

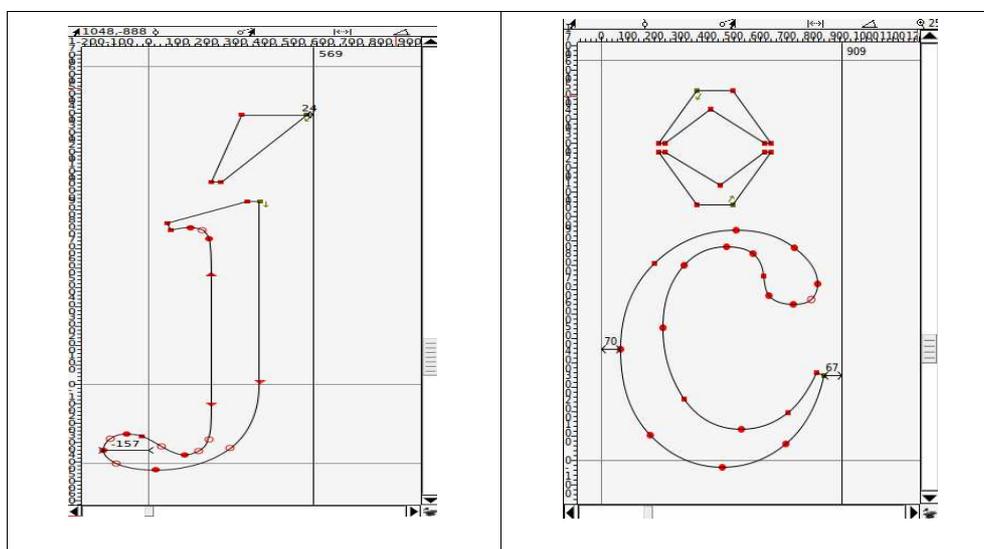
El corpus bíblico del español sefardí: De la planificación a la edición crítica

Francisco Javier Pueyo Mena
(CCHS, CSIC, Madrid)

1. INTRODUCCIÓN: JACOB HASSÁN Y EL MUNDO DIGITAL

De las muchas materias en las que el profesor Jacob Hassán era un experto y maestro singular quizá la faceta menos conocida por aquellos que no tuvieron la fortuna de trabajar a su lado sea la de *filólogo computacional*. Enemigo acérrimo del ratón, apasionado de las combinaciones de teclas más esotéricas para cualquier mínima interacción con el ordenador, el profesor Hassán visto desde la puerta de su despacho emergía entre una montaña de libros, carpetas y papeles que daban fe de su trabajo incansable. Acercándose un poco más uno podía descubrir en el centro de su escritorio un desgastado teclado y finalmente sus queridas y ruidosas impresoras. Aunque prefería redactar a mano todo cuanto escribía, finalmente su trabajo acababa insertándose digitalmente en alguno de los cientos de directorios arbolados, carpetas y subcarpetas infinitas que solo una mente como la suya podía comprender. Junto a él creamos la primera versión de las fuentes tipográficas digitales para la representación del sistema de transcripción de la Escuela Española de Filología Sefardí (Figura 1):

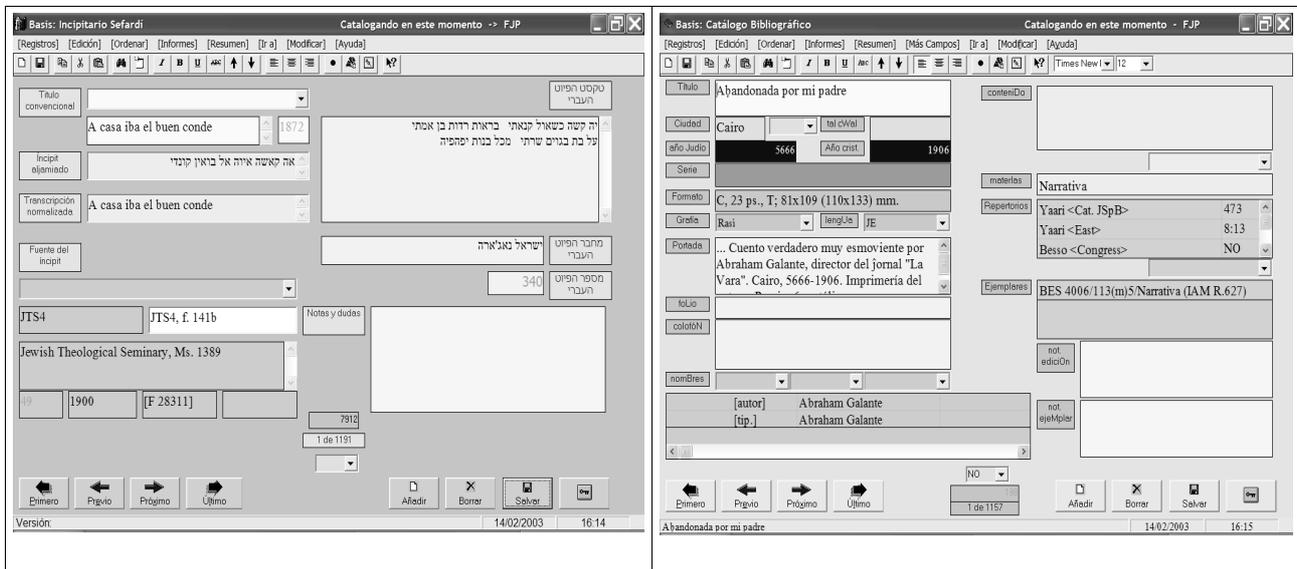
Figura 1: Creación de fuentes tipográficas



Bajo su dirección se desarrollaron también los diferentes programas que transformaban los códigos de los antiguos y ya descatalogados procesadores de texto al indispensable estándar Unicode. De su mente surgió la iniciativa de digitalización de las fuentes literarias sefardíes, y la idea de verter a una base de datos relacional los detallados listados bibliográficos, los ficheros lexicográficos y las incontables notas filológicas que cada vez resultaba más difícil consultar manualmente. En este nuevo contexto digital fue posible enlazar y relacionar entre sí todas las fuentes sefardíes acumuladas en años de cuidadoso trabajo filológico (desde los íncipit sefardíes en *piyutim* hebreos a las coplas, desde los

ingente labor de completar en lo posible el incomparable laboratorio de investigación filológica del departamento de estudios sefardíes del CSIC.

Figura 3: Basis: incipitario sefardí y catálogo bibliográfico



Como se muestra en la Figura 2, *Basis* es una base de datos compleja a cuyos módulos de trabajo y a su interfaz de comunicación dedicamos años de diseño y desarrollo. A pesar de ser un proyecto abierto y necesariamente inacabado, la idea digital de Jacob Hassán ya ha sido probada y usada con éxito en muchos de los trabajos del grupo y aún en el estado de desarrollo en el que se encuentra es una herramienta de trabajo filológico muy destacable.

