

La edición enriquecida y en paralelo para el estudio del patrimonio documental vizcaíno*

(The enriched parallel edition for the study of Biscayan documentary heritage)

Fernández Carrelo, Patricia

Automatic Trans. Fontanella 10, entlo. Puerta 3ª. 08010 Barcelona

Gómez Seibane, Sara

Univ. de Castilla la Mancha. Fac. de Letras de Ciudad Real. Dpto.

Filología Hispánica y Clásica. Avda. Camilo José Cela, s/n. Campus

Universitario. 13071 Ciudad Real

BIBLID [1137-4454 (2009), 24; 87-98]

Recep.: 27.07.2006

Acep.: 26.05.2009

El presente artículo plantea una reflexión sobre el tratamiento digital del patrimonio documental vizcaíno a partir de la presentación de un estudio que conjuga el análisis del diasistema de un documento notarial y su tratamiento informático por medio del lenguaje de marcación XML.

Palabras Clave: Patrimonio documental vizcaíno. Diasistema de un documento. Edición. Textos enriquecidos. XML. TEI. Corpora paralelos.

Artikulu honek Bizkaiko dokumentuek osatzen duten ondarearen tratamendu digitalari buruzko hausnarketa jasotzen du. Horretarako, notario dokumentu baten diasistema hartu du oinarri gisa, eta XML markazio lengoaiaren tratamendu informatikoa eman dio agiriri.

Giltza-Hitzak: Bizkaiko dokumentu ondarea. Dokumentu baten diasistema. Edizioa. Testu aberratuak. XML. TEI. Corpora paraleloak.

Cet article lance une réflexion sur le traitement numérique du patrimoine documentaire en se basant sur la présentation d'une étude qui conjugue l'analyse du diasystème d'un document notarial et son traitement informatique à travers le langage à balises XML.

Mots Clé : Patrimoine documentaire biscaïen. Diasystème d'un document. Édition. Textes enrichis. XML. TEI. Corpus parallèles.

*. Este trabajo ha contado con una ayuda a la investigación de Eusko Ikaskuntza, 2005.

1. LA CONSERVACIÓN DEL PATRIMONIO DOCUMENTAL

Actualmente, los archivos están considerados patrimonio de la humanidad y los documentos que contienen representan la expresión de la actuación humana en lo relativo a las relaciones de los hombres y de las instituciones sociales. En la misma línea, la sociedad está manifestando un interés creciente por la apertura y la buena conservación de los archivos de las administraciones públicas, así como de archivos históricos, regionales, municipales, eclesiásticos, empresariales y privados.

El tratamiento archivístico del patrimonio documental en lo relativo a conservación, almacenamiento y difusión se ha planteado durante décadas de forma exclusiva desde la óptica de la microfilmación. Esta técnica, consistente en la fotografía de documentos en formato de película de 16 a 35 mm, representó en su momento una gran contribución para el depósito y preservación de la información. Sin embargo, la aplicación de las TIC al manejo documental ha suscitado una serie de inquietudes respecto a cuál es la herramienta más recomendable para la conservación documental. La práctica parece demostrar que, si bien el deterioro por uso es menor en los microfilmes, actualmente existen nuevas técnicas, como la digitalización, que ofrecen numerosas ventajas añadidas. A lo largo de estas líneas, especialmente en 3.1, se hará referencia a las características y aportaciones de la versión digital de las fuentes documentales de esos mismos archivos, así como a la edición en sí misma, que en los últimos años ha experimentado un notable incremento.

1.1. Eusko Ikaskuntza y la edición de Fuentes documentales medievales del País Vasco

En el caso del País Vasco, diversas instituciones están llevando a cabo un esfuerzo en lo que se refiere a esta edición de *corpora* documentales. Así, el objetivo de Eusko Ikaskuntza desde 1982 es la transcripción e impresión del acervo documental medieval del área vasca -desde la Alta Edad Media hasta 1520-, seleccionado por unidades archivísticas y de acuerdo a un criterio histórico o paleográfico.

Estas ediciones (en total 126 hasta el año 2005) constituyen una ayuda de gran importancia para la localización y primer acercamiento a los documentos puesta al servicio de investigadores de diversas áreas. Este es el caso del Seminario Alfonso Irigoien (SAI), coordinado por C. Isasi, que ha generado una serie de trabajos que, por un lado, contribuyen a la reconstrucción de la historia del romance en las áreas vizcaína y guipuzcoana y, por otro lado, aportan noticias sobre la presencia de elementos de la lengua vasca en una etapa en la que la documentación romance es el único vehículo escrito¹.

