



Gana el contrato la empresa alemana Lucy Software que trabajará en colaboración con las vascas Emergiatech, UZEI y Gizartebitek

En quince meses disponible el traductor automático castellano-euskera promovido por el Gobierno Vasco

Patxi Baztarrika: hemos creado una nueva ocasión y herramienta para el fomento del uso del euskera

La Viceconsejería de Política Lingüística del Gobierno Vasco ha abierto el camino para disponer de un traductor automático de textos castellano-euskera que estará disponible antes de 15 meses. El proceso para dotar al euskera de esta herramienta electrónica ha sido largo y para su desarrollo la Viceconsejería de Política Lingüística del Gobierno abrió un concurso público para empresas especializadas que ha sido adjudicado a la alemana Lucy Software que para construir el traductor trabajará en colaboración con las empresas vascas Emergiatech, UZEI y Gizerbitek. El Gobierno Vasco invertirá 395.000 euros para la financiación de la nueva herramienta del euskera, en dicha cantidad se incluye el diseño, implantación y desarrollo del traductor automático.

En la presentación de la nueva herramienta del euskera, Patxi Baztarrika, Viceconsejero de Política Lingüística ha subrayado la importancia que tienen las tecnologías de la información y comunicación para todas las lenguas y especialmente para una lengua como el euskera, *en un futuro muy próximo las tecnologías no sólo diferenciarán los países desarrollados y los no desarrollados, sino también las lenguas desarrolladas y las no desarrolladas. Evidentemente, el euskera sólo encontrará y afianzará el lugar que le corresponde entre las demás lenguas, entre las lenguas desarrolladas, en la medida en que utilice estas nuevas tecnologías. La globalización no es una amenaza para los euskaldunes, sino una realidad, un reto. Y en este nuevo mundo, hay que considerar a las TIC como una oportunidad, aceptarlas y hacerlas nuestras. No estamos parados, vamos bien, estamos avanzando, sin ninguna duda. La presencia que el euskera hoy en día tiene en Internet y en las tecnologías lingüísticas es mucho mayor de la que le correspondería por su número de hablantes con relación al peso de número de hablantes del mundo. De todos modos, el euskera no podrá ocupar más que una pequeña parte en esa red; no nos conviene olvidar nuestra dimensión. La cuestión es que tenemos que ser capaces de aprovechar todas las oportunidades, en la mayor medida que podamos.*

El Viceconsejero de Política Lingüística ha puesto el acento en el uso de estas posibilidades, *el Gobierno Vasco ya ha cumplido su cometido, es decir, crear una nueva ocasión para utilizar la lengua, para que el euskera sea un idioma moderno, para poner las tecnologías al servicio de nuestra lengua, para que el euskera cuente con una herramienta (el traductor automático) de la que dispone cualquier otro idioma moderno. Por lo tanto, hemos creado una nueva ocasión y una nueva vía para el euskera, con el fin de que sea una lengua utilizada. Ahora corresponde a la ciudadanía aprovechar dichas opciones. Existen ya cantidad de herramientas tecnológicas*



en euskera, además dentro de año y medio podremos añadir a la lista el traductor automático, es por ello que quiero animar a la ciudadanía a utilizarlas. Crear y utilizar las oportunidades: esos son los dos cimientos sobre los que construirá su futuro el euskera y eso nos incumbe a todos y todas.

La Viceconsejería de Política Lingüística del Gobierno Vasco ha invertido en los últimos tres ejercicios (2006-2008) 4.740.723 euros para el desarrollo del euskera en el campo de las Tecnologías de la Información y Comunicación, de esa cantidad, 3.990.723 euros se han gestionado desde la convocatoria para las TIC de la Viceconsejería y otros 1.450.000 euros se han destinado al desarrollo de instrumentos, herramientas y localización de software.

El traductor automático del euskera desarrollado para el Gobierno Vasco se utilizará de tres maneras: para traducir textos cortos en Internet, para traducir páginas web y, a través de la intranet del Gobierno Vasco, para traducir textos largos. Estos últimos que serán corregidos y supervisados por un traductor humano lo que servirá además para alimentar y desarrollar el sistema. Esta herramienta de traducción español-euskera es un primer paso de cara a próximos desarrollos: euskera-español, euskera-inglés, inglés-euskera, euskera-francés, francés-euskera...

