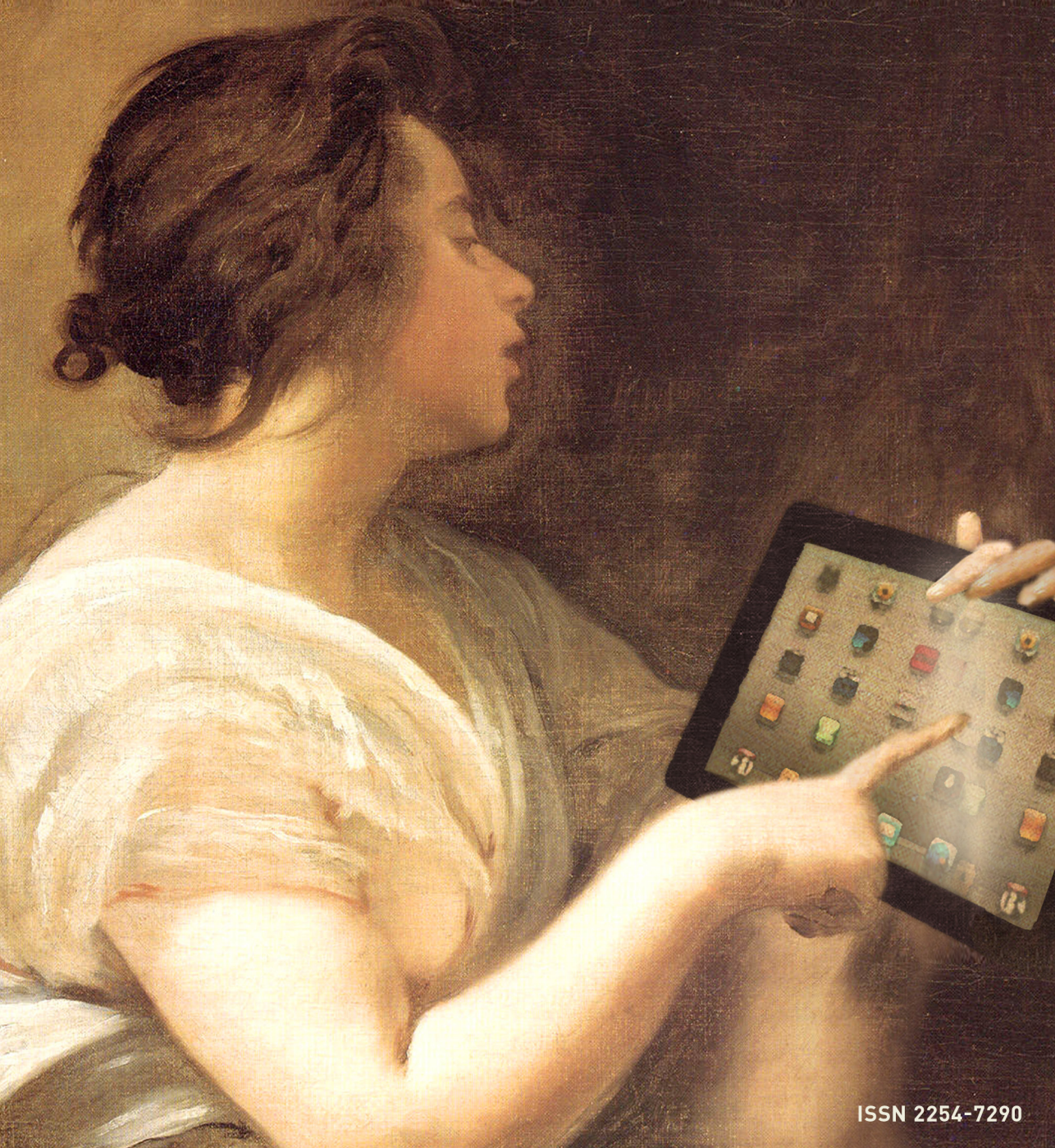




HUMANIDADES DIGITALES: DESAFÍOS, LOGROS Y PERSPECTIVAS DE FUTURO

SAGRARIO LÓPEZ POZA Y NIEVES PENA SUEIRO (EDITORAS)



Humanidades digitales: desafíos, logros y perspectivas de futuro



Sagrario López Poza
y Nieves Pena Sueiro
(editoras)

Janus. Anexo 1 (2014)

ISSN 2254-7290

Diseño de la cubierta: Paula Lupiáñez López

Edita:

SIELAE (Seminario Interdisciplinar para el Estudio de la Literatura Áurea Española).
Facultad de Filología. Universidade da Coruña. A Coruña (España)

© 2014 Los autores (cada uno de su trabajo)

© De esta edición: SIELAE-JANUS

Reservados todos los derechos.

SIELAE

<http://www.bidiso.es/sielae>

Maquetación: Dosgraphic, s.l.

ÍNDICE

Sagrario López Poza	
Presentación	7
José Luis Canet	
Reflexiones sobre las humanidades digitales	11
John Griffiths	
Musicología, informática y la vihuela en el siglo XXI	21
Paul Spence	
Centros y fronteras: el panorama internacional de las humanidades digitales..	37
	
Diego Agudelo Grajales	
Proyectos sociales y de aula a partir del aprendizaje colaborativo virtual. Experiencias desde la tutoría en Diplomado de AUSJAL.....	63
Montserrat Amores	
El buscador de GICES XIX, herramienta digital sobre el cuento español del siglo XIX	79
Gonzalo A. Aranda Corral, Joaquín Borrego Díaz y Juan Galán Páez	
Extracción y organización del conocimiento de etiquetados. Aplicación a etiquetados en repositorios digitales sobre arte	87
Consolación Baranda y Eduardo Rodríguez	
Red ARACNE: retos y objetivos de un proyecto de coordinación en letras hispánicas digitales.....	101
Gidor Bilbao y Ricardo Gómez	
Textos antiguos vascos en Internet.....	111
Jean-Marc Buiguès	
Indexación y códigos en bases de datos bibliográficas.....	123