1. Naturalmente, para numerosas cuestiones los trabajos han exigido la comprobación en los documentos originales, lo que no disminuye la utilidad de las ediciones publicadas.

2. EL ESTUDIO DE FUENTES ROMANCES VIZCAÍNAS

Como se acaba de señalar, en los últimos años las aportaciones de Isasi y de los investigadores del grupo SAI han permitido conocer algunos de los rasgos del romance de la *scripta* vizcaína de época medieval y posmedieval, de gran interés dado que la documentación civil romance anterior a esta época es muy escasa para la mayor parte del País Vasco.

Este romance se ha identificado con el castellano en el marco del continuo dialectal norteño (Isasi 2006c). Asimismo, se ha descrito la presencia de fenómenos lingüísticos divergentes de las cronologías comúnmente aceptadas y/o elementos coincidentes con peculiaridades de difusión geográfica restringida (Isasi 2000, 2002c; Ramírez Luengo 2001, 2002c). Los documentos contienen, además, características presentes en áreas orientales –riojana y navarra– (Gancedo 2006; Isasi 1998, 2002a, 2002b), occidentales –asturiana y cántabra– (Gómez Seibane 2006b) y, en ocasiones, propias del castellano del norte de Burgos (Isasi 2006c)².

Por añadidura, se han detectado otros fenómenos que se encuentran en relación directa con la presencia de la lengua vasca, como el componente léxico (Isasi *et al* 2005) o la hipótesis de que la peculiar resolución de las oposiciones de sibilantes dentales en el ámbito vizcaíno tenga relación con la presencia de la lengua vasca (Isasi 2006a, 2006b).

3. LA EDICIÓN DIGITAL PARA EL ESTUDIO LINGÜÍSTICO

No nos detendremos aquí en la conocida polémica sobre la validez de las ediciones, tanto de textos notariales como de obras literarias, como base para la investigación “rigurosa”. No obstante, coincidimos con Isasi (2006c) en que el rechazo del uso de estas ediciones, por un lado, obliga a los lingüistas a consultar el manuscrito (donde quiera que se encuentre), y por otro, anula una de las labores más reivindicadas por la comunidad científica, como es la elaboración de *corpora*, que se verían relegados a fuentes de información léxica. Por tanto, al igual que C. Isasi (2006c), sostenemos que tales *corpora* son aprovechables en los casos en los que los criterios de transcripción adoptados sean homogéneos dentro del propio corpus y se expliciten, además, a los usuarios.

3.1. La edición en los nuevos soportes

Con la aplicación de las TIC a la preservación de las obras manuscritas surge una nueva perspectiva de análisis de los productos editoriales. A cuestiones

2. Hay que señalar, sin embargo, que esta labor de caracterización resulta difícil, debido a las limitaciones actuales del conocimiento de los textos vascos, la escasa posibilidad de contraste homogéneo con descripciones sobre textos notariales de áreas próximas, así como el hecho de que los datos sobre el castellano del área vasca occidental continúa fuera de las reconstrucciones diacrónicas del complejo castellano (Isasi 2004).

filológicas clásicas como el “concepto” del texto que se pretende editar o el tratamiento de los aspectos formales, se suma un nuevo componente definitorio de los distintos tipos de edición: el soporte digital.

Al revisar la tipología editorial existente vemos cómo la perspectiva desde la que tratar las publicaciones textuales se ve condicionada en mayor o menor medida por los nuevos medios. La edición bedierista, clave metodológica de la mayor parte de las ediciones en papel generadas en la modernidad, conserva su primacía en la etapa digital; la neolachmanniana, edición crítica en sentido estricto, casi desaparece de entre las propuestas digitales, aunque su estructura comienza a elaborarse ya por medio de herramientas de cotejo automático; la creación de la edición genética se ve facilitada por el soporte informático; y la múltiple, citada con frecuencia como representante de la llamada edición sinóptica, sirve como base o precedente de las ediciones de textos alineados o en paralelo, de las que se hablará en 4.3.2.

