= La Human-Languages Page

Creada por Tyler Chambers en mayo de 1994, la Human-Languages Page (Página de las Lenguas Humanas) es un catálogo de 1.800 recursos lingüísticos en una centena de lenguas. Los recursos están clasificados en varias secciones: lengua y literatura, escuelas e instituciones, productos y servicios, agencias de empleo y formación, diccionarios, cursos de idiomas.

Otro proyecto web de Tyler es el Internet Dictionary Project (Proyecto de Diccionarios en el Internet), un proyecto colaborativo abierto a todos para la creación de diccionarios gratis en la web, del inglés a otros idiomas (alemán, español, francés, italiano, latín y portugués).

Según su sitio web, el propósito de este proyecto es "la creación de diccionarios de traducción con la ayuda del internet. Este sitio permite a la gente en el mundo entero utilizar estos diccionarios y participar en la traducción de términos ingleses a otros idiomas. Las listas de términos en inglés y sus equivalentes en otros idiomas están entonces a disposición de todos en este sitio web, sin restricción de ningún tipo. (...) El

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