El *Corpus bíblico del español sefardí*, y su interfaz gráfica de trabajo *Biblioteca Hispánica*, son parte de este proyecto global de digitalización y aprovechamiento computacional de las fuentes literarias sefardíes emprendido durante el magisterio de Jacob Hassán¹. Si bien ambos recursos digitales y filológicos eran y son responsabilidad mía, no hubieran podido emprenderse sin todos los recursos digitales emanados de su visión de futuro: las fuentes tipográficas, los estándares de codificación, las fuentes literarias digitalizadas, los diversos sistemas de transcripción, el acceso a la información bibliográfica y a los ficheros, o los sistemas de conversión de textos en diversas fases de trabajo editorial que de otra manera se hubieran perdido para siempre.

Pero por encima de la deuda digital se sitúa la deuda intelectual: esta iniciativa no se hubiera llevado a cabo de forma satisfactoria si no hubiéramos prestado la atención debida a sus silencios interminables y a sus miradas profundamente sarcásticas. Silencios y miradas que finalmente se materializaban en propuestas concretas que el profesor Hassán siempre formulaba en forma de preguntas engañosamente sencillas, cuyas complejas respuestas conocía de antemano, y que revelaban que el diseño o las implementaciones que se le habían

¹ Otra de las iniciativas que se está desarrollando en este sentido es el *Corpus Histórico Judeoespañol* (CORHIJE). Puede encontrarse una descripción del mismo en García Moreno & Pueyo Mena (2014) y un prototipo de su interfaz de consulta en García Moreno (2012-).

presentado no superaban el listón de la exigente crítica filológica que caracteriza cada uno de sus trabajos. Su atención al detalle textual o bibliográfico más insospechado, nos obligaba a descartar cualquier avance computacional si éste no era capaz de contemplar los casos más excepcionales. Trabajando junto a él aprendimos que, a pesar de las posibilidades de investigación que se abren al trasvasar al mundo digital los materiales de trabajo de las humanidades, cualquier necesidad computacional debe supeditarse a la precisión del método filológico.

En esta ocasión, el medio de publicación digital en el que aparece este trabajo, nos permite ofrecer una descripción panorámica e ilustrada del corpus, describiendo en las páginas que siguen tanto sus fases de diseño y planificación, como las herramientas empleadas para su desarrollo y algunas de las posibilidades de explotación del mismo.

2. PLANIFICACIÓN Y DISEÑO

El *Corpus Bíblico del Español Sefardí* se desarrolla dentro del grupo de investigación de Estudios Hispanojudíos y Sefardíes del Consejo Superior de Investigaciones Científicas. En este apartado, describimos concisamente su diseño: sus límites cronológicos, geográficos y tipológicos, los criterios de selección de los textos y la problemática de representación de sus fuentes.

2.1. Cronología, geografía y tipología

Definimos la lengua sefardí como la variedad hispánica que se habla o se escribe en las comunidades sefardíes fuera de España desde la expulsión de los judíos de los territorios peninsulares, a partir de 1492, y que se mantiene de forma minoritaria a comienzos del s. XXI. El origen de la lengua sefardí no es otro que las variedades hispánicas habladas en la Península ibérica por judíos y no judíos, pero más significativamente la lengua castellana. Evidentemente su evolución a partir de finales del s. XV es diferente a la del español peninsular, porque el judeoespañol, no hace falta recordarlo, no es español del s. XV, ni mucho menos una variedad arcaizante del español actual (Hassán 1994, 1995, 2006). El corpus se compone, por lo tanto, de textos escritos en judeoespañol dentro de los límites temporales que acabamos de definir y no se tienen en cuenta los textos bíblicos hispanojudíos realizados en castellano entre los siglos XIII y XV.

Una vez delimitado el marco cronológico y lingüístico del corpus necesitamos determinar los límites geográficos de los manuscritos y ediciones bíblicas en judeoespañol. Se tienen en cuenta todos los textos producidos en las comunidades sefardíes fuera de España (Italia, antiguo Imperio otomano, Austria, Países Bajos, Londres e Israel) con la excepción del norte de África, donde evidentemente existían comunidades judías en las que se hablaba judeoespañol, pero no imprenta ni, por lo tanto, ediciones bíblicas o manuscritos que se hayan preparado para la misma. Es por lo tanto intencionadamente complementario, geográfica y cronológicamente, al *Corpus Biblia Medieval* (Enrique-Arias, 2010-) que recoge, entre otros, los textos bíblicos hispanojudíos medievales.

En cuanto a la tipología textual de las fuentes que se incluyen en el corpus, el núcleo central lo conforman las traducciones bíblicas sefardíes y también la fuente hebrea de la que derivan directamente dichas traducciones. En una capa más externa del corpus se irán incorporando otro tipo de fuentes de interés para los estudios de traducción que quieran realizarse a partir del corpus central: por ejemplo, las paráfrasis en arameo o *targumim*, o los comentarios judíos medievales tanto en hebreo (por ejemplo, el de Raší) como en castellano (como el de Mosé Arragel de Guadalajara), textos que pueden explicar muchas de las

elecciones léxicas de los traductores sefardíes (y también de los traductores hispanojudíos medievales). En una capa todavía más externa del corpus situamos otras fuentes de interés comparativo (como el texto latino de la *Vulgata* o el griego de la *Septuaginta*), algunas versiones españolas que han hecho uso de las traducciones sefardíes (como la *Biblia del Oso*), o comentarios bíblicos cristianos que influyeron de forma notable en las traducciones cristianas a las lenguas vernáculas (como el de Nicolás de Lira), todos ellos de gran utilidad a la hora de converger, en el futuro, nuestro corpus con el corpus bíblico medieval (véase la Figura 4).

Figura 4: Tipología textual del corpus



2.2. Selección de los textos

Respecto a los criterios de selección de los textos que formarán parte del corpus, se optó por la incorporación de obras completas y no solo de fragmentos textuales de las mismas. Todos los textos bíblicos sefardíes son susceptibles de inclusión en el corpus, siempre y cuando caigan en alguna de las siguientes categorías²:

a. Biblias completas sefardíes: Las ediciones (y algunos manuscritos) en las que se ha transmitido el texto completo del canon bíblico hebreo en lengua sefardí se reducen en occidente a la *Biblia de Ferrara* y a las sucesivas reediciones y adaptaciones que de la misma se realizaron en la ciudad de Amsterdam³. En Oriente contamos con la que podríamos denominar *Biblia aljamiada del s. XVI*⁴. A esta última se unen en el corpus las diferentes

² En este trabajo nos referimos a las ediciones y a los manuscritos bíblicos por su lugar y fecha de producción, para una descripción completa de las fuentes véase Pueyo Mena (2008); los detalles bibliográficos de las ediciones pueden consultarse en Lazar (1964, 1994).

³ *Biblia en lengua española* (Ferrara 1553). Reediciones y adaptaciones: Amsterdam 1611, 1630, 1646, 1661 y 1726; edición bilingüe de Proops (Ámsterdam 1762).