El entorno digital, además, facilita y transforma en un proceso dinámico la inserción de un aparato de notas enfocado bien a la aclaración semántica (de palabras desusadas o de aquellas cuyo significado en el texto sea diferente del que tienen en el léxico actual), bien a proporcionar información sobre entidades extratextuales a las que el texto haga referencia (sucesos, personajes, instituciones etc.) -ediciones anotadas-; permite, de forma sencilla, acompañar al texto de una reproducción de las imágenes de la fuente, -edición facsímil-; y/o mantener, resaltar o subrayar visualmente el tratamiento dado a los aspectos formales en función de los deseos del editor -ediciones paleográfica, crítico-interpretativas, conservadoras y modernizadoras-.

3.2. La edición digital enriquecida

El marco de los nuevos recursos tecnológicos aporta múltiples posibilidades de reutilización textual gracias a la digitalización tanto de manuscritos como de obras anteriormente editadas en papel. Entre otras cuestiones, favorecen la divulgación de textos relegados al olvido o de difícil acceso, permiten su difusión y accesibilidad a través de medios públicos, como la World Wide Web, o facilitan el análisis documental por medio de determinadas herramientas.

Actualmente esta digitalización, proceso de distintos grados de complejidad en función del mecanismo adoptado, abarca desde la simple memorización de un documento en procesadores de estilo WYSIWYG³, hasta la preparación de un texto enriquecido mediante la inserción de *metacontenidos*⁴. Para hacer explícitos los contenidos de los objetos documentales se emplean los denominados

3. Con este término “what you see is what you get” queremos hacer referencia a aquellos procesadores que muestran durante la propia edición una representación gráfica del documento que pretende ser un reflejo de la versión impresa que producirá.

4. Información sobre el texto mismo. Para más información sobre este concepto *vid.* Abaitua 2003.

lenguajes de etiquetado o metalenguajes, como XML (*eXtensible Markup Language*), herramientas que permiten por medio de etiquetas (*tags*) añadir información metatextual.

En función de la información textual que se quiera anotar y entre las diversas propuestas que recogen y sistematizan estos datos para ofrecer un estándar de etiquetado que responda a los objetivos de los investigadores y estudiosos del texto, destaca, dentro del ámbito filológico, la propuesta TEI (*Text Encoding Initiative*)⁵. Esta iniciativa, que goza de una gran difusión y apoyo por parte de la comunidad internacional de humanistas e investigadores de este ámbito, ha convertido en lenguaje de marcación -mediante su repertorio de etiquetas- procesos como la descripción bibliográfica o de material de libros y manuscritos, la transcripción de fuentes primarias o la resolución de exigencias específicas de algunos métodos como la inclusión del aparato de variantes en la edición crítica neolachmanniana. Así, siempre y cuando se persiga alguno de los objetivos definidos en las directrices de TEI, es posible aplicar este código de anotación a cualquier lengua, en diacronía o sincronía, y a cualquier texto o conjunto de textos, tal y como se muestra en el siguiente apartado.

4. FUENTES VIZCAÍNAS Y EDICIÓN DIGITAL: NUEVA APROXIMACIÓN A LAS FUENTES ROMANCES VIZCAÍNAS

Como ejemplo de una nueva aproximación al estudio de las fuentes romances vizcaínas que conjuga el tratamiento informático del patrimonio documental, el enriquecimiento textual (todo ello a través de lenguajes de marcación estandarizados) y el estudio del romance vizcaíno, se expone a continuación un proyecto financiado por Eusko Ikaskuntza-Sociedad de Estudios Vascos, actualmente en su segunda fase de desarrollo: *Un ejemplo de reutilización del patrimonio documental vizcaíno. El análisis lingüístico del diasistema de un documento por medio de la marcación en XML*⁶ (P. Fernández Carrelo y S. Gómez Seibane).

Este trabajo constituye un acercamiento al análisis de documentos notariales a través de una perspectiva interdisciplinar, pues se aproxima al estudio lingüístico de las diferencias presentes en las sucesivas copias realizadas por escribanos vizcaínos de un documento real original por medio del tratamiento informático del texto, que ha sido enriquecido con el lenguaje de marcación XML.

Entre sus objetivos, se encuentra, por una parte, confeccionar la edición en formato digital de las copias de un documento real de confirmación y mejora del fuero de Bilbao (1372) que datan de los años 1499, 1543, 1562, 1570, 1661, 1736, y una más del siglo XVIII; y por otra, llevar a cabo el estudio lingüístico contrastivo a través de herramientas informáticas que automaticen el proceso. La finalidad última de este ensayo radica, por tanto, en el tratamiento informático

5. tei-c.org

6. Primera fase: marzo 2005-marzo 2006. Segunda fase: marzo 2006-marzo 2007, EI-SEV.

del diasistema de un documento notarial vizcaíno con vistas a comprobar la validez del análisis lingüístico a través de una metodología digital.