La **traducción automática**, es un desarrollo tecnológico crucial y necesario para el euskera, y por ese motivo ese ámbito se recoge tanto en el *Plan General de Promoción del Uso del Euskera* como en el Plan Euskadi en la Sociedad de la Información del Gobierno Vasco. La Viceconsejería de Política Lingüística lleva varios años en el proyecto, y en este sentido, su puesta en marcha ha sido un proceso largo y costoso, teniendo en cuenta que el nivel de desarrollo de estos instrumentos para nuestra lengua es inferior al de las lenguas mayoritarias.

Con el fin de desarrollar el traductor automático para el euskera, la Viceconsejería de Política Lingüística ha utilizado la siguiente metodología: en primer lugar contactó con las empresas que habían mostrado interés en integrar el euskera en sus productos y les pidió que elaboraran un informe técnico sobre cómo desarrollar con su tecnología el traductor para el euskera, así como proponer un plan de viabilidad. Sobre esa base y una serie de proyectos piloto, convocó un concurso público en el que ha valorado la calidad de la propuesta técnica presentada (estrategia, calidad de propuesta técnica: estrategia, metodología, tareas, acciones previstas y planteamiento general de la prestación) el programa de trabajo y el calendario de actuaciones, la oferta económica y las mejoras propuestas por cada licitador.

Teniendo en cuenta que la herramienta que el Gobierno Vasco pretende obtener será para traducir textos generales, no especializados, y que para la construcción de un sistema basado en las memorias de traducción y en los métodos estadísticos, sería necesario disponer de un gran corpus (del cual no se dispone en la actualidad), el modelo que en este momento permitirá obtener mejores resultados es el basado en reglas y transferencia. También se ha valorado especialmente que la tecnología utilice un análisis potente en la lengua de origen. La empresa que garantiza estas prestaciones ha resultado Lucy Software. Además, su tecnología dispone de un motor de gestión de memorias, por lo que a medida que el corpus aumente y mejore, ese recurso podrá ser utilizado para mejorar la herramienta. El sistema será un sistema vivo, que irá mejorando continuamente, siendo conscientes de que para que el resultado de la traducción automática entre dos lenguas de estructuras tan diferentes sea de un nivel aceptable, es necesario entrenar y utilizar el sistema para llevar a cabo las adecuaciones necesarias.



Lucy Software, Emergiatech, UZEI, Gizerbitek

Lucy Software es una empresa alemana que trabaja en el área de traducción automática y que ofrece servicios en dicha área. Su filial (100%) en Barcelona, es la empresa **Translendum S.L.** En la elaboración del traductor de euskera participarán tres empresas vascas:

- **Emergiatech:** es una spin-off del Grupo DELi de la Universidad de Deusto; su ámbito de trabajo es el procesamiento inteligente de la información basado en la ingeniería lingüística.

- **UZEI:** es el Centro de Terminología y Lexicografía. Se encarga de la actualización del Banco Terminológico Público Vasco EUSKALTERM. Ha elaborado corpus del euskera y diccionarios lexicográficos. Es el autor de la mayoría de los diccionarios terminológicos en euskera.

- **Gizerbitek:** empresa tecnológica; está especializada en desarrollos informáticos.

La tecnología de Lucy Software y Translendum esta presente desde hace años en el mercado con muy buenos resultados. Éstos son los pares de lenguas de los que dispone en la actualidad:

Catalán	→	inglés, español, francés,
Inglés	→	alemán, catalán, español, francés, ruso, italiano
Alemán	→	inglés, español, francés, ruso
Español	→	alemán, inglés, catalán, francés
Francés	→	alemán, inglés, español, catalán
Ruso	→	alemán, inglés

La Generalitat de Cataluña y la Xunta de Galicia tienen implantado el traductor de Translendum. También otras muchas administraciones públicas. Con la tecnología Translendum se traduce el diario de Lleida 'El Segre' y un usuario importante de esta tecnología también es la Agencia Catalana de Noticias.