Paloma Centenera Centenera y María Luisa Díez Platas

El proceso de desarrollo de una biblioteca digital y su apertura e integración en el recolector *Hispana*: el caso de la *Biblioteca Digital Ovidiana Abierta*..... 137

Daniel Escandell Montiel

Tecnoestética de la blogoficción: imitación y suplantación del bloguero en la narrativa del avatar 151

Carme Fernández Pérez-Sanjulián, M^a Antonia Pérez Rodríguez, Eduardo Rodríguez López y Ánxeles Saavedra Places

Recursos para a clasificación da produción editorial na Galiza durante a etapa franquista: diseño e alimentación da base de datos 161

Carlota Fernández Travieso

Base de datos sobre Inventarios y Bibliotecas del Siglo de Oro (IBSO). Utilidad y posibilidades 175

Carolina Ferrer

El canon literario hispanoamericano en la era digital 185

Juana María González García

El proyecto *Epístola*: edición digital de los epistolarios de la Edad de Plata 197

Elena González-Blanco García, Clara Isabel Martínez Cantón, María Dolores Martos Pérez y María Gimena del Río Riande

Una propuesta de integración del sistema de formularios de bases de datos MYSQL con etiquetado TEI: ReMetCa, Repertorio digital de la métrica medieval castellana..... 209

M^a Isabel Hidalgo Urbaneja

Análisis crítico de los discursos y narrativas digitales. *The Gallery of Lost Art* –Tate– como estudio de caso 221

M^a Eugenia Iglesias Moreno, Pilar Azcárate Aguilar-Amat y Sonia Sánchez Cuadrado

Primera aproximación para la extracción automática de *Entidades Nombradas* en corpus de documentos medievales castellanos 229

Xurxo Insua, Mireia Freixa y Teresa-M. Sala

Experiencias de e-research y divulgación digital en historia del arte: GracmonDocs y Viquimodernisme 239

Raquel Jimeno Revilla

Los clubes del libro en la era digital: nuevas estrategias y aspectos sociales. El caso de Círculo de Lectores (1998-2012)..... 247

Sonia Lavandeira Rojo

Publicidad digital..... 257

Laura Llanos Casado y Milka Villayandre Llamazares

Nuevas tecnologías y escritura académica 263

Álvaro Llosa SanzDe cómo Ramón Gómez de la Serna se convirtió en un fan de Twitter:
un ejemplo de textualidad digital..... 277**Sagrario López Poza y Ángeles Saavedra Places**Recursos digitales ofrecidos por el SIELAE para el estudio del Siglo de Oro.
Creación, gestión y evolución de BIDISO..... 285**Ana Mancera Rueda y Ana Pano Alamán**Las redes sociales como corpus de estudio para el Análisis del discurso
mediado por ordenador 305**Patricia Meilán Jácome y Fátima Díez Platas**La *Biblioteca Digital Ovidiana* y el patrimonio bibliográfico español 317**Isa Bernardino Mestre**Literatura digital o la reinención de la lectura: desde la poética
de la obra abierta hacia la hipermedia..... 329**Nieves Pena Sueiro y Sandra Álvarez García**El *Catálogo y Biblioteca digital de relaciones de sucesos*: bases de datos
bibliográficas, textos e imágenes..... 335**Santiago Pérez Isasi y José Ramón Carriazo Ruiz**Ensayo de un índice conceptual para un corpus literario de textos
alineados multilingües 347**Mónica Poza Diéguez**The Littera Project: *Webpapers* o trabajos académicos 2.0 361**Ernesto Priani Saisó y Ana María Guzmán Olmos**

TEI como una nueva práctica de lectura..... 373

Begoña Regueiro Salgado

Innovación y creatividad: enseñar literatura en el siglo XXI..... 383

María José Rodríguez-Espiñeira y Yolanda López ArcaEl proyecto MORFOGEN. Diseño de una aplicación web para visualizar
familias léxicas 395**Dolores Romero López**Hacia la *Smartlibrary: Mnemosine*, una biblioteca digital de textos literarios
raros y olvidados de la Edad de Plata (1868-1936). Fase I..... 411**Laura Sánchez Gómez**

Poéticas Digitales: el texto como imagen de una obra cinética..... 423

**Margarita Santos Zas (coord.), Francisca Martínez Rodríguez
y Rosario Mascato Rey**

La creación y gestión del archivo digital valleincliniano: *corpus manuscrito
e impreso* 435

Sabine Thuillier

De una obra digitalizada a una obra digital:
reflexiones en torno al *Diccionario Griego-Español* en línea 459

**Gael Vaamonde, Ana Luísa Costa, Rita Marquilhas, Clara Pinto
y Fernanda Pratas**

Post Scriptum: Archivo Digital de Escritura Cotidiana 473

Textos antiguos vascos en Internet¹

GIDOR BILBAO

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)
gidor.bilbao@ehu.es

RICARDO GÓMEZ

Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)
ricardo.gomez@ehu.es

1. TEXTOS ANTIGUOS VASCOS

Conviene comenzar este trabajo definiendo lo que entendemos como *texto antiguo vasco* y recordando algunas características del corpus al que nos referimos.