Conviene señalar que no se ha tratado en ningún momento de realizar una labor de reconstrucción de las filiaciones o relaciones que se establecen entre los diferentes testimonios, de ahí que el interés del trabajo se centre en el diseño de un sistema de etiquetado aplicable al patrimonio documental vizcaíno con la intención, además, de que sea reutilizable.

4.1 Descripción de las fuentes

Así pues, fundamentalmente por su dilatada tradición textual, se seleccionó un documento emitido por Juan I en 1372 que, tras la invocación a la Trinidad y a la Virgen, expone, mediante una fórmula protocolaria de notificación universal muy frecuente en la Edad Media, que confirma y mejora los privilegios concedidos por sus antecesores.

Dado que el original en pergamino de 1372 resultaba ilegible⁷, el texto tomado como base para el cotejo de las copias fue el de 1499 (B)⁸, seleccionado también por los transcritores del texto en Eusko Ikaskuntza⁹, por ser copia autenticada por el bachiller Juan de Herrera, alcalde de Bilbao, y rubricada por dos escribanos.

Una vez establecido el documento base, se realizó el cotejo del resto de las copias conservadas en el Archivo Municipal de Bilbao, un total de ocho documentos. La mayor parte de estas copias se conserva como parte de la documentación presentada en procesos de pleitos sobre jurisdicciones de la villa de Bilbao, por ejemplo con la anteiglesia de Begoña, de Abando o simplemente forma parte del conjunto textual del *Libro Copiador de Privilegios*.

De estos documentos, se prescindió del estudio del manuscrito C, por ser copia incompleta y mal datada¹⁰, y del G, cuya mala conservación impide su lectura¹¹. Los documentos cotejados, de cronología extendida de 1543 al siglo XVIII, se encuentran todos ellos rubricados por escribano, y en su mayor parte, en buen estado de conservación, salvo el documento J, que presenta la tinta desvaída, lo que dificulta en gran medida la lectura.

7. A-Original de 1372 en pergamino (555x549mm); letra gótica; mala conservación, ilegible; Bilbao Sección Pergaminos 0029.

8. B-Copia judicial autenticada de 1499 (290x210mm), rubricada por Juan Sánchez de Balparda y Fortún Martínez de Bilbao la Vieja; letra cortesana, disposición en líneas y sin columnas; buena conservación; Bilbao Sección Antigua 0001/002/010/03 (fol. 3^o-5^o).

9. Enríquez Fernández, J.; Hidalgo de Cisneros, C.; Martínez Lahidalga, A.: *Colección documental del Archivo Histórico de Bilbao*, San Sebastián: Eusko Ikaskuntza, v. I, 1999, pp. 1300-1473.

10. C-Copia incompleta y mal datada (290x200mm), por Pedro Martínez de Bilbao la Vieja; letra corteana, buena conservación; Bilbao Sección Antigua 0304/001/014.

11. G-Copia inserta en pleito de 1660 (315x200mm), letra humanística; mala conservación, ilegible; Bilbao Sección Antigua 0233/002/001.

Presentamos una relación de los manuscritos empleados en el proyecto:

D-Copia inserta en ejecutoria de 1543 (310x210mm) por Juan de la Cuadra; letra procesal, disposición en líneas y sin columnas; buena conservación, sin encuadernar; Bilbao Sección Antigua 0050/001/001 (fol. 26v^o-31v^o).

E-Copia en el Libro Copiador de Privilegios de 1562 (415x270mm), rubricada por Diego de Olábarri; letra cortesana, disposición en líneas y sin columnas; buena conservación; Bilbao Sección Antigua 0199/001/024 (fol. 40r^o-41r^o).

F-Copia inserta en ejecutoria de 1570 (295x200mm) por Benito Lopez de Basarte; letra humanística, disposición en líneas y sin columnas; buena conservación, hojas sueltas, necesidad de cosido; Bilbao Sección Antigua 0038/001/011 (fol. 36v^o-41v^o). Reconoce que copia de B.