Metodología para el desarrollo del traductor automático utilizada por la Viceconsejería de Política Lingüística

Con el fin de desarrollar el traductor automático para el euskera, la Viceconsejería de Política Lingüística ha utilizado la siguiente metodología: en primer lugar contactó con las empresas que habían mostrado interés en integrar el euskera en sus productos y les pidió que elaboraran un informe técnico sobre cómo desarrollar con su tecnología el traductor para el euskera, así como proponer un plan de viabilidad. Además, les pidió que realizaran una demostración práctica del uso de su tecnología. Para la elaboración de dicha prueba, cada empresa percibió la cantidad de 10.000 euros más IVA. En esta fase para conocer el estado del arte de la traducción automática participaron las siguientes empresas:

- AutomaticTrans (Cataluña)
- Eleka Ingeniaritza Linguistikoa (Euskadi)



- Emergiatech (Euskadi)
- Incyta (Cataluña)
- Interlan System (Euskadi)

Con el fin de valorar y evaluar la información obtenida en esta primera fase, la Viceconsejería de Política Lingüística, organizó una comisión compuesta por representantes de la administración y expertos externos. Toda la información adquirida durante la fase de pruebas y conocimiento de las diferentes tecnologías ha sido la base para la elaboración del pliego de condiciones técnicas de la contratación que, posteriormente, puso en marcha el Gobierno Vasco.

Así, el 4 de enero del 2008, se publicó en el BOPV el anuncio para la contratación de una herramienta de traducción automática del castellano al euskera, de textos generales, no especializados, a través de concurso público. La contratación incluía el suministro, implantación y mantenimiento inicial de dicha herramienta.

Se recibieron cuatro ofertas que respondían a lo requerido por el pliego de condiciones técnicas:

- AutomaticTrans (Cataluña)
- Lucy Software (Alemania-Cataluña)
- SHS Polar (Madrid)
- UTE Zelai Haundi (Euskadi)

La Viceconsejería de Política Lingüística ha llevado a cabo la valoración de las ofertas, con ayuda de EJIE, teniendo en cuenta los criterios prescritos, con el resultado de la adjudicación antedicha.

Patxi Baztarrika: *Crear y utilizar las oportunidades: esos son los dos cimientos sobre los que construirá su futuro el euskera.*

Las Tecnologías de la Información y de la Comunicación son un ámbito estratégico para el euskera. Todos somos conscientes de que vivimos en la era de la Tecnología de la Información y de la Comunicación y de que los idiomas que no consigan integrarse en dicho ámbito desaparecerán.

En un futuro muy próximo las tecnologías no sólo diferenciarán los países desarrollados y los no desarrollados, sino también las lenguas desarrolladas y las no desarrolladas. Evidentemente, el euskera sólo encontrará y afianzará el lugar que le corresponde entre las demás lenguas, entre las lenguas desarrolladas, en la medida en que utilice estas nuevas tecnologías.

La globalización no es una amenaza para los euskaldunes, sino una realidad, un reto. Y en este nuevo mundo, hay que considerar a las TIC como una oportunidad, aceptarlas y hacerlas nuestras. Esas tecnologías, por tanto, pueden ser grandes aliadas si hacemos bien nuestros deberes. Si utilizamos bien las tecnologías, el euskera tiene mucho que ganar y nada que perder. Y ése es, precisamente, nuestro objetivo. En efecto, entre todas las lenguas poderosas, el euskera necesita protección, tenemos que proteger el euskera; tenemos que utilizar la oportunidad que nos brindan estas tecnologías para proteger y reforzar el euskera.



Las TIC, la red, Internet: esa es la clave. Todo lo que quede fuera de esa dinámica está destinado a ser marginal. O lo que es peor, condenado a desaparecer. Por eso tienen una vital importancia la presencia del euskera en la red y la utilización de las tecnologías lingüísticas para el euskera. Son primordiales no sólo porque conviene estar ahí sino porque sin estar ahí el euskera se quedaría sin aliento y sin futuro.

Efectivamente, si queremos que el euskera siga ganando un futuro firme, es absolutamente necesario, indispensable, crear ocasiones de uso en todos los lugares y en todas las actividades. Y son muchas las opciones que están a nuestro alcance. Ya disponemos de muchas herramientas: diccionarios, el banco terminológico, el corrector ortográfico, aplicaciones para utilizar en diversos ámbitos... Sin embargo, el mismo reto que debemos afrontar con el euskera, debemos afrontarlo con los recursos tecnológicos, es necesario pasar de disponer de dichas herramientas a utilizarlas. La mayoría de los euskaldunes aún deben dar el paso cualitativo entre disponer de dichas herramientas y utilizarlas, ya sea por falta de costumbre, ya sea porque desconocen que esas herramientas también están disponibles en euskera, o por el motivo que sea.