⁴ Representada por los siguientes testimonios: *Pentateuco* (Constantinopla 1547), *Profetas anteriores* (Ms. JTS 36106 – Constantinopla 1580), *Salmos* (Constantinopla [1540]), *Profetas posteriores y Escritos* (Salónica 1568-1572, reimpr. 1583-1585).

ediciones de la traducción realizada por Abraham Asá en el siglo XVIII y la traducción (o adaptación de esta última) de Yisrael Bajar Ḥayim⁵.

b. *Libros sueltos sefardíes*: A lo largo de los siglos XVII, XVIII y XIX son numerosas las ediciones de libros bíblicos independientes⁶ o de secciones más amplias de la Biblia hebrea⁷. Los libros sueltos que se imprimieron con mayor frecuencia fueron el *Cantar de los cantares*, *Rut* y *Ester*, es decir, aquellos que más se necesitaban debido al uso litúrgico o paralitúrgico de los mismos durante las fiestas religiosas.

c. *Biblias cristianas*: Se incluyen también algunas de las Biblias e impresiones de libros sueltos realizadas por misioneros cristianos en judeoespañol⁸ y se incluyen igualmente otras Biblias cristianas que se deriven o que muestren una influencia destacable de la tradición sefardí⁹. Además de la significación social, histórica o exegética¹⁰ del fenómeno de

⁵ *Biblia de Abraham Asá* (Constantinopla 1739-1745) [reimpr. *Séfer Ḥaméš meguil-lot* (Salónica 1808), reimpr. *Séfer Ketubim* (Salónica 1809)]; *Biblia de Yisrael Bajar Ḥayim* (Viena 1813-1816).

⁶ Algunas de estas ediciones son: *Ester* (Viena ca. 1900), *Job* (Liorna 1778), *Salmos* (Ms. BZI 2275, Viena 1822, Ms. JNUL 8° 920, Salónica 1888, 1898 y 1930, Constantinopla 1936), *Rut* (Venecia 1753, Liorna 1777, Salónica 1878, 1884, Viena 1890, Salónica 1894 y 1906, Viena 1909, Salónica 1921 [reimpr. 1927, reimpr. 1932]), *Cantar de los cantares* (Salónica 1600, Venecia 1619 [Abraham Lañado, *Séfer Necudot hakésef*], 1655, 1672, 1695, 1721, 1756 y Liorna 1769 [reimpr. 1797], Venecia 1778, Salónica 1796 y 1800, Venecia 1804, Salónica 1805, Pisa 1822, Salónica 1828, Liorna 1856 [reimpr. 1929, reimpr. 1930], 1860, Salónica 1863 y 1867 [reimpr. 1872], Liorna 1870, Salónica 1875, Jerusalén 1891, Salónica 1895 y 1897), *Eclesiástico* (Viena 1808).

⁷ Representativos de este tipo de ediciones, tanto en Occidente como en Oriente, serían: *Pentateuco* (Menasseh ben Israel, Ámsterdam 1655), *Ḥumáš con haftarot* (Ámsterdam 1627, 1643, 1655, 1691, 1697, 1705, 1718, 1724 y 1733), *Pentateuco* (Mantua 1742).

⁸ La primera traducción fue realizada por Wilhelm Gottlieb Schaffler (Misión Americana): Viena 1841 (reimpr. Esmirna 1843), *Pentateuco* (Viena 1845). La segunda traducción la llevó a cabo, probablemente, Alexander Thomson (Misión Anglicana Escocesa): Constantinopla 1873 (reimpr. [sin hebreo] 1873, reimpr. 1905), *Pentateuco* (reimpr. 1895 y 1905), Viena 1931. Algunas ediciones de libros sueltos de las misiones protestantes en el antiguo Imperio Otomano son: *Salmos* (Constantinopla 1836, 1873 y 1878 [reimpr. 1885, reimpr. 1900, reimpr. 1909], 1922 y Esmirna 1853). También se tradujo al judeoespañol el *Nuevo Testamento: Firmamiento Muevo* (Corfú 1829), *Libro de el Nuevo Testamento* (Atenas 1844), *Nuevo Testamento* (Constantinopla 1877), *Los cuatro Evangelios* (Londres 1891), y *Nuevo Testamento* (Constantinopla 1894; reimpr. 1922).

⁹ En particular la conocida *Biblia del Oso* (Basilea 1569). El propio Casiodoro de la Reina declara en su *Amonestación del intérprete de los Sacros Libros al lector*, el importante influjo lingüístico que la versión sefardí de Ferrara ejerció sobre su traducción: «De la vieja translación española del *Viejo Testamento*, impresa en Ferrara, nos avemos ayudado en semejantes necessidades más que de ninguna otra que hasta ahora ayamos visto, no tanto por aver ella siempre acertado más que las otras en casos semejantes, quanto por darnos la natural y primera significación de los vocablos hebreos y las diferencias de los tiempos de los verbos, como están en el mismo texto, en lo qual es obra digna de mayor estima (a juicio de todos los que la entienden) que quantas hasta aora ay; y por esta tan singular ayuda, de la qual las otras translaciones no han gozado, esperamos que la nuestra por lo menos será inferior a ninguna de ellas [...] Esto nos pareció dezir de la translación de Ferrara en este lugar, no privándola de la alabança que justamente merece, ni encubriendo con invidia el ayuda que en la nuestra avemos tenido de ella.» (*Biblia del Oso*, hs. xxi-xxii).

¹⁰ En la citada *Amonestación*, Casiodoro de la Reina ejemplifica el tipo de consideraciones exegéticas que se derivan del uso de las versiones sefardíes y la necesidad de su adaptación desde la perspectiva cristiana: «Fuera de esto tiene también grandes yerros, algunos afectados en odio de Cristo, como en el capítulo 9 de *Isaías*, versículo 4, donde trasladó ansí: “Y llamó su nombre el Maravilloso, el Consejero, el Dio Barragán, el Padre eterno, Sar-Salom”, añadiendo de suyo, con malicia rabínica, este artículo *el* en todos estos nombres y no en el postrero Sar-Salom”, siendo otramete assaz diligentes los autores de ella en no dexar los tales artículos quando el texto los pone, y en no poner lo que no hallan en él. Mas en este lugar este sacrilegio, al parecer tan pequeño, no les es de poca importancia para defenderse en su pertinacia, porque se vee claro que todos aquellos nombres son títulos clarísimos del Messías, algunos de los cuales testifican abiertamente su naturaleza divina y poniendo el artículo *el* en cada uno de los precedentes y callándolo en el postrero, da a entender que los precedentes todos son nombres de Dios, y el último solo del Messías, como si dixesse: “El Maravilloso, el

cristianización del texto bíblico sefardí para el uso de las comunidades judías o cristianas, hay aspectos lingüísticos destacables que podrán ser estudiados mediante la incorporación al corpus de los textos cristianos en judeoespañol o de las versiones castellanas derivadas de las sefardíes. Las traducciones de Schauffler o de Casiodoro de la Reina son muestra de un proceso de reelaboración de los materiales tradicionales sefardíes desde un cierto enfoque académico o erudito, donde se consultan, cotejan o discuten fuentes y lecturas diversas. Por otro lado, la Biblia de la Misión Escocesa resulta de interés precisamente por lo contrario: la reelaboración de un material, cuya transmisión y cuya forma lingüística había alcanzado con el paso de los siglos un alto grado de sacralización (y por lo tanto una cierta concentración de arcaísmos, tanto léxicos como gramaticales), que los misioneros cristianos consideraban seguramente anticuado y que quisieron renovar mediante la sustitución de buena parte del vocabulario tradicional con formas léxicas extraídas tanto del registro popular del judeoespañol como del español peninsular de la época y mediante un intento consciente de deshebraización de la sintaxis y de la morfología del judeoespañol bíblico tradicional.