H-Copia de 1661 (315x210mm), rubricada por Pedro de Bascaran; letra humanística; buena conservación, disposición en líneas y sin columnas; Bilbao Sección Antigua 0001/002/006 (fol. 1r^o-6v^o).

I-Copia inserta en ejecutoria de 1575 pero en copia simple del siglo XVIII (315x205mm); letra humanística, disposición en líneas y sin columnas; buena conservación; Bilbao Sección Antigua 0041/001/001 (fol. 121r^o-140v^o). Nombre de escribano ilegible por la tinta.

J-Copia inserta en ejecutoria de 1736 (295x200mm) por Jeronimo de Espinar Mendiola; letra humanística, disposición en líneas y sin columnas; regular conservación, tinta desvaída o corrida, sin encuadernar; Bilbao Sección Antigua 0061/001/001 (fol. 117r^o-137v^o).

4.2. La edición

La primera fase del proyecto, tras la selección y confirmación de la disponibilidad de los textos, consistió en llevar a cabo la digitalización de los mismos. En primer lugar, se procedió a la memorización informática de las seis copias a la vez que se incluían las etiquetas XML indicadas por la propuesta TEI y se llevaba a cabo la preparación del documento para su futuro tratamiento y análisis digital. Se adoptó el estándar de codificación ISO-8859-1¹²; se reemplazaron los signos σ y $\&$ por sus correspondientes entidades ($\&\#963$; y $\&\#38$), necesarias para el futuro procesamiento informático del texto; y se definió la estructura textual del documento en función del estándar de marcación elegido.

De forma paralela a la digitalización de los textos, se procedió a la selección de determinados criterios de edición, todos ellos dependientes del formato digital y con intención conservadora en los aspectos lingüísticos conflictivos y modernizadora en cuestiones como las mayúsculas o la puntuación¹³.

12. Norma de la ISO (Organización Internacional para la Estandarización) que define la codificación del alfabeto latino, incluyendo los diacríticos (como letras acentuadas, ñ, ç) y letras especiales (como ß, Ø), necesarios para la escritura de las siguientes lenguas originarias de Europa occidental (Información tomada de la Wikipedia -<http://es.wikipedia.org>).

13. Tanto los criterios de edición como otras decisiones tomadas, se especifican y desarrollan por extenso en la memoria entregada en Sociedad de Estudios Vascos-Eusko Ikaskuntza en marzo de 2006.

4.3. El etiquetado

Como ya se ha mencionado, en el proceso de memorización del texto se insertaron una serie de anotaciones filológicas por medio del lenguaje de etiquetado XML a través del estándar TEI.

4.3.1. Descripción documental

Se dotó a cada documento de una *cabecera* (*teiHeader*) con una completa descripción bibliográfica del documento. Cada una de estas cabeceras incluye:

- descripción de la edición digital dentro del proyecto en el que se inserta, con datos como el título, autor, editor, investigadores, entidades responsables y financiadoras, características de la edición, extensión, características de la publicación, anotaciones complementarias y descripción del texto fuente.
- especificación de algunos aspectos de la codificación tales como: descripción del proyecto, criterios editoriales y otros referidos a la creación y organización del corpus digital y su propio etiquetado.
- descripción del archivo digital: fecha de creación, lengua del texto y definición de la tipología textual por medio de palabras clave.

4.3.2. Segmentación para el cotejo y alineamiento: los corpora paralelos

Con la intención de realizar después una anotación lingüística, el texto fue segmentado en el nivel de la oración. A cada una de estas unidades, consideradas en función de un criterio tipográfico (existencia de punto (.), punto y coma (;) o dos puntos (:)) en la edición de Eusko Ikaskuntza, 1999), se le asignó un identificador que se utilizó en un segundo momento para establecer las correspondencias entre cada una de las oraciones de los distintos textos y su equivalente en el texto tomado como base, la copia B de 1499. De este modo, se generó un corpus similar al de los *corpora* paralelos. Este tipo de *corpora* se componen de un texto y sus traducciones a una o varias lenguas, son, por tanto, documentos multilingües en los que la unidad de alineamiento la suele constituir el párrafo u oración. Sus características y filiación son equiparables a las que definen el diastema de un documento, formado por diversas versiones de un mismo texto en las que se encuentran más o menos variantes, en una misma lengua, pero en distintos documentos. Por ello establecemos aquí una correspondencia entre nuestro corpus textual y los *corpora* alineados.