No estamos parados, vamos bien, estamos avanzando, sin ninguna duda. La presencia que el euskera hoy en día tiene en Internet y en las tecnologías lingüísticas es mucho mayor de la que le correspondería por su número de hablantes con relación al peso de número de hablantes del mundo. De todos modos, el euskera no podrá ocupar más que una pequeña parte en esa red; no nos conviene olvidar nuestra dimensión. La cuestión es que tenemos que ser capaces de aprovechar todas las oportunidades, en la mayor medida que podamos.

El Gobierno Vasco ya ha cumplido su cometido, es decir, crear una nueva ocasión para utilizar la lengua, para que el euskera sea un idioma moderno, para poner las tecnologías al servicio del euskera, para que el euskera cuente con una herramienta de la que dispone cualquier otro idioma moderno (el traductor automático). Por lo tanto, hemos creado una nueva ocasión y una nueva vía para el euskera, con el fin de que sea una lengua utilizada. Ahora corresponde a la ciudadanía aprovechar dichas opciones. Hemos mencionado anteriormente que hoy en día existen ya cantidad de herramientas tecnológicas en euskera, además dentro de año y medio podremos añadir a la lista el traductor automático, es por ello que quiero animar a la ciudadanía a utilizarlas. Crear y utilizar las oportunidades: esos son los dos cimientos sobre los que construirá su futuro el euskera.

*La Viceconsejería de Política Lingüística es consciente de la importancia de las Tecnologías de la Información y de la Comunicación para la normalización del euskera. Por ello, cada año otorga mayores ayudas económicas para este ámbito, y también, por eso, cada año promueve nuevas herramientas y aplicaciones en euskera y las pone a disposición de la ciudadanía: diccionarios on line (monolingües: Euskal biztegia, el Diccionario Enciclopédico Harluxet; y diccionarios bilingües: el diccionario Elhuyar, el diccionario 3000 , el Diccionario Morris); el Banco Terminológico Público Vasco EUSKALTERM; todas las versiones del corrector ortográfico Xuxen (para Microsoft Office, para OpenOffice.org, para Mac Word, para QuarkXPress); plugins de diccionarios (diccionarios Elhuyar, diccionario de sinónimos de UZEI, OCR, hyphenator); sistemas operativos (Windows y Linux); paquetes de ofimática (Microsoft Office y OpenOffice.org); aplicaciones para la gestión empresarial (5 módulos de SAP, PRISMA), herramientas para el ámbito socio-económico (el software BDP Táctil), la biblioteca Miatu para el desarrollo de webs, etc. Todas ellas están a disposición del usuario en la página web www.euskadi.net/euskara_soft. Así mismo, La Viceconsejería de Política Lingüística elaboró también un **Inventario de las Tecnologías de la Información y de la Comunicación en Euskera** con la intención de reunir y ofrecer información sobre el área de las TIC. (www.euskara.euskadi.net/ikt_inbentarioa).*

Si mencionamos cifras, tenemos que decir que la Viceconsejería de Política Lingüística en los tres últimos años (2006, 2007, 2008) ha realizado las siguientes inversiones en la promoción de las TIC en euskera:



SUBVENCIONES (dentro de las convocatorias de las TIC)	3.540.723 euros
Otros proyectos (herramientas, localizaciones de software...)	1.990.000 euros

*La **traducción automática**, es un desarrollo tecnológico crucial y necesario para el euskera, y por ese motivo ese ámbito se recoge tanto en el Plan General de Promoción del Uso del Euskera como en el Plan Euskadi en la Sociedad de la Información.*

El proceso de puesta en marcha del proyecto del traductor automático para el euskera ha sido largo y costoso, teniendo en cuenta que el nivel de desarrollo de los soportes técnico de nuestra lengua es inferior al de las lenguas mayoritarias. Sobre la base de la metodología de trabajo anteriormente citada, la Viceconsejería de Política Lingüística ha concluido el proceso que posibilitará que antes de 15 meses la primera herramienta de traducción automática de textos del castellano al euskera sea una realidad. ¡El camino está abierto!

Donostia-San Sebastián, 24 de julio de 2008