La histórica y aún vigente recopilación de textos vascos de Luis Michelena, publicada en 1964, recogía textos breves escritos en lengua vasca antes del año 1700, y se llamó *Textos Arcaicos Vascos*. Actualmente, en los estudios de historia de la lengua vasca se acepta comúnmente que el año 1745, fecha de publicación del *Diccionario Trilingüe del castellano, bascuence y latín* de Manuel Larramendi, marca el final del euskera antiguo y el comienzo del llamado "primer euskera moderno" (Lakarra, 1997: 516). La periodización propuesta por Lakarra cuenta hoy en día con una amplia aceptación y es la utilizada en la monografía más importante sobre la historia del euskera (Gorrochategui, Igartua y Lakarra, 2013). Por ello hemos preferido utilizar el término *texto antiguo vasco*, refiriéndonos así a todo texto, no necesariamente literario, de diversa extensión, escrito en lengua vasca antes de 1745.

Para el investigador menos familiarizado con los textos vascos, quizás no resulta superfluo recordar aquí que, a pesar de la amplitud del período abordado, el corpus es muy reducido. Sirva como referencia un estudio de Ibon Sarasola sobre la producción editorial vasca en el período antiguo (1976: 177-183), que, a pesar del tiempo transcurrido desde su publicación, aún sigue siendo significativo; recogemos en una tabla de elaboración propia algunos datos que nos parecen relevantes, referidos en este caso a publicaciones de más de 48 páginas:

¹ Los autores han contado con la ayuda de estos proyectos y grupos de investigación: "Monumenta Linguae Vasconum (IV): textos arcaicos vascos y euskera antiguo" (MINECO, FFI2012-37696), "Historia de la lengua vasca y lingüística histórico-comparada" (HLMV-LHC) (Gobierno Vasco, GIC. IT698-13) y "Lingüística Teórica y Diacrónica: Gramática Universal, lenguas indoeuropeas y lengua vasca" (LingTeDi) (UPV/EHU, UFI11/14).

Período	Publicación total	Producción total (sin reediciones)	Producción original (sin traducciones)	Producción original NO religiosa
1545-1599	3	3	1	1
1600-1649	11	7	6	–
1650-1699	25	13	6	1
1700-1749	34	17	5	1

Actualmente no disponemos de datos exactos, con número de caracteres, del corpus que nos ocupa atendiendo a la definición que hemos adelantado más arriba; pero los datos referidos a la producción editorial son perfectamente extrapolables, de modo que podemos afirmar que los textos en lengua vasca de la Antigüedad y de la Edad Media son escasísimos y se reducen a nombres propios de persona, de divinidades o de lugar, a los que se unen, a partir de la Edad Media, glosas, vocabularios y frases, expresiones o palabras insertas en documentos latinos y romances. En los siglos XVI y XVII, encontramos esta misma tipología, aunque en mayor cantidad, pero aparecen también algunas declaraciones en procesos judiciales, cantares o fragmentos de cantares antiguos, poesías sueltas, colecciones de refranes y sentencias, métodos para el aprendizaje de la lengua vasca, y las primeras obras de cierta extensión (el primer libro impreso en euskera, *Linguae Vasconum Primitiae* de Bernard Etxepare, data de 1545). La progresión es notable a partir de principios del siglo XVIII, pero siempre en cantidades tan ínfimas que cualquier texto escrito en lengua vasca antes de 1745, por pequeño y aparentemente insignificante que pueda parecer, es importante para la historia de la lengua vasca.

2. EL ACCESO A REPRODUCCIONES DE ORIGINALES: BIBLIOTECAS DIGITALES

Los facsímiles de textos antiguos vascos, tanto manuscritos como impresos, accesibles desde internet son, como sucede con otras lenguas del entorno, cada vez más numerosos.

Las bibliotecas digitales y los proyectos internacionales de tipo generalista son, por supuesto, útiles para este fin. Por ejemplo, gracias al proyecto *Europeana*², podemos acceder en menos de un minuto, a través de *Gallica*³, a una reproducción de calidad excelente del *unicum* de la primera edición de la obra de Bernard Etxepare (1545) que hemos mencionado más arriba, para cuya consulta debíamos acudir hasta hace bien poco a París, a la Biblioteca Nacional de Francia. Desde esa misma plataforma *Europeana*, necesitaremos un minuto más para poder consultar otra reproducción ejemplar del manuscrito *Grammaire cantabrique* de Pierre d'Urte (c. 1714), disponible en los fondos de la Fundación Sancho el Sabio de Vitoria-Gasteiz, a través de su proyecto *Memoria Digital Vasca*⁴, integrado en *Hispana*⁵. Otros proyectos como Google Books o Internet Archive nos ayudan a diario en nuestras búsquedas.

² <<http://www.europeana.eu/>>.

³ <<http://www.gallica.bnf.fr/>>.

⁴ <<http://www.memoriadigitalvasca.es/>>.

⁵ <<http://hispana.mcu.es/>>.

También las bibliotecas digitales locales y especializadas en textos vascos han apostado en los últimos años por permitir el acceso digital a facsímiles de ediciones impresas y manuscritos de textos antiguos vascos. Por razones de espacio, sólo podemos mencionar las más importantes⁶:

- **Liburuklik – Biblioteca Digital Vasca**⁷: reúne y facilita el acceso a parte de los recursos desarrollados en la Comunidad Autónoma del País Vasco, entendiendo el patrimonio bibliográfico vasco en su sentido más amplio (patrimonio bibliográfico existente en bibliotecas de la Comunidad Autónoma Vasca, independientemente de su origen o su lengua). Ha integrado fondos tan importantes como los Fondos Digitales del Parlamento Vasco (en su totalidad), la biblioteca del Santuario de Loyola, Eresbil (archivo de la música vasca), o la Fundación Sancho el Sabio, entre otros, pero aún no ha conseguido convertirse en referente único, puesto que hay instituciones importantes que no se han integrado en el proyecto (es el caso del Repositorio Digital de la Diputación Foral de Gipuzkoa), e incluso las instituciones que participan lo hacen, en algunos casos, con parte de sus fondos, y no con la totalidad de su patrimonio catalogado, de modo que el investigador se ve obligado a repetir la búsqueda en el catálogo de la propia institución o incluso en otros catálogos más generales.
- **Memoria Digital Vasca**⁸: La Fundación Sancho el Sabio, especializada en estudios vascos, fue una de las instituciones pioneras en la publicación digital de reproducciones de textos antiguos vascos. Se ha integrado en *Liburuklik* y en *Hispana*, pero mantiene su catálogo exclusivo, que sigue siendo la única fuente de información totalmente segura para acceder a la totalidad de sus fondos digitales.
- **Repositorio Digital de la Diputación Foral de Gipuzkoa**⁹: Es una institución pública local que conserva, entre otros, el Fondo Julio Urquijo, de modo que incluye un rico patrimonio de textos vascos, impresos y manuscritos.
- **Biblioteca Navarra Digital**¹⁰: Tiene como finalidad difundir el patrimonio bibliográfico de Navarra a través de las nuevas tecnologías de la información, de tal forma que incluye las reproducciones de varios ejemplares de textos antiguos que difícilmente podrían ser consultados de otra manera.
- **Biblioteca Foral de Bizkaia. Biblioteca Digital**¹¹: La Diputación Foral de Bizkaia es heredera de diversos legados públicos y privados que incluyen textos antiguos vascos impresos y manuscritos, cuyos facsímiles están siendo publicados en internet paulatinamente en las colecciones "Manuscritos vascos" y "Fondo antiguo vasco", fundamentalmente.
- **Biblioteca Azkue de Euskaltzaindia**¹²: La biblioteca de la Academia de la Lengua Vasca es depositaria, entre otros, de los fondos de la Fundación

⁶ Dejamos al margen los sistemas de consulta de la documentación de los archivos (valga como ejemplo el sistema *Badator* de la Comunidad Autónoma de Euskadi [<http://dokuklik.snae.org/>], cuya labor de los últimos años hace accesibles desde internet documentos que contienen textos antiguos vascos inéditos [Bilbao, 2013]).

⁷ <http://www.liburuklik.euskadi.net/>.

⁸ <http://www.memoriadigitalvasca.es/>.

⁹ <http://meta.gipuzkoakultura.net/>.

¹⁰ <https://administracionelectronica.navarra.es/binadi/busqueda.aspx>.

¹¹ <http://bibliotecaforal.bizkaia.net/>.

¹² <http://gorbeia.euskaltzaindia.net/cgi-bin/abwebp/LLeu-ES/G0/>.

Olaso Dorrea, una de las colecciones más ricas de textos vascos antiguos. Se puede consultar el catálogo en internet, y las reproducciones pueden ser consultadas en CD-ROM, en la sala de la biblioteca.

Finalmente, hay reproducciones de originales de textos vascos disponibles en bibliotecas digitales aisladas, en las cuales el estudioso difícilmente buscaría si no es porque tiene conocimiento fehaciente de su existencia. Es el caso, por ejemplo, de los manuscritos del siglo XVII, conocidos como *Basque-Icelandic Glossaries*, que han circulado entre los filólogos interesados gracias a unas copias microfilmadas a mediados del siglo XX, de calidad aceptable pero sin punto de comparación con las fotografías que actualmente podemos consultar en nuestro ordenador, en la página web de *The Árni Magnússon Institute for Icelandic Studies*¹³.

Desgraciadamente, el investigador interesado en estudiar textos vascos antiguos debe necesariamente, en cada caso, conocer o indagar en las diferentes posibilidades, puesto que, a pesar de tratarse de un corpus reducido, no existe aún un directorio especializado y exhaustivo que permita localizar fácilmente los cada vez más numerosos facsímiles accesibles desde internet, pero dispersos por diferentes bibliotecas y proyectos de ámbito nacional o internacional, generalistas y especializados.

3. EL ACCESO A LOS TEXTOS: FORTALEZAS Y DEBILIDADES DE DIVERSOS PROYECTOS

3.1. *Euskal Testuen Gordailua* [Depósito de textos vascos]

Este portal de textos antiguos vascos¹⁴ se encuentra alojado en la web de la UPV/EHU y su responsable es Patxi Salaberri Muñoa, profesor de la misma institución. El portal tiene como objetivo reunir en un único repositorio versiones electrónicas de textos vascos de diversa procedencia, o bien confeccionadas *ad hoc* a partir de distintas fuentes impresas. Hace años que no se renuevan sus contenidos, siendo la última actualización de 29 de diciembre de 2000.

El portal cuenta con tres secciones:

1. "Idazleak eta idazlanak" [Autores y obras]: son cincuenta y tres obras escritas desde el s. XVI a la primera mitad del XX, organizadas alfabéticamente por autor. Gran parte de los "clásicos" que incluye fueron proporcionados por Josu Lavín, el auténtico precursor en la digitalización de textos vascos.
2. "Eresiak, bertsoak eta kantak" [Elegías, versos y cantos]: se trata de versos antiguos, en su mayor parte tomados de las colecciones publicadas por Luis Michelena (1964) e Ibon Sarasola (1983).
3. "Kutxatila" [Arqueta]: frases sueltas, cartas y otros textos breves, casi todos tomados también de Michelena (1964) y Sarasola (1983).

Junto a cada texto se menciona la versión (o edición) publicada en papel utilizada (bajo el epígrafe "Bertsio honen iturria") o, en su caso, el autor de la

¹³ <http://www.arnastofnun.is/page/basknesk_islensk_ordasofn_en>.