d. Glosarios bíblicos: Entre los glosarios bíblicos se tienen en cuenta aquellos que siguen la ordenación bíblica en libros, capítulos y versículos, que son las divisiones estructurales que nos permiten alinear todos los textos del corpus. Tendremos en cuenta, por lo tanto: a) el *Hěšec Šelomó*, que es el más antiguo de ellos, y que apareció en Venecia en 1588, y su reedición en la misma ciudad en 1617, b) las explicaciones lingüísticas y traducciones de palabras al ladino de Ya‘acob Lombroso a la Biblia hebrea publicada en Venecia en 1639), y c) el *Séfer Or tob* (Ámsterdam 1675) de Yehudá Arié de Módena, que al estilo del *Hěšec Šelomó*, traduce el texto bíblico siguiendo el orden del canon hebreo.

e. Fuentes excluidas: Quedan, por lo tanto, excluidas del corpus las siguientes fuentes bíblicas sefardíes: los fragmentos bíblicos insertos en comentarios sefardíes (como el *Me‘am lo‘ež*), las *hağadot* de Pésaḥ, los *maḥzorim* y *siḏurim* (libros de oraciones), las *haftarot* (lecturas de los Profetas), los glosarios alfabéticos y otras glosas en ladino a Biblias en hebreo.

2.3. Naturaleza multilingüe, multigráfica y polimórfica

En un trabajo anterior (Pueyo Mena 2014), hemos estudiado de forma extensa lo que hemos denominado la *problemática de representación de los textos bíblicos en ladino*, es decir, los problemas que la naturaleza misma de las fuentes bíblicas sefardíes plantean a la hora de incorporarlas a un corpus digital: el uso de varias lenguas en las fuentes, la utilización en los textos de diversos sistemas gráficos, el polimorfismo inherente a la historia del judeoespañol y la preservación de los textos originales en distintos medios digitales. También se describen allí las soluciones implementadas en el diseño y el desarrollo del corpus, mediante los procesos de normalización, lematización, alineación y la inserción de metadatos en los propios artefactos digitales.

En cuanto a la problemática del multilingüismo en las fuentes, el corpus incorpora como parte central del mismo el texto bíblico hebreo, el cual ha sido marcado con el mismo nivel de profundidad que las versiones sefardíes (véase la sección 3.2). Cuando explicaba anteriormente que hemos derivado a las capas exteriores del corpus otro tipo de fuentes

Consejero, el Dio Barragán, el Padre eterno llamó su nombre [*del Mesías*] Sar-Salom”. Esta maldita malicia no ha lugar, si se traslada fielmente como está en el texto sin poner el artículo *el* sobre ningún nombre, como nosotros avemos trasladado, porque entonces, aunque pese a quien le puede pesar de la gloria del Messías, el nombre Sarsalón va con todos los precedentes de esta manera: “Y llamó (*Dios*) (o *será llamado*, activa por pasiva como es frecuentísimo uso de la Escritura) su nombre [*del Messías*] Maravilloso, Consejero, Dios, Fuerte (o *Valiente* o *Valeroso*), Padre eterno, Príncipe de paz.”» (*Biblia del Oso*, hs. xxi-xxii).

bíblicas (como los comentarios medievales o la *Vulgata*), debe entenderse que dichas fuentes recibirán la marcación mínima necesaria para asegurar su alineación en la búsquedas con los demás textos (es decir, la marcación de libro, capítulo y versículo bíblicos). Por lo tanto, no serán lematizados ni anotados gramaticalmente, ya que ambos procesos en muy poco podrían contribuir a la caracterización lingüística o de traducción de los textos sefardíes, que son el objetivo central de nuestro corpus. Por lo tanto, el corpus es plenamente multilingüe en el sentido de que el hebreo (o el arameo) son parte del núcleo central del corpus y quedan anotados en profundidad, y lo es más periféricamente porque en cualquiera de las consultas realizadas sobre los textos sefardíes o sobre la fuente hebrea se obtienen de forma paralela pasajes correspondientes en latín, griego y español medieval.

Respecto a la naturaleza multigráfica de las fuentes originales, nos enfrentamos en primer lugar a la grafía hebrea en letras cuadradas vocalizadas del texto bíblico hebreo (y de algunas de las versiones sefardíes). Puede verse un ejemplo de dicha grafía en la Figura 5.

Figura 5: Grafía hebrea en letras cuadradas vocalizadas (*Pentateuco de Constantinopla, 1547*)

	<p>^{1]} Estas las palabras que habló Mošé a_todo Yisra'el en parte del Yardén, en_el dišierito, en_la llanura a_escuentera de Suf, entre Parán y entre Tofel, y Labán, y Ḥašerot, y Di-žahab. ^{2]} Onće días de Ḥoreb, carera de monte de Se'ír, hasta Cadéš Barnea'. ^{3]} Y fue en cuarenta años, en onće meses, en uno a_el mes, habló Mošé a_hijos de Yisra'el como todo lo_que encomendó Adonay a_él aellos; ^{4]} Enpués de su herir a Siḥón, rey de el 'Emorí, que están en Ḥešbón, y a 'Og, rey del Bašan, que están en 'Aštarot, en 'Edre'i. ^{5]} En parte de el Yardén, en tierra de Mo'ab, en-veluntó Mošé declarar a_la ...</p> <p>(Deuteronomio 1:1-5)</p>
--	--

En segundo lugar, la mayoría de las fuentes bíblicas sefardíes aljamiadas se presentan en tipografía raší, que es en la que la lengua sefardí se ha expresado por escrito tradicionalmente en la imprenta, como se muestra en la Figura 6.

Figura 6: Grafía aljamiada en letra raší
(Biblia de Asá, Constantinopla 1744)

<p>אי פ'ואי אין דיאש די ג'וזגאר לוי ג'ואיוש אי פ'ואי אמברי אינלה טיירה אי אנדוב'ב' ארון די בית לחם יהודה פור מוראר אין קאמפוס די מואב איל אי סו מו'יר אי דוש שוס אי'וש : (ב) אי נומברי דג איל ב' ארון אלימלך אי נומברי די סו מו'יר נעמי אי נומברי די דוש סוס אי'וש מחלון אי כליון אפרתים די בית לחם יאודה איונירון אקאמפוס די מואב אי פ'ואירון אלוי : (ג) אי מוריין אלימלך מא'ידו די נעמי אי פ'ואי רומא ניסידה אלייא אי דוש שוס אי'וש : (ד) אי טומארון פארה אלייוש מו'יריש מואביות נומברי די לה אונה ערפה אי נומברי די לה סיגודה רות אי אישטונירון אלוי קומו דייש אניוש : (ה) אי מוריירון</p>	<p>¹) Y fue en días de juzgar los jueces, y fue hambre en la tierra, y anduvo varón de Bet Léhem Yehudá por morar en campos de Mo'ab, él y su mujer y dos sus hijos. ²) Y nombre de el varón Elimélej, y nombre de su mujer Na'omí, y nombre de dos sus hijos Maḥlón y Jilyón, efratim de Bet Léhem Yehudá; y vinieron a campos de Mo'ab, y fueron allí. ³) Y murió Elimélej, marido de Na'omí, y fue romaneada ella y dos sus hijos. ⁴) Y tomaron para ellos mujeres mo'abiyot: nombre de la una, 'Orpá y nombre de la segunda, Rut; y estuvieron allí como diez años. ⁵) Y murieron ...</p> <p style="text-align: right;">(Rut 1:1-5)</p>
---	--

