

El euskera y las tecnologías del habla

Inmaculada Hernaez

Universidad del País Vasco / Euskal Herriko Unibertsitatea

1 Introducción

Según datos de 1996¹, el euskera se habla por más de un millón de personas, considerando tanto vasco-parlantes activos como pasivos, sólo en las zonas de la Comunidad Autónoma Vasca (con un 50,6% de vasco-parlantes) y Navarra (con un 16,5% de vasco-parlantes). Todo ello excluyendo las zonas francesas. Siempre según las mismas fuentes oficiales citadas, de entre los vasco-parlantes de la CAV (es decir, sin considerar Navarra y las zonas francesas), un 24% hablan su lengua materna, mientras que el resto han aprendido el euskera *batua* o euskera estándar en la enseñanza primaria o ya de adultos.

En el último mapa de dialectos publicado por Koldo Zuazo se distinguen actualmente 6 dialectos, y de acuerdo con la clasificación de Bonaparte realizada en 1869, 25 subdialectos y 50 variedades dialectales. Considerando además que el euskera estándar normaliza y reglamenta únicamente la escritura de la lengua, la realidad hablada del euskera presenta una enorme diversidad.

El impacto de esta diversidad en el desarrollo de las tecnologías del habla puede considerarse desde dos puntos de vista: el fonético, y el prosódico. Con el respecto al primero, podemos destacar la dificultad para generar un sistema de transcripción fonética automática que satisfaga a todos los usuarios, y las consecuencias de ello en el campo de la conversión de texto a voz, y en el reconocimiento de voz, con las dificultades asociadas para la elaboración de las bases de datos. Con respecto a la prosodia, los modelos existentes distan mucho entre sí. Así, por ejemplo, el acento, tan estudiado por los lingüistas, no está tampoco reglamentado, por lo que se realiza por los individuos en función de sus experiencias y educación personal.

¹ Datos del Instituto Vasco de Estadística y del Instituto de Estadística de Navarra, basados en el censo de 1996.

En el resto de apartados de este artículo se presentan de forma breve los recursos orales disponibles y las aplicaciones informáticas, así como algunos de los productos y servicios en el mercado.

2 Recursos Orales

Uno de los pilares del desarrollo de las tecnologías del habla para una lengua son los recursos orales o bases de datos conteniendo voz debidamente etiquetada y segmentada en función de la finalidad para la que se desee desarrollar. En este apartado se describen brevemente los recursos actualmente existentes para el euskera, así como aquellos que se encuentran en fase de desarrollo.

2.1 Bases de datos para síntesis de voz

En los laboratorios del grupo de investigación TSR de la UPV/EHU se han generado de una forma progresiva recursos orientados a obtener los datos necesarios para el desarrollo de un sistema de conversión de texto a voz. Los recursos orales disponibles pueden clasificarse en dos grupos: aquellos que permiten realizar el modelado prosódico de la lengua (modelado de la entonación, la duración y el ritmo) y por otra parte, las bases de datos de unidades para la síntesis por concatenación de unidades. Las bases de datos están descritas con más detalle en (Etxebarria *et al.*, 2000), y la *Tabla 1* muestra un resumen de las características de las mismas.

| Data Base | Locutor | Dominio | #frases | #palabras |
|--------------------------|---------|----------|----------|-----------|
| <i>Julen-unidades</i> | H (*) | U.S. (*) | 1758 | 12351 |
| <i>Julen-prosodia</i> | H | Prosodia | 1096 | 11136 |
| <i>Ion</i> | H | U.S. | - | 1403 |
| <i>Amaia</i> | M | U.S. | - | 1228 |
| <i>Juanjo</i> | H | U.S. | - | 1228 |
| <i>Euskadi-irratia</i> | H | Noticias | 1633 | 9776 |
| <i>Jokin</i> | H | Prosodia | 343 | 2499 |
| <i>Bermeo</i> | M | Prosodia | 348 | 1376 |
| <i>Variedades vascas</i> | H & Me | Prosodia | 120 * 25 | - |

Tabla 1.- Características de las bases de datos para síntesis de voz en los laboratorios de investigación de la ESIB. (*)Locutor: H: Hombre, M. Mujer. Dominio: U.S.: Unidades de Síntesis,

Las tres primeras bases de datos se desarrollaron dentro de contratos de investigación con entidades privadas (las dos primeras para Telefónica I+D, y la tercera para Euskaltel, S.A.). Todas las demás se encuentran disponibles en los laboratorios mencionados.