¹⁴ <<http://www.vc.ehu.es/gordailua/>>.

versión electrónica preparada ex profeso para el repositorio y, en ocasiones, la fuente original ("Non"). Se ofrece únicamente la transcripción del texto, sin introducción ni notas de ningún tipo. Por otro lado, no hay un criterio uniforme para la presentación de los textos, conservando las características muy diversas de la fuente impresa o de la versión electrónica utilizada en cada caso.

3.2. Corpus del *Diccionario General Vasco – Orotariko Euskal Hiztegia (OEH)*

Fue un subproducto del *Diccionario General Vasco* de la Academia de la Lengua Vasca, diseñado a partir del corpus elaborado para la confección de dicho diccionario. Fue editado en 2000, en soporte CD-ROM, pero no fue comercializado y tuvo una distribución limitada, como versión de trabajo exclusiva para las diferentes comisiones de la Academia de la Lengua Vasca. No obstante, merece un lugar privilegiado en este trabajo, porque fue pionero en muchos aspectos y buena parte de sus materiales fueron posteriormente utilizados por *Klasikoen Gordailua*.

El CD-ROM incluye 304 documentos, lo cual suponía en el 2000 un volumen más que considerable. Todos los textos incluidos son obras literarias (en su acepción más amplia y generosa, acorde con las limitaciones de la producción en lengua vasca), escritas entre 1545 y 1970, pero quedaron excluidos textos editados por Michelena (1964), y otros textos de pequeño volumen como romances, etcétera.

Este trabajo incorporaba un motor de búsqueda muy rudimentario desde la perspectiva actual, pero novedoso y útil en aquel momento. El investigador puede organizar sus búsquedas con filtros cronológicos, de variedades dialectales o géneros, y puede restringir la búsqueda a una obra concreta. Por supuesto, el buscador no podía incorporar entonces ningún tipo de lematizador, de modo que se buscan palabras enteras (secuencias de caracteres), pudiendo combinar hasta dos palabras, relacionadas con diferentes operadores. En la tabla de resultados obtenidos, que presenta una lista alfabética de palabras, el investigador debe buscar manualmente otras palabras relacionadas con el lema buscado. El motor de búsqueda permite entonces consultar los resultados en su contexto e incluso obtener la obra completa en formato RTF.

Los textos, que como hemos apuntado más arriba, no habían sido preparados para su publicación, sino como material de trabajo para la confección del *Diccionario General Vasco*, son presentados siempre en grafía modernizada y deben ser manejados con cuidado, puesto que la carencia de labor de edición es evidente: la modernización de la grafía no siempre es correcta, no hay notas ni aparato crítico, ni siquiera explicaciones sobre las modificaciones operadas en cada texto. Debemos insistir, sin embargo, en que este trabajo cumplió en su momento una función importante y, usado con sentido crítico, resultó ser una herramienta muy útil para los investigadores.

3.3. *Klasikoen Gordailua* [Depósito de clásicos vascos]

Este segundo repositorio¹⁵ forma parte del más amplio portal *Armiarma* [La tela de araña]¹⁶, dedicado a la literatura vasca, y todos sus contenidos son de

¹⁵ <<http://klasikoak.armiarma.com/>>.

¹⁶ <<http://www.armiarma.com/>>.

dominio público. *Klasikoen Gordailua* es el heredero de *Euskal Testuen Gordailua*, pero supone un gran avance con respecto a su predecesor. Por un lado, se han incorporado nuevos textos, hasta completar el corpus de mayor volumen de textos antiguos vascos (el límite *ante quem* es 1936) accesible hoy en día en internet. Por otra parte, el repositorio dispone de un motor de búsqueda relativamente avanzado, denominado "Corpus arakatzaillea" [El rastreador del corpus]. Permite buscar hasta un máximo de tres palabras (o inicios de palabra) a la vez y establecer la distancia entre ellas. Además, la búsqueda puede restringirse mediante varios filtros: autor/obra, género, época y dialecto. Un problema que plantea este motor de búsqueda es que sólo funciona dentro de la sección de autores y obras, a la que nos referiremos a continuación, pero no en el resto de los textos recogidos en el repositorio.

El repositorio cuenta, en efecto, con tres secciones, que se corresponden, corregidas y muy aumentadas, a las tres de *Euskal Testuen Gordailua* que hemos mencionado en la sección anterior.

En primer lugar, "Idazleak eta idazlanak" [Autores y obras], que constituye el corpus principal del repositorio. Esta sección consta en la actualidad de 496 obras, 11,9 millones de palabras y 73,9 millones de caracteres. Puede descargarse completa en formato TEI (.sgml) o en formato RTF. El listado de obras de esta sección puede consultarse ordenado alfabéticamente por autor, cronológicamente o por géneros. Sin embargo, no todos los textos disponen de la misma cantidad de información. En el mejor de los casos encontraremos la siguiente: (1) El texto en sí, que, a diferencia de *Euskal Testuen Gordailua*, se presenta con la grafía modernizada; el problema es que los textos contienen muchos errores, ya que en ocasiones la modernización parece haberse realizado de forma (semi)automática. (2) En el margen derecho, se sitúan los enlaces a las secciones en las que se divide el texto, si este es de cierta extensión. (3) El nombre del autor enlaza con su posición en el listado alfabético. (4) El año, a su vez, nos sitúa la obra en el listado cronológico. (5) Enlace para descargar la obra completa en formato RTF. (6) Enlace para descargar la obra completa en formato PDF. (7) En algunos casos, enlace para descargar una edición facsimilar en formato PDF. (8) Se complementa, en el caso de los autores más estudiados, con otro servicio incluido en el macro-proyecto *Armiarma*, llamado *Literaturaren zubitegia* [Los puentes de la literatura], que incluye información diversa sobre la obra o el autor, bibliografía, traducciones, un directorio de enlaces a facsímiles de diferentes bibliotecas o a otras páginas web relacionadas, e incluso a textos en formato EPUB. (9) Fuente impresa utilizada y/o autor de la versión electrónica incluida.