4.3.3. Segmentación para el análisis

Se adoptaron en este punto las pautas de etiquetado que deben aplicarse a un texto al que se quiere añadir información lingüística contenida en la sección

de TEI dedicada a los métodos de análisis lingüístico, *Simple Analytic Mechanisms* (capítulo 15)¹⁴.

En primer lugar, se anotó el segmento o unidad lingüística (15.1 *Linguistic Segment Categories*), a la que, a continuación, se asignó su correspondiente interpretación lingüística (15.4 *Linguistic Annotation*).

Así pues, se emplearon las siguientes unidades recogidas en la propuesta TEI:

- Grafías o caracteres: *c*
- Morfemas: *m*
- Palabras (secuencia de caracteres rodeada por dos espacios en blanco): *w*
- Sintagmas o locuciones: *ph*
- Cláusulas u oraciones (no entramos en la polémica distinción de ambas, denominaremos así a la unidad mayor que el sintagma): *cl*

La atribución de interpretaciones lingüísticas se llevó a cabo a través de marcas especiales que se situaron fuera del texto mismo, al final del documento, manteniendo, de este modo, una de las máximas de Leech (1993) que indica que “it should be possible to extract the annotations by themselves from the text”.

4.4. Análisis lingüístico

Concluidas estas tareas, se anotaron una serie de fenómenos lingüísticos pertenecientes a distintos niveles de lengua (gráfico-fonético, morfológico, morfosintáctico, sintáctico y pragmático)¹⁵, algunos de los cuales han sido objeto de estudio por parte de los miembros del grupo SAI. Cabe citar, entre otros, la secuencia de consonante velar + <oa> para la representación gráfica de <gua, gua>, aspecto gráfico coincidente con el área navarra (Isasi 1998, 2002; Romero 2002), el uso de *so* para el masculino y femenino y el uso antepuesto de *mio* y *mi*, manifestaciones de un hipotético sistema de posesivos propio, por el momento, del área vizcaína (Isasi 1994), o los casos de leísmo singular y plural (Gómez Seibane 2006a).

Otros elementos etiquetados fueron los tradicionalmente atribuidos al área norteña, como el tratamiento gráfico de la /f/ inicial latina como <h> o como <O>, y los considerados frecuentes y generales en castellano medieval y tardo-

14. Se menciona aquí este proceso ya que se trata de una consideración metodológica de gran utilidad para entender la tarea del enriquecimiento textual.

15. Hay que señalar que no se han anotado todas las diferencias lingüísticas existentes entre el texto base y las copias, aunque no se descarta abordar en trabajos futuros aspectos como el seseo, la combinación de artículo y posesivo, las fórmulas de tratamiento o las preposiciones y locuciones prepositivas.

medieval, como los fenómenos morfofonéticos que afectan a la categoría verbal o las variantes de los latinos ONDE, UBI e IBI.

En una segunda fase, actualmente en desarrollo, los elementos lingüísticos anotados serán procesados por medio de una herramienta informática que permita realizar, de forma automatizada, tareas como las siguientes:

- revisión de los fenómenos concretos que engloban cada una de las clasificaciones establecidas (en función de niveles lingüísticos y cambios anotados) para la propuesta de una clasificación más exhaustiva en caso de ser necesario,
- revisión de los criterios empleados en la clasificación de unidades y fenómenos,
- cotejo de la relevancia y número de fenómenos etiquetados dentro de cada una de las divisiones planteadas,
- extracción de las primeras conclusiones filológicas en torno a los resultados aportados por la herramienta.

5. REFLEXIÓN FINAL

A lo largo de este artículo, se han abordado de forma somera aspectos como la aplicación de las TIC a la recuperación del patrimonio documental, el tratamiento textual automatizado o el análisis lingüístico del romance vizcaíno a través de un ensayo metodológico de perspectiva interdisciplinar.

Si bien la fase actual de desarrollo del proyecto aún no nos permite adelantar resultados lingüísticos, conviene tener en cuenta que nuestro interés no sólo es analizar un estado de lengua, sino también presentar un método de trabajo que aúne la conservación del patrimonio documental y su tratamiento automatizado para el análisis lingüístico.

6. BIBLIOGRAFÍA

ABAITUA, Joseba. "Tratamiento de corpora bilingües". En: M. A. Martí y J. Llisterri, *Tratamiento del lenguaje natural*. Barcelona: Edicions Universitat de Barcelona, 2002; pp. 61-90.