2.2 Bases de datos para Reconocimiento de Voz

La base de datos para el reconocimiento de la lengua vasca probablemente de mayor volumen existente hoy en día es propiedad de Telefónica I+D. Es una base de datos telefónica (8kHz), constituida por unas 19000 llamadas, aunque sólo una parte de ella se encuentra etiquetada. Esta base de datos se desarrolló en el contexto de un proyecto de colaboración entre la UPV/EHU² y Telefónica I+D, con la participación de EITB³ y UZEI⁴. Además, existen otras bases de datos desarrolladas en la UPV/EHU, véase por ejemplo (López de Ipiña, 1995).

En la actualidad, en el laboratorio TSR de la UPV/EHU estamos elaborando una base de datos SpeechDat⁵ para el euskera de 1000 locutores. Para su desarrollo se han considerado las zonas de habla vasca ubicadas en la Comunidad Autónoma Vasca y Navarra.

2.3 Otras bases de datos: Bizkaifon

Bizkaifon es una base de datos constituida por archivos sonoros del euskera vizcaíno accesible por web. Es todavía un proyecto en desarrollo⁶, pero un prototipo de ella puede visitarse en <http://bips.bi.ehu.es/fonoteka>.

El objetivo del proyecto Bizkaifon es recoger, preservar y clasificar material sonoro existente y desarrollar una arquitectura de software abierta en la que puedan ser integrados diferentes sonoros. La principal motivación del desarrollo de este proyecto la constituyeron las mas de 5000 horas de grabaciones realizadas por el lingüista Iñaki Gaminde a lo largo de los últimos 20 años de trabajo de campo por las zonas del habla

² Dpto. de Electricidad y Electrónica y Dpto. de Filología Vasca

³ Euskal Irrati Telebista

⁴ Centro de terminología y lexicografía. <http://www.uzei.com>

⁵ <http://www.speechdat.org>

⁶ Desarrollado en el grupo TSR, Dpto. de Electrónica y Telecomunicaciones de la UPV/EHU. Proyecto financiado por la Diputación Foral de Bizkaia /Bizkaiko Foru Aldundia.

de euskera vizcaino. Las grabaciones se encontraban en diferentes estados de conservación, clasificación y formatos: desde las cintas de cassette de los primeros años a las más modernas realizadas en Minidisk, y algunas de ellas digitalizadas y ya etiquetadas con transcripciones, e incluso publicadas en formatos digitales (Gaminde, 1999-2001). El proyecto unificaría formatos, y criterios y métodos de clasificación y etiquetado de todas esas grabaciones, constituyendo al mismo tiempo la estructura de base que permitiera la integración de otros recursos, nuevos o existentes. Creemos que las aplicaciones y usos que una base de datos de este tipo puede generar (aparte de la propia aplicación web de acceso a los datos) son muchas y muy variadas: en los campos de la educación, la investigación socio-lingüística, la dialectología etc.

La información de las señales se encuentra almacenada en ficheros de texto con formato SGML. Además de incluir diversos datos sobre el origen y procedencia de la grabación, estos ficheros contienen transcripciones de las señales sonoras, realizadas a varios niveles. Así, en función del tipo de documento sonoro (palabras sueltas, frases, textos diversos, literatura popular), se incluyen las transcripciones al euskera euskera batúa y al euskera vizcaíno, junto con enlaces que asocian la transcripción los instantes de tiempo en que se producen, soportando el sistema diferentes niveles de transcripción o etiquetado en general y segmentación.

3 Reconocimiento de voz

Los trabajos más importantes en el campo de reconocimiento de voz para el euskera se han desarrollado en la UPV/EHU, véase por ejemplo (López de Ipiña, 1995, López de Ipiña, 2000).

Cabe citar también los trabajos desarrollados en el departamento de Ing. de Sistemas y Automática en Escuela de Ing. Técnica de Vitoria para el desarrollo de aplicaciones basadas en el SDK de ViaVoice de IBM (Anega, 2001).

4 Conversión de texto a voz

Además del sistema de conversión de texto a voz que se describirá a continuación, cabe citar al grupo de investigación de la UPNA⁷, en el

⁷ Grupo de Procesado de Señal, microelectrónica e instrumentación, Dpto. Ing. Eléctrica y Electrónica

cual se han realizado trabajos basados en las técnicas de síntesis por concatenación PSOLA.