La segunda sección lleva por título "Herri literatura" [Literatura popular] y está dividida en cuatro subsecciones: (1) "Baladak, poesiak, kantak" [Baladas, poesías, canciones], compuesta principalmente por baladas y poesías antiguas de los ss. XIV al XVII y algunas colecciones de canciones populares y *bertsos* del s. XIX. (2) "Ipuinak, kondairak" [Cuentos, leyendas], que enlaza con tres obras que aparecen también en la sección de autores y obras. (3) "Antzerkia" [Teatro], integrada por una colección de *astolasterrak*, un tipo de teatro popular en verso propio del País Vasco Continental. (4) "Atsotitzak eta beste" [Refranes y otras cosas], en la que se incluyen ejemplos de refranes pertenecientes a varias colecciones paremiológicas de los ss. XVI y XVII.

La tercera sección es la llamada genéricamente "Testuak" [Textos], dividida a su vez en dos subsecciones: (1) "Testu zaharrak" [Textos antiguos], que contiene

frases sueltas, oraciones, cartas y otros textos breves, tanto literarios como administrativos, de todos los períodos abarcados por el repositorio (es decir, hasta la Guerra Civil). (2) "Epigrafikoak" [Epigráficos], inscripciones de época romana que incluyen algún nombre en euskera, principalmente teónimos.

Estas últimas dos secciones presentan unas características especiales. Primeramente, como hemos señalado más arriba, estas secciones no están incluidas en el motor de búsqueda. En segundo lugar, aparte del texto, la única información complementaria que ofrecen es la fuente impresa ("Bertsio honen iturria") o el autor de la versión electrónica utilizada ("Bertsio informatiko honen egilea") y, en ocasiones, la fuente original ("Non"). En tercer lugar, los textos se presentan sin modernizar, respetando la grafía y la disposición de la fuente impresa u original utilizada. Por último, en cada subsección las obras se ordenan cronológicamente (excepto las obras epigráficas, todas de época romana, que aparecen dispuestas por orden alfabético).

3.4. Portal Andrés de Poza

El Portal Andrés de Poza¹⁷ es un proyecto desarrollado en la Universidad de Deusto, y ha ido creciendo paulatinamente desde 2007, siempre bajo la dirección de la doctora Carmen Isasi. Entre los proyectos aquí expuestos, este es sin duda el más ambicioso en la aplicación combinada de criterios filológicos y estándares de marcación, siguiendo las propuestas de la Text Encoding Initiative (TEI).

Se trata de un corpus reducido de textos múltiples, entendiendo como tales aquellos que permiten la presentación alineada de diferentes versiones o ediciones del mismo texto, o bien diferentes traducciones del mismo texto. Por tanto, el corpus no es exclusivamente (ni siquiera mayoritariamente) en lengua vasca, y de hecho los textos antiguos vascos incluidos son contados: *Platica lelengoá* de Rafael Mikoleta (1653), *San Frances Sales Genevaco Ipizpicuaren Philotea eta Chapeletaren Andre Dana Mariaren ohoretan devocionearequin erraiteco antcea* de Silvain Pouvreau (1664), *Gudu espirituala* de Silvain Pouvreau (1665) y *Philotea edo devocioneraco bide erakusçaillea* de Joannes Haraneder (1749).

Todos ellos comparten la particularidad de ser traducciones y son presentados en versión múltiple alineada, junto con sus originales respectivos o con traducciones de otras lenguas, que pueden ser cotejados fácilmente por pares, tríos y hasta cinco versiones diferentes. Los textos vascos han sido editados *ad hoc* para este proyecto por filólogos de prestigio (Josu Bijuesca y Ana Toledo), y los criterios de edición se explicitan detalladamente. Todos ellos cuentan con un aparato crítico cuidadoso (quizás no suficientemente audaz en la *emendatio*), aunque la presentación de variantes y notas paleográficas no es muy lograda desde el punto de vista visual. Según apuntan los responsables del proyecto, se optó por elaborar una edición conservadora de los testimonios, manteniendo en general la ortografía de los originales.

Además de la cabecera que recoge los datos fundamentales de cada texto y edición, han tratado mediante etiquetas TEI elementos como intervenciones de los editores, deterioros, citas o cambios de lengua, identificación y anotación de antropónimos y topónimos y correspondencias de segmentos.

¹⁷ <<http://andresdepoza.com/>>.

En todo caso, se trata de un proyecto en fase de experimentación que merece los apoyos necesarios para explorar nuevas posibilidades.

El portal ha optado por una licencia Creative Commons BY-NC-ND.

3.5. *Lazarragaren eskuizkribua: edizioa eta azterketa* [El manuscrito de Lazarraga: edición y estudio]

El sitio web¹⁸ que vamos a tratar en esta sección está dedicado a la edición y estudio de la obra de un único autor, pero tiene la particularidad de ser el primer texto antiguo vasco cuya edición filológica ha sido pensada para ser editada en internet como *work in progress*. Es, seguramente, a día de hoy el proyecto más ambicioso y completo desde el punto de vista de la filología vasca y se enmarca en el grupo de investigación *Monumenta Linguae Vasconum* de la UPV/EHU (Bilbao *et al.*, 2010).

La obra editada es el llamado *Manuscrito de Lazarraga*¹⁹, un cartapacio incompleto descubierto en 2004, compuesto por textos que pueden datarse, seguramente, entre la segunda mitad del siglo XVI y principios del XVII (entre 1567 y 1602, aproximadamente).