(Artículo consultado en la página web del autor: <<http://paginaspersonales.deusto.es/-abaitua/konzeptu/ta/soria00.htm>> [julio 2006]).

ABAITUA, Joseba *et alii*. "Contenidos y metacontenidos en la edición digital". En: *La edición digital. Letras de Deusto*, nº 100, Julio-Septiembre, 2003; pp. 11-52.

ABAITUA, Joseba; ISASI, Carmen (coords.). *La edición digital*, Monográfico de Letras de Deusto, n.º 100, Bilbao: Universidad de Deusto, 2003.

AMORRORTU, Estibaliz; FERNÁNDEZ, Patricia; ISASI, Carmen; PÉREZ ISASI, Santiago. "Autoitzulpenaren zenbait aspektu Unai Elorriagaren SPraiko tranbia-Un tranvia en SP

Fernández, P; Gómez, S.: La edición enriquecida y en paralelo para el estudio del patrimonio...

lanean: XML azterketa literarioari aplikatuta”, Non: *Unai Elorriagaren kontagintza*, Donostia: Eusko Ikaskuntza, 2006; p. 83-126.

BLECUA, Alberto. *Manual de crítica textual*. Madrid: Castalia, 1990.

BLECUA, J. M.; CLAVERÍA, G.; SÁNCHEZ, C.; TORRUELLA, J. (eds.). *Filología e informática. Nuevas tecnologías en los estudios filológicos*. Bellaterra: Editorial Milenio et Universitat Autònoma de Barcelona, 1999; pp. 189-210.

CIOTTI, F. “La codifica del testo. XML e la Text Encoding Initiative”. En Gigliozzi, G. *Introduzione all'uso del computer negli studi letterari*, (edición de F. Ciotti). Milán: Bruno Mondadori, 2005.

CIULA, A. “Digital palaeography: using the digital representation of medieval script to support palaeographic analysis”. En *Digital Medievalist* 1.1, <<http://www.rdigitalmedievalist.org/article.cfm?ReclD=2#robinson2004>, 2005.> [julio 2006].

CIVIT TORRUELLA, M. *Criterios de etiquetación y desambiguación morfosintáctica de corpus en español*. Madrid: SEPAN, 2003.

ESCOLAR FRANCO, A. “La edición filológica en la empresa privada. Pasado y futuro”. En: *II Congreso Internacional de la Lengua Española*, Valladolid, 2001. <http://cvc.cervantes.es/obref/congresos/valladolid/ponencias/activo_del_espanol/2_la_edicion_en_espanol/escolar_a.htm> [julio 2006].

FERNÁNDEZ, Patricia. *Preparación de una edición digital bilingüe de Il Romulo del marqués Virgilio Malvezzi traducido por Teodoro dell'Aula para su estudio sintáctico-retórico*. Bilbao: Universidad de Deusto, 2004 (Tesina inédita).

FIORMONTE, Domenico. *Scrittura e filologia nell'era digitale*. Torino: Bollati Boringhieri, 2003.

—. *Informatica Umanistica. Dalla ricerca all'insegnamento. Atti dei convegni “Computers, Literature and Philology”*, Università di Roma “La Sapienza” (1999) e Università di Alicante, Spagna (2000), con la collab. di Giulia Buccini. Roma: Bulzoni, 2003.

GANCEDO NEGRETE, Marisol. “Abreviaturas y signos abreviativos en la documentación original del siglo XV del Archivo Municipal de Durango”. En: C. Isasi Martínez y S. Gómez Seibane (eds.), *Lingüística Vasco-Románica / Euskal-Erromantze Lingüistika*. Donostia: Eusko Ikaskuntza, 2006; pp. 101-116.

GÓMEZ SEIBANE, Sara. “Uso de los pronombres átonos de tercera persona en la Vizcaya del siglo XV: Bilbao y Lekeitio”, *Diacronía, lengua española y lingüística*. Madrid: Síntesis, 2006a; pp. 575-587.

—. «Análisis del registro coloquial de un expediente de Portugalete (1708)», En: C. Isasi Martínez y S. Gómez Seibane (eds.), *Lingüística Vasco-Románica / Euskal-Erromantze Lingüistika*. Donostia: Eusko Ikaskuntza, 2006b; pp. 123-139.

ISASI, Carmen. “Nueva aproximación a la documentación vizcaína medieval: una nota de morfología”. En: *Letras de Deusto*, 64. 1994; pp. 215-22.

—. “Para la historia del castellano de Bilbao”. En: C. García Turza, F. González, J. Manguado (eds.), *Actas del IV Congreso Internacional de Historia de la Lengua Española*, Logroño: Universidad de La Rioja, II, 1998; pp. 741-49.

—. “Los documentos notariales: entre el formulismo y la innovación”. En: E. Artaza (ed. lit.), *Estudios de Filología y Retórica en Homenaje a Luisa López Grigera*. Bilbao: Universidad de Deusto, 2000; pp. 281-94.

- . “Castellano y euskera en los documentos de Bilbao”. En: A. Arejita; A. Elejabeitia; C. Isasi; J. Otaegí (eds.), *Bilbao. El espacio lingüístico. Simposio 700 Aniversario*. Bilbao: Universidad de Deusto, 2002a; pp. 135-52.
 - . “Documentos navarros y documentos vizcaínos: algunas coincidencias”. En: M^a T. Echenique; J. Sánchez (eds.), *Actas del V Congreso Internacional de Historia de la Lengua Española*. Madrid: Gredos, II, 2002b; pp. 1541-52.
 - . “Criterios filológicos y edición electrónica”. En: *Lamusa Digital*, 3. 2003. <<http://www.uclm.es/lamusa/>> [julio 2006].
 - . “Seseo vizcaíno en documentos del siglo XVIII”. En: J. J. de Bustos Tovar / J. L. Girón Alconchel (eds.): *Actas del VI Congreso Internacional de Historia de la Lengua*. Madrid: Arco Libros, 2006a; pp. 2462-2471.
 - . “A vueltas con el seseo vizcaíno y otras cuestiones de fonética y morfología históricas en documentos de Bizkaia”. En: J. Lakarra; J. I. Hualde (eds.), *Studies in Basque and Historical Linguistics in Memory of R. L. Trask. Anuario Seminario de Filología Vasca Julio de Urquijo*, 40. 2006b; pp. 531-546.
 - . “El romance de la documentación vizcaína en el espacio variacional castellano”. En: C. Isasi Martínez y S. Gómez Seibane (eds.), *Lingüística Vasco-Románica / Euskal-Erromantze Lingüistika*. Donostia: Eusko Ikaskuntza, 2006c; 209-227.
- ISASI, Carmen; ÁLVAREZ, David; GANCEDO, Soledad; GÓMEZ, Josu; GÓMEZ, Sara; RAMÍREZ LUENGO, José Luis; ROMERO, Asier. “Léxico vizcaíno medieval”. En: *Oihenart. Cuadernos de Sección. Lengua y literatura*, nº 20. Donostia: Eusko Ikaskuntza, 2005; 73-201.
- LEECH, Geoffrey. “Corpus annotation schemes”, *Literary and Linguistic Computing. Special Section on Corpora*, 8. 1993; 275-282.
- LUCÍA MEGÍAS, J. M. *Manual de informática humanística*. Madrid: Castalia, 2005.
- MARTÍ, M. A. (coord). *Tecnologías del lenguaje*. Barcelona: Ed. UOC, 2003.
- RAMÍREZ LUENGO, J. Luis. “Casos de indicativo por subjuntivo en la Vizcaya bajomedieval”. En: J. D. Luque Durán et alii (eds.), *Nuevas tendencias en la Investigación Lingüística*. Granada: Granada Lingüística, 2001; pp. 241-50.
- . “Notas sobre el futuro de subjuntivo en la Vizcaya bajomedieval”. En: M. Villayandre Llamazares (ed.): *Actas del V Congreso de Lingüística General*, 3. Madrid: Arco Libros, 2002; pp. 2322-2329.
- ROMERO ANDONEGUI, Asier. “Notas gráfico-fonéticas en documentación testamentaria de Bilbao”. En: *Res Diachronicae* 1. 2002; pp. 316-23.
- SAI (Seminario Alfonso Irigoyen) <<http://littera.deusto.es/sai>> [julio 2006].
- TEI (Text Encoding Initiative) <<http://www.tei-c.org>> [julio 2006].
- TEUBERT, W. “Comparable or parallel corpora?”. En: *Corpus to corpus: a study of translation equivalence: International Journal of lexicography*. London: Oxford University Press, 1996.