El primer sistema de conversión de texto a voz para el euskera fue bautizado *Ahozka* (Hernaiz, 1994) y está basado en un sintetizador de formantes. Desde entonces, el sistema ha evolucionado considerablemente, hasta convertirse en *AhoTTS*, un sistema modular de conversión de texto a voz, multiplataforma, multilingüe y multihilo. AhoTTS puede ser utilizado bien como plataforma de desarrollo de síntesis, o como una API (*Application Program Interface*) para ser llamada desde cualquier aplicación. La arquitectura modular del sistema permite trabajar de forma independiente con cada uno de los bloques que forman el conversor: el módulo normalizador, el módulo de análisis lingüístico, el módulo de conversión a unidades de síntesis y finalmente el motor de síntesis. Es por tanto muy flexible, y puede ser configurado con diferentes lenguajes, voces (bases de datos de unidades), modelos prosódicos (modelos de entonación, duraciones, pausado) y motores de síntesis (Formantes, PSOLA, MBROLA). Una descripción más detallada puede encontrarse en (Hernaiz, 2001).

5 Productos y servicios

Dejando a un lado los conocidos sistemas de audiotext que emiten mensajes pregrabados, el único servicio que utiliza tecnologías de síntesis y reconocimiento en euskera lo ofrece Telefónica S.A.. El servicio ofrece básicamente la posibilidad de leer el correo electrónico a través del teléfono, y es ofertado en las cuatro lenguas oficiales (castellano, gallego, catalán y euskera). Las tecnologías de síntesis y reconocimiento que utiliza han sido desarrolladas por Telefónica I+D, en proyectos de colaboración con diferentes universidades, y para el caso del euskera con la UPV/EHU⁸.

Además, se han ofertado algunos servicios (por ejemplo, el servicio de *cita previa* de la Diputación Foral de Bizkaia para la realización de la declaración de la renta) que empleando sistemas de reconocimiento desarrollados para el castellano, se han adaptado para el reconocimiento de algunas palabras en euskera.

⁸ Con el Dpto. de Electricidad y Electrónica, y Filología Vasca para el caso de reconocimiento de voz, y con el Dpto. de Electrónica y Telecomunicaciones para la conversión de texto a voz.

6 Conclusiones

El euskera dispone actualmente de aplicaciones y recursos suficientes para el desarrollo de servicios en el campo de las tecnologías del habla. El interés de las empresas locales por este tipo de tecnologías es actualmente bajo y no parecen ver un mercado interesante en el público de habla vasca. Por eso, es imprescindible contar con el apoyo de las instituciones para el desarrollo de los recursos y aplicaciones de base. Afortunadamente, la ingeniería lingüística se ha declarado como una de las líneas prioritarias en el último Plan de Ciencia y Tecnología del Gobierno Vasco, y esperamos que con este nuevo impulso, en poco tiempo el euskera presente un desarrollo en este campo tecnológico al nivel del resto de las lenguas europeas.

Referencias

- Anega, I. Juan José Juguete (2001) Ahots bidez kontrolaturiko etxe domotikoa . *Proyecto Fin de Carrera. Escuela Ingeniería Técnica de Vitoria (UPV/EHU)*.
- Etxebarria B., E. Navas, A. Armenta, I. Madariaga, I. Gaminde, I. Hernaez. (2000) "Tools and Basque language databases developed in the AhoLab laboratory" *Second International Conference on Language Resources and Evaluation. Workshop Proceedings, pp.62-70, Atenas 2000*
- Gaminde, I. (1999-2001) Serie "Berbarik Berba": Ayuntamientos de Bakio, Zamudio, Meñaka, Leioa, Mungia, Arrieta, Jatabe.
- Hernández, I., J. C. Olabe, P. Etxeberria, B. Etxebarria, A Cuesta (1994) "AHOZKA: Un sistema de conversión de texto a voz para el euskara" *SEPLN, Boletín nº 14, 1994*.
- Hernaez, I., E. Navas, J.L. Murugarren, B. Etxebarria (2001) "Description of the AhoTTS Conversión System for the Basque Language" *Proc. of the 4th Speech Synthesis Workshop*,
- López de Ipiña, K, I. Torres, L.Oñederra (1995) "Design of a phonetic corpus for a speech database in Basque language" *Proc. of the EUROSPEECH' 95, Vol. 1 , pp.8541-854 Madrid*
- López de Ipiña, K., I. Torres, L. Oñederra, N. Ezeiza, M. Hernández, M. Peñagarikano, A. Varona, L.J. Rodríguez, (2000) "First Selection of Lexical Units for Continuous Speech Recognition of Basque" *Proc. of the ICSLP, Vol.2 pp. 531-535, Pekín*.