El sitio web *Lazarragaren eskuizkribua: edizioa eta azterketa* está publicado bajo una licencia Creative Commons BY-SA, lo cual no ha evitado que sea citado de manera deficiente por varios investigadores. Está estructurado en varias secciones, que comentaremos brevemente a continuación: (1) La introducción reseña la vida y obra del autor principal del manuscrito, Juan Pérez de Lazarraga, y expone los criterios seguidos en la edición. (2) La gramática se encuentra todavía en fase de elaboración; en estos momentos, el sitio web sólo dispone de un índice provisional de la futura gramática. (3) El diccionario también está en fase de elaboración; por ahora se ofrecen las concordancias del texto "en bruto", descargables en formato PDF. (4) La sección "Versiones" detalla el registro de los cambios realizados en el sitio web en cada nueva versión. (5) Una bibliografía recoge las ediciones previas del manuscrito y los estudios sobre este o sobre Lazarraga en general, así como todas las obras que se han utilizado y mencionado en el trabajo; cuando existe una versión en la red de alguno de los trabajos recogidos en la bibliografía, se añade el enlace correspondiente. (6) La "Ayuda" brinda algunos consejos sobre cómo utilizar el sitio web. (7) Finalmente, la sección central está dedicada a la edición del manuscrito, que se ofrece en distintas versiones, algunas más avanzadas que otras, y que pasamos a detallar.

La primera versión del texto a la que nos referiremos es la edición crítica, que se presenta tanto en HTML para leer directamente en la red, como en versión PDF para descargarse o imprimir. Esta primera versión se centra en la transcripción semipaleográfica del texto (con regularización –no modernización– de la grafía, actualización de la puntuación, del uso de mayúsculas y de la separación de palabras, entre otros cambios). A la transcripción de cada texto precede una introducción que explica brevemente su contenido, estructura, métrica, fuentes, etc. Por último, la edición cuenta con un aparato de notas paleográficas y otro

¹⁸ Accesible desde estas dos direcciones: <<http://www.ehu.es/monumenta/lazarraga/>> y <<http://lazarraga.com>>.

¹⁹ Para más detalles sobre el manuscrito, ver Gómez y Urgell (2010) y Gómez (2011).

de notas explicativas. Las notas paleográficas, en la versión HTML, se anuncian mediante números volados y están situadas al final de cada texto.

Por su parte, en esta presentación en HTML las notas explicativas están marcadas mediante hiperenlaces y situadas en la columna derecha de la página web. El objetivo de estas notas es ofrecer al lector todas las claves para poder entender el texto del modo más completo y correcto posible. Para decidir la cantidad, tamaño y profundidad de las notas, se ha tenido en mente a un lector potencial que ya posee un conocimiento mínimo de la historia del euskera y de sus dialectos.

Como hemos señalado, la edición crítica cuenta también con una presentación en formato PDF. Se trata de la misma edición que se presenta en HTML, pero con algunas modificaciones en cuanto a su disposición: las notas explicativas son aquí notas a pie de página y las notas paleográficas aparecen al final de cada texto.

La segunda versión corresponde al texto en euskera con grafía modernizada, presentado en formato PDF. En esta versión se han suprimido las notas paleográficas, y las notas explicativas se han reducido a las necesarias para la comprensión del texto.

La tercera versión es la traducción del texto al castellano (actualmente en fase de elaboración). Se presenta únicamente el texto, sin aparato crítico.

Por último, cada texto se enlaza con la reproducción facsimilar digitalizada que publicó la Diputación de Gipuzkoa en 2004, tras presentar el manuscrito en sociedad.

4. NUEVA RECOPIACIÓN DE TEXTOS ANTIGUOS VASCOS

El equipo *Monumenta Linguae Vasconum* de la UPV/EHU, dirigido por el catedrático Joseba A. Lakarra y mencionado en el apartado anterior, trabaja actualmente en una nueva recopilación de textos antiguos vascos, que provisionalmente denominamos con el nombre genérico correspondiente en lengua vasca: *Euskal Testu Zaharren Bilbuma Berria*.

Este trabajo pretende recoger el testigo de *Textos Arcaicos Vascos* de Michelena (1964) y de *Contribución al estudio y edición de textos antiguos vascos* de Ibon Sarasola (1983), concebido en su momento precisamente como continuación del anterior. El proyecto aspira a recopilar y editar todo texto²⁰, no necesariamente literario, escrito en lengua vasca antes de 1745, de menos de 14.000 caracteres; será a su vez complementario con una crestomatía de textos vascos más extensos del mismo período y parte de la experiencia de la edición y estudio de la obra de Lazarraga, expuesta en el apartado inmediatamente anterior a este.

El equipo de trabajo simultaneará la elaboración de una edición en papel y una edición *online*, complementarias una de la otra, pero concebidas desde su inicio pensando en las características del soporte respectivo. La edición en papel recogerá la edición filológica convencional (semipaleográfica) de los textos, y la edición *online* ofrecerá no sólo la edición filológica de los textos, sino también

²⁰ Quedan excluidas del concepto de texto las palabras sueltas que no llegan a formar al menos una frase, así como los topónimos, antropónimos, teónimos... que no formen parte de una frase redactada en lengua vasca.

la edición paleográfica y los textos en grafía modernizada, así como las reproducciones de originales y/o primeras ediciones, junto con las notas bibliográficas completas y, en su caso, las reproducciones o los enlaces de las principales ediciones modernas de cada uno de los textos.

En la edición *online*, la edición filológica de cada texto en formato HTML permitirá hacer uso de todas las virtualidades de este tipo de edición (hiperenlaces, multimedia...), pero se adjuntarán también los PDF de cada texto y de la obra completa, en sus diferentes versiones (edición filológica, paleográfica y modernizada).