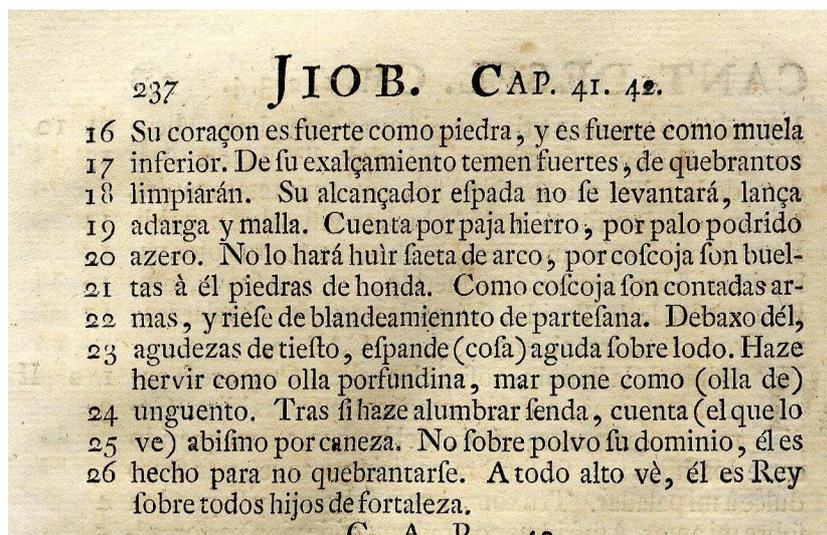
En tercer lugar, hay que tener en cuenta el hecho de que algunas fuentes bíblicas sefardíes solamente se han conservado en escritura aljamiada manuscrita. Sirva como ejemplo el tipo de caligrafía profesional que se muestra en la Figura 7 –colofón ya conocido–, presumiblemente empleada con la intención de enviar el texto a la imprenta y que dice lo siguiente: «Fue acabada esta obra en día de jueves, / vente y ocho del mes de 'elul de cinco mil trecientos / y cuarenta [= 18 sept. 1580], en caña de la ilustrísima señora / marat ['doña'] Reina, almanat ['viuda de'] hasar vegadol beYisrael ['el príncipe y grande en Israel] hadukos ['el duque'] haadón ['el señor'] / don Yosef Hanasí žlhh' ['su recuerdo perdure en la vida del mundo futuro']. Hecho a requerimiento del / yacar venijbad ['preciado y honorable'] kh'r ['su excelencia rabí'] Ya'acob 'Imanuel yŠv ['le guarde su Roca (Dios) y le dé vida']. El Šy't ['Dios'] le dé / žejut ['el mérito de'] que siempre esté en su gracia. Amén».

Figura 7: Aljamía manuscrita (Constantinopla, 1580)
Profetas Anteriores - Colofón

פ'ואי אקאבארה אישטה אוסרה אין האה ה ג'א'תם
ניטוי אי אוש הל מיש ה אלול ה שינקו מיל טבה-טל
אי קואהנמה אין קאזה הלה אלעס טבה שיהיה שינוה
מרת היטא שמה השד ונתל בשדא הרוכוס האון
דין יוסף הנשיא זלה' הי' אה הקיהינענו חל
יקר ועבד כהר' ישרב עמינאל א' אל שיה' ל' ה
זלות קי שיימטה אישטי אין סו נראטיאל אמן

Sin embargo, en ocasiones y en relación con la reincorporación de judíos ex-conversos españoles (y deshebraizados) a las prácticas religiosas del judaísmo, primero en Italia y más tarde en los Países Bajos, la Biblia se imprime en letras latinas, como muestra la Figura 8.

**Figura 8: Grafía latinada
(Edición de Proops, Ámsterdam 1762)**



Como solución a esta variación en el plano gráfico y para lograr la uniformidad de los cuatro tipos de grafías señalados, el corpus no incorpora la aljamía de forma textual sino que los textos aljamiados se representan en forma de transcripción normalizada. No me detengo en la explicación del método de transcripción de la aljamía¹¹, aunque sí quisiera dejar constancia de dos aspectos del sistema que son relevantes para nuestro corpus:

a) Está basada en la grafía hispánica estándar, a la cual se le añaden signos diacríticos para expresar aquellos sonidos del español sefardí que no existen en español. Hassán (1988, 1998, 2008) ya había previsto las ventajas de la eliminación de los signos diacríticos a la hora de usar los textos en un contexto de estudios o difusión hispánicos. En nuestro corpus prescindimos internamente de los diacríticos para la realización de los procesos de categorización, lematización y alineación. Hacemos uso, por lo tanto, de textos completamente hispanizados, factor que facilita mucho los procesos citados, ya que podemos aprovechar y adaptar las herramientas de procesamiento lingüístico desarrolladas para el sistema de escritura hispánico estándar. El uso de otros sistemas de transcripción o edición que se han popularizado, por ejemplo el de la revista *Aki Yerushalayim*, dificultaría notablemente la anotación de los textos del corpus.

b) El método de transcripción es filológicamente acertado, ya que va más allá de la mera transliteración y ofrece una lectura interpretada de la grafía hebrea (la cual puede darse, sin duda, a diversas interpretaciones), lo que nos permite ofrecer una presentación crítica de los textos, más allá de la mera transliteración. En cualquier caso en el corpus se mantiene una versión adicional de los textos en la que se preservan los cortes de línea y de hoja, así como la unión y separación originales de las palabras, las divisiones estructurales del original y, en el

¹¹ Para una explicación detallada del mismo, véase Hassán (1978) y sobre su aplicación concreta en el corpus bíblico sefardí puede consultarse Pueyo Mena (2014).

caso de los textos latinados, las grafías propias del original. Este tipo de presentación pseudo-paleográfica se ofrece para facilitar el cotejo de las transcripciones con las fotografías digitales del corpus y nos permite una integración más homogénea con el corpus *Biblia Medieval* que está constituido, en la fase de desarrollo en la que se encuentra, por las transcripciones paleográficas de los textos. En nuestro caso, sin embargo, precisamos de la edición plenamente normalizada ya que nuestra intención es mantener un corpus anotado en profundidad. A partir de la presentación crítica podemos realizar con mayor precisión las labores de alineación, lematización y marcación gramatical de las que hablaremos más adelante.

El polimorfismo característico de la lengua sefardí es fruto de fenómenos de variación bien conocidos, como la neutralización de las vocales cerradas en posición átona, la existencia de formas diptongadas y sin diptongar para el mismo vocablo, el refuerzo articulatorio consonántico de diptongos en posición inicial de palabra, la metátesis de consonantes, la pérdida o conservación de [y] (o su aparición como refuerzo antihiático), la conservación o pérdida de la F- inicial latina, la labialización de nasales, y otros muchos procesos de variación característicos del judeoespañol, como la síncope, la prótesis de a-inicial, la aféresis, etc.

A pesar de que dicho polimorfismo resulta más frecuente en otro tipo de registros de la lengua que en el registro bíblico, es un fenómeno que presenta un reto evidente a la hora de desarrollar cualquier corpus histórico y geográfico del judeoespañol. En el caso del corpus bíblico, la mejor solución para uniformar los distintos resultados de los procesos citados –sin descartar las variantes diastráticas, diatópicas, diafásicas y diacrónicas– es mediante la realización de un proceso de lematización, es decir, la asignación de una entrada léxica única para todas las formas existentes de un mismo vocablo: *golor / olor* → OLOR, *día / diya* → DÍA, etc.

3. DESARROLLO E IMPLEMENTACIÓN

Antes de abordar la implementación del corpus, es necesario detenerse brevemente en el tipo de traducciones que nos ofrecen los textos bíblicos sefardíes. Quizá pueda parecer fuera de lugar hablar de los contenidos de los textos en un trabajo centrado en la descripción de la propia construcción del corpus y no en las particularidades de sus contenidos, pero en el caso específico del corpus bíblico sefardí resulta imprescindible hacerlo, ya que precisamente es el tipo de traducción que encontramos en los textos sefardíes lo que dirige desde el primer momento tanto su diseño como su implementación.

3.1. *El método de traducción tradicional*

La característica fundamental de la traducción, no solo de las Biblias sefardíes sino de cualquier traducción al español que siga el método tradicional de los judíos, se resume en la portada de la Biblia de Ferrara de 1553 (Figura 9):

Figura 9: Detalle de la portada de la Biblia de Ferrara 1553



«Biblia en lengua española traducida palabra por palabra de la verdad hebraica», y que en realidad puede entenderse como «palabra por *palabra(s)* de la verdad hebraica». Es decir, que cada vocablo del texto hebreo debe traducirse con una sola palabra romance si es posible, o con más de una si no lo es, y que debe seguirse en lo posible el orden de palabras del original hebreo. Cuando el traductor tradicional sefardí se enfrenta al primer versículo de la Biblia hebrea necesita encontrar un *ladino* (es decir, una equivalencia en judeoespañol) para cada una de las siete palabras hebreas que lo componen (véase la Tabla 1, donde se muestra el mismo proceso en dos testimonios, el primero sefardí y el segundo medieval, separados por más de un siglo).

**Tabla 1: Proceso tradicional de traducción (Génesis 1:1)
Pentateuco de Constantinopla 1553 / Escorial I.i.4, c. 1450**

Hebreo	Be-re'sit	→	bará'	→	'Elohim	→	'et	→	ha-šamáyim	→	ve-'et	→	ha-'areš
	↓		↓		↓		↓		↓		↓		↓
PC	En * <i>prencipio</i>		<i>crió</i>		<i>el Dio</i>		* <i>a</i>		<i>los cielos</i>		y * <i>a</i>		<i>la tierra</i>
E4	En * <i>comienço</i>		<i>crió</i>		<i>Dios</i>		* <i>a</i>		<i>los cielos</i>		e * <i>a</i>		<i>la tierra</i>

El ladinador no se detiene a examinar si es necesario o no traducir todas y cada una de las palabras hebreas para acertar con el significado global del versículo en romance y tampoco se detiene a considerar si el resultado de la traducción es gramaticalmente aceptable en judeoespañol (o en español medieval). En el ejemplo hemos indicado con un asterisco dos casos típicos de agramaticalidad en las traducciones: la ausencia del artículo determinado *el* ante el sustantivo (*prencipio* / *comienço*), y la traducción del marcador hebreo de objeto directo ('et ↔ *a*) donde en español no es necesario, ya que en ambos casos se refieren a objetos directos inanimados (*los cielos* / *la tierra*).

Las flechas verticales y horizontales del ejemplo simbolizan lo que se conoce como el método de traducción posicional progresiva (Kohring, 1996), en el que cada palabra hebrea se traduce en su lugar (↓), por lo que el orden del original hebreo (→) se preserva estrictamente. Esto supone sin duda forzar la sintaxis romance y se traduce en el empleo de

un orden de palabras altamente artificial o, mucho mejor dicho, altamente hebraizado. En el versículo 1:1 del libro de *Rut*: «... y partió de Belén de Judá un hombre a morar en la campiña de Mo'ab, él, su mujer y *sus dos hijos*» (Cantera & Iglesias, 2000), la versión de la *Biblia de Asá* presenta la siguiente traducción: «y anduvo varón de Bet Léhem Yehudá por morar en campos de Mo'ab, él y su mujer y **dos sus hijos*». En la Tabla 2, observamos como el traductor ladino en primer lugar el vocablo *šené* (↓ ‘dos’) y solamente después (→) puede pasar a traducir la siguiente palabra *banav* (↓ ‘sus hijos’), quedando la construcción en ladino claramente hebraizada.

Tabla 2: Ejemplo de construcción hebraizada

šené	→	banav
↓		↓
<i>Dos</i>		<i>sus hijos</i>
<i>sus dos hijos</i>		

Tanto el orden estrictamente secuencial como la necesidad de ajustarse a cada palabra hebrea lo hemos aprovechado en los procesos de alineación del corpus. Sin embargo, hay algunas otras características del método tradicional que nos han ayudado durante los procesos de categorización y lematización. Por ejemplo, la necesidad de adecuar las categorías gramaticales de la traducción a las del original hebreo, fenómeno que se extiende a todo tipo de palabras: sustantivos, verbos, adjetivos, adverbios, preposiciones, conjunciones, artículos, pronombres e interjecciones. Además de la búsqueda de categorías gramaticales equivalentes para cada tipo de palabra, el método tradicional busca ajustar lo más posible la morfología del judeoespañol a la del original hebreo, usando diferentes técnicas lingüísticas, como la creación de numerosos neologismos mediante la derivación de verbos a partir de sustantivos, o el uso en el sistema verbal de correspondencias romances estrictas y (deliberadamente) limitadas para la traducción de los tiempos y conjugaciones verbales del hebreo.

Algunas de estas correspondencias habituales son: el uso del participio de presente apocopado romance para traducir el participio hebreo (*Gé 1:2 merahéfet* ↔ *avolán*), el uso del infinitivo ladino para traducir el infinitivo hebreo (*Gé 9:14 be'anení* ↔ en mi *anubear*), el uso de participios de pasado en ladino allí donde el hebreo también los presenta (*Gé 4:12 na' vanad* ↔ *meneado* y *esmovido*), la traducción al romance de las conjugaciones *nif'al* y *hof'al* hebreas como pasivas analíticas (*Gé 1:9 yicavú* ↔ *sean apañadas*) o de las conjugaciones *hitpa'el* e *hitpol'el* como formas reflexivas (*Gé 9:21 vayitgal* ↔ *descubrióse* / *Gé 2:25 yitbošašu* ↔ *se avergonzaban*), la traducción del perfecto o del imperfecto consecutivo hebreos como pretéritos perfectos romances (*Gé 4:11 paštá* ↔ *abrió* / *Gé 2:7 vayipah* ↔ *y asopló*), y de igual modo la consistencia en traducir el imperfecto y el perfecto consecutivo hebreos como futuros romances (*Gé 3:15 'ašit* ↔ *porné* / *Gé 3:5 venifquehú* ↔ *y serán abiertos*), o la equivalencia entre los imperativos hebreos y romances (*Gé 1:28 vejibšuah uredú* ↔ *y sochiguadla y podestad*)¹².

Teniendo presente este peculiar método de traducir, común (de una u otra manera) a todos los textos de nuestro corpus, pasamos a describir los procesos concretos de desarrollo o implementación del corpus.

¹² En los ejemplos corresponden al glosario *Hěšec Šelomó*, véase Pueyo Mena (2011) para una explicación detallada de los mismos.

3.2. Niveles de marcación

Respecto al nivel de marcación de los textos, nuestro corpus está diseñado con una marcación intermedia, es decir, no se ha planificado anotar sintácticamente los textos. En el caso de la Biblia las unidades estructurales de marcación vienen dadas por sus propias divisiones estructurales:

a. Libro bíblico: Para asegurar la futura integración de los corpus bíblicos sefardí y medieval, se ha realizado la misma asignación numérica de los libros bíblicos que la empleada en el corpus *Biblia Medieval*. Además nos hemos adherido al mismo uso de abreviaturas estándar para cada uno de los libros bíblicos (véase la Tabla 3):

Tabla 3: Numeración y abreviaturas de los libros bíblicos

<i>Canon hebreo</i>		<i>Nuevo Testamento</i>
1. (<i>Gé</i>) Génesis	27. (<i>Sal</i>) Salmos	50. (<i>Mt</i>) San Mateo
2. (<i>Éx</i>) Éxodo	28. (<i>Jb</i>) Job	51. (<i>Mc</i>) San Marcos
3. (<i>Le</i>) Levítico	29. (<i>Pr</i>) Proverbios	52. (<i>Lc</i>) San Lucas
4. (<i>Nú</i>) Números	30. (<i>Rt</i>) Rut	53. (<i>Jn</i>) San Juan
5. (<i>De</i>) Deuteronomio	31. (<i>Ca</i>) Cantar de los cantares	54. (<i>Hech</i>) Hechos de los Apóstoles
6. (<i>Jos</i>) Josué	32. (<i>Ece</i>) Eclesiastés	55. (<i>Ro</i>) A los Romanos
7. (<i>Ju</i>) Jueces	33. (<i>La</i>) Lamentaciones	56. (<i>Cor1</i>) I a los Corintios
8. (<i>Sam1</i>) Samuel 1	34. (<i>Est</i>) Ester	57. (<i>Cor2</i>) II a los Corintios
9. (<i>Sam2</i>) Samuel 2	35. (<i>Dan</i>) Daniel	58. (<i>Gá</i>) A los Gálatas
10. (<i>Re1</i>) Reyes 1	36. (<i>Esd</i>) Esdras	59. (<i>Ef</i>) A los Efesios
11. (<i>Re2</i>) Reyes 2	37. (<i>Ne</i>) Nehemías	60. (<i>Flp</i>) A los Filipenses
12. (<i>Is</i>) Isaías	38. (<i>Cr1</i>) Crónicas 1	61. (<i>Col</i>) A los Colosenses
13. (<i>Je</i>) Jeremías	39. (<i>Cr2</i>) Crónicas 2	62. (<i>Te1</i>) I a los Tesalonicenses
14. (<i>Ez</i>) Ezequiel		63. (<i>Te2</i>) II a los Tesalonicenses
15. (<i>Os</i>) Oseas	<i>Deuterocanónicos o apócrifos</i>	64. (<i>Tim1</i>) I a Timoteo
16. (<i>Jl</i>) Joel	40. (<i>To</i>) Tobit	65. (<i>Tim2</i>) II a Timoteo
17. (<i>Am</i>) Amos	41. (<i>Jdt</i>) Judit	66. (<i>Tit</i>) A Tito
18. (<i>Ab</i>) Abdías	42. (<i>SupEst</i>) Suplementos a Ester	67. (<i>Flm</i>) A Filemón
19. (<i>Jon</i>) Jonás	43. (<i>Sab</i>) Sabiduría	68. (<i>Heb</i>) A los Hebreos
20. (<i>Mi</i>) Miqueas	44. (<i>Eco</i>) Eclesiástico	69. (<i>San</i>) Santiago
21. (<i>Na</i>) Nahum	45. (<i>Ba</i>) Baruc	70. (<i>Pe1</i>) I de San Pedro
22. (<i>Ha</i>) Habacuc	46. (<i>EpiJe</i>) Carta de Jeremías	71. (<i>Pe2</i>) II de San Pedro
23. (<i>So</i>) Sofonías	47. (<i>SupDan</i>) Adiciones a Daniel	72. (<i>Jn1</i>) I de San Juan
24. (<i>Ag</i>) Ageo	48. (<i>Mac1</i>) Macabeos 1	73. (<i>Jn2</i>) II de San Juan
25. (<i>Za</i>) Zacarías	49. (<i>Mac2</i>) Macabeos 2	74. (<i>Jn3</i>) III de San Juan
26. (<i>Mal</i>) Malaquías	77. (<i>Esd3</i>) Esdras 3	75. (<i>Jds</i>) San Judas
		76. (<i>Ap</i>) Apocalipsis

b. Capítulo: Cada capítulo recibe una marca numérica, según la numeración que le corresponde siguiendo la Biblia hebrea (o la *Vulgata* para los libros fuera del canon hebreo). El libro de *Salmos*, con 150 capítulos (151 en la *Vulgata*), es el libro con el mayor número de capítulos en el corpus.

c. *Versículo*: Del mismo modo cada versículo bíblico recibe una marca numérica siguiendo la numeración de la Biblia hebrea (o de la *Vulgata* en los casos de libros que no pertenecen al canon hebreo). En el corpus, el salmo 119, con 176 versículos, es el que contiene un mayor número de versículos.

d. *Palabra hebrea*: Más allá del nivel estructural de marcación, estamos procediendo a la marcación léxica y morfológica a nivel de palabra. Los procedimientos son diferentes para el texto hebreo y para los textos hispánicos. Para el texto hebreo, seguimos el uso del método de traducción tradicional en el que una palabra se entiende estrictamente como la secuencia de caracteres situados entre dos espacios en blanco. En el texto bíblico hebreo una palabra, tal y como acabamos de definirla, puede presentar uno o más formantes léxicos o morfológicos. En el caso de la forma *bere'sit*, que veíamos anteriormente, nos encontramos ante una palabra gráfica constituida por dos formantes gramaticales: una preposición (*be-*) y un sustantivo (*re'sit*). Cada palabra recibe un identificador único {libro-capítulo-versículo-palabra}, por ejemplo *bere'sit* {01-001-001-01} y a cada uno de los formantes de dicha palabra se le asigna también su propio identificador {libro-capítulo-versículo-palabra-formante}, por ejemplo *be* {01-001-001-01-a} y *re'sit* {01-001-001-01-b}. Se marca también la lengua de cada palabra {hebreo | arameo}, se asigna un número estandarizado correspondiente al lema de cada formante, por ejemplo {B658} y la acepción del vocablo en el contexto en el que aparece {a | b | ...}, se anota la forma superficial hebrea {יְאָמַר}, su transcripción {yó'mer} y finalmente el lema hebreo que le corresponde {אמר}. Cada formante hebreo recibe además una etiquetación gramatical completa, como muestra la Figura 10, donde se puede observar la categorización del término *yicavú* (*Gé* 1:9): *verbo, hebreo, conjugación nif'al, aspecto imperfectivo, significado yusivo, 3ª persona, masculino, plural y sin sufijos pronominales*.

Figura 10: Bibli@teca
Categorización del término hebreo יְאָמַר / yicavú

The screenshot shows the 'Bibli@teca' web application interface for Hebrew word categorization. At the top, there are navigation tabs: 'Ediciones', 'Búsquedas y comparaciones', 'Alineación del hebreo y las versiones', 'Categor. y lemat. del hebreo', and 'Categor., normal. y lemat. versiones'. Below this is the 'Módulo de categorización y lematización del hebreo'. The main area is divided into two sections: 'Elegir pasaje' and 'Ayuda'. The 'Elegir pasaje' section shows 'Libro: Génesis', 'Capítulo: 1', and 'Versículos: 9'. The 'Ayuda' section provides instructions on how to use the module. Below these are two rows of word forms with their identifiers. The first row shows: wayyô'mer (0100100901), 'elohim (0100100902), yiqqawú (0100100903), hammáyim (0100100904), mitahaj (0100100905), haššamayim (0100100906), 'el (0100100907), māqôm (0100100908), 'ehäd (0100100909), wəjə'rā'eh (0100100910), hayyabbāšāh (0100100911). The second row shows: 1 (0100100911a), אמר (0100100911b), אלהים (0100100902a), קוה (0100100903a), ה (0100100904a), מים (0100100904b), מן (0100100905a), תחת (0100100905b), ה (0100100906a), שמים (0100100906b), אל (0100100907a), מקום (0100100908a), אחד (0100100909a), ו (0100100910a), ראה (0100100910b), ה (0100100911a), ושה (0100100911b). Below the word forms is a table with columns: 'Partícula', 'Pronombre', 'Nombre', 'Adjetivo', 'Verbo', and 'Lematización'. The 'Verbo' section is expanded, showing 'Tipo hebreo: nifal (n)', 'Aspecto: imperfecto (i)', 'Persona: 3ª persona (3)', 'Género: masculino (m)', 'Número: plural (p)', and 'Estado: absoluto (a)'. The 'Sufijo verbal' section is also expanded, showing 'apocopado (Aa)', 'he paragógica (Ah)', 'nun paragógica (An)', and 'he direccional (Ad)'. The 'Sufijo pronominal' section is expanded, showing 'Persona: s. 3ª persona (3)', 'Género: s. común (c)', and 'Número: s. plural (p)'. At the bottom, there is a footer: 'Biblioteca (Biblia Medieval y Biblia Sefardí) - © Francisco J. Pueyo Mena 2007-2010'.

e. *Palabra en ladino*: Respecto a las versiones sefardíes, los versículos están igualmente segmentados en palabras. Por ejemplo, la forma {קוויירון}, recibe también un identificador

único del tipo {libro-capítulo-versículo-palabra}, en este caso {01-003-007-11}, una transcripción normalizada {*cošieron*} y su equivalente ortográfico en español estándar {*cosieron*}, un lema en judeoespañol normalizado {*cošer*}, y el lema normalizado según el español normativo {*coser*}. La Figura 11 muestra la ficha de categorización gramatical de la palabra en ladino correspondiente al hebreo *yicavú*, según la traducción de Abraham Asá: *apáñense* (Gé 1:9): *verbo principal* (es decir, no auxiliar), modo *imperativo*, 3ª persona del plural (el género en el verbo se reserva para el participio), con pronombre *enclítico* de 3ª persona singular.

Figura 11: Bibli@teca
Categorización del término *apáñense*

Ediciones Búsquedas y comparaciones Alineación del hebreo y las versiones Categor. y lemat. del hebreo Categor., normal. y lemat. versiones

Módulo de categorización, normalización y lematización de versiones

Elegir versiones Elegir pasaje Ayuda

Fazlenda GE2 E19 BNM BNJob PC CantSalonica Libro Génesis

SalmosXIII GE3 E7 RAH BNJam Sal1540 Asa Capítulo 1

Este módulo nos permite comparar paralelamente pasajes de la Biblia en 20 versiones medievales y sefardíes. Seleccione las versiones que desea mostrar en la sección "Elegir versiones".

Y (E100100901) djo (E100100902) el (E100100903) Dio (E100100904) **apáñense** (E100100905) las (E100100906) aguas (E100100907) de (E100100908) debajo (E100100909) de (E100100910) los (E100100911) cielos (E100100912) a (E100100913) lugar (E100100914) uno (E100100915) y (E100100916) sea (E100100917) aparecido (E100100918) el (E100100919) seco (E100100920) Asa

Partículas Pronombre Nombre Adjetivo **Verbo** Determinante Lematización y normalización

Tipo principal (p) auxiliar (a) semiauxiliar (s)

Modo indicativo (i) subjuntivo (s) imperativo (m) participio (p) gerundio (g) infinitivo (f)

Tiempo presente (p) pasado (s) futuro (f) imperfecto (i) condicional (c)

Persona 1ª persona (1) 2ª persona (2) 3ª persona (3)

Género masculino (m) femenino (f)

Número singular (s) plural (p)

Sufijo pronominal

Persona s. 1ª persona (1) s. 2ª persona (2) s. 3ª persona (3)

Género s. masculino (m) s. femenino (f) s. neutro (n)

Número s. singular (s) s. plural (p)

Caso s. dativo (d) s. acusativo (a) s. oblicuo (o)

Biblioteca (Biblia Medieval y Biblia Sefardí) - © Francisco J. Pueyo Mena 2007-2010

3.3. Estándares de numeración

Las referencias estándar en las que basamos la anotación del corpus son los siguientes: para la división estructural de los textos y la marcación de versículos seguimos la *Biblia Hebraica Stuttgartensia* en su edición electrónica (Westminster-Claremont-Michigan Electronic Text), para la lematización y anotación morfológica del hebreo se sigue el *Groves-Wheeler Westminster Electronic Hebrew Morphology* (WHM), el identificador numérico del lema hebreo y de sus acepciones se anota siguiendo el diccionario *Hebrew and Aramaic Lexicon of the Old Testament* (HALOT).