Solamente la versión *online* acogerá, también por motivos obvios, el léxico de todo el corpus, que será presentado de dos formas: por una parte, las concordancias bruto, tal como son generadas por la herramienta informática, para que los investigadores puedan generar sus propias búsquedas y sus propios subproductos; por otra parte, un léxico elaborado por el equipo, con criterios unitarios para todo el corpus.

Finalmente, una base de datos exhaustiva permitirá habilitar un sistema de búsquedas con filtros de cronología, variedad dialectal, editor moderno... así como palabras, secuencias de palabras...

BIBLIOGRAFÍA

- Bilbao, Gidor, "Textos Arcaicos Vascos 50 urte geroago", en J. Gorrochategui, R. Gómez, J.A. Lakarra y C. Mounole (eds.), *Actas del congreso Koldo Mitxelena*, Bilbao, UPV/EHU, 2013, pp. 717-728.
- Bilbao, Gidor *et al.*, *Lazarraga eskuizkribua: edizioa eta azterketa (1.0)*, Vitoria-Gasteiz, UPV/EHU, 2010, <<http://lazarraga.com>> [24/09/2013].
- Euskal Testuen Gordailua*, Patxi Salaberri Muñoa (resp.), Vitoria-Gasteiz, UPV/EHU, 2000, <<http://www.vc.ehu.es/gordailua/>> [23/09/2013].
- Gómez, Ricardo, "Presencia e influencia del castellano en el manuscrito Lazarraga", *Oihenart*, 26 (2011), pp. 231-258.
- Gómez, Ricardo y Urgell, Blanca, "Hacia una edición crítica del manuscrito Lazarraga", *Oihenart*, 25 (2010), pp. 63-88.
- Gorrochategui, Joaquín; Igartua, Iván y Lakarra, Joseba A. (eds.), *Euskararen historia*, Vitoria-Gasteiz, Gobierno Vasco, 2013.
- Klasikoen Gordailua*, Zarautz, Susa, 2013, <<http://klasikoak.armiarma.com/>> [24/09/2013].
- Lakarra, Joseba A., "Euskararen historia eta filologia: arazo zahar, bide berri", *ASJU*, 31:2 (1997), pp. 447-535.
- Michelena, Luis, *Textos Arcaicos Vascos*, Madrid, Minotauro, 1964.
- OEH-ko testu corpusa* [CD-ROM], Bilbao, Euskaltzaindia/Aurten Bai Fundazioa, 2000.
- Sarasola, Ibon, *Historia social de la literatura vasca*, Madrid, Akal Bolsillo, 1976.
- Sarasola, Ibon, "Contribución al estudio y edición de textos antiguos vascos", *ASJU*, 17 (1983), pp. 69-212.



RESUMEN

En este artículo se presentan y evalúan los diferentes proyectos desarrollados hasta la actualidad para facilitar el acceso en internet a textos antiguos vascos, entendiendo el concepto de "texto antiguo vasco" como texto no necesariamente literario, de diversa extensión, escrito en lengua vasca antes de 1745.

Se presentan las principales bibliotecas digitales que permiten el acceso a facsímiles de ediciones impresas y manuscritos de textos antiguos vascos; se subraya la necesidad de un directorio exhaustivo que permita localizar fácilmente estos documentos.

A continuación, se ofrece una evaluación exhaustiva de las fortalezas, debilidades y virtualidades de cinco proyectos muy diferentes en su concepción, pero que tienen en común la presentación en formato digital de textos antiguos vascos: *Euskal Testuen Gordailua* [Depósito de textos vascos], *Corpus del Diccionario General Vasco – Orotariko Euskal Hiztegia* (OEH), *Klasikoen Gordailua* [Depósito de clásicos vascos], *Portal Andrés de Poza* y *Lazarragaren eskuizkribua: edizioa eta azterketa* [El manuscrito de Lazarraga: edición y estudio].

Finalmente, presentamos el proyecto en el que trabaja actualmente el equipo *Monumenta Linguae Vasconum*: una nueva recopilación de textos antiguos vascos, que pretende avanzar a partir de la experiencia parcial de la edición del manuscrito de Lazarraga.

Palabras clave: Textos antiguos vascos, internet, edición filológica, repositorios digitales.

ABSTRACT

This article presents and evaluates the different projects developed so far with the aim of facilitating access to old Basque texts on the Internet, the concept of "old Basque text" meaning a text not necessarily literary, of diverse length and written in Basque before 1745.

First of all, the article presents the different digital libraries that allow access to facsimiles of print editions and manuscripts of old Basque texts, highlighting the necessity of an exhaustive directory that will help find those documents easily.

Next, it offers a detailed evaluation of the strong and weak points as well as the potentialities of five projects which, although they were conceived very differently, are similar in that they all present old Basque texts in digital format: *Euskal Testuen Gordailua* [Depository of Basque texts], *Corpus del Diccionario General Vasco – Orotariko Euskal Hiztegia* (OEH) [Corpus of the Basque General Dictionary], *Klasikoen Gordailua* [Depository of Basque classics], *Portal Andrés de Poza* [Andrés de Poza web portal] and *Lazarragaren eskuizkribua: edizioa eta azterketa* [Lazarraga's manuscript: editing and study].

Finally, the article presents the project on which the *Monumenta Linguae Vasconum* research group is currently working: a new compilation of old Basque texts that aims to make progress from the partial experience of the editing of Lazarraga's manuscript.

Keywords: Old Basque texts, Internet, scholarly editing, digital repositories.



La revista digital
Janus. Estudios sobre el Siglo de Oro
está patrocinada por el
SIELAE



SIELAE

www.bidiso.es/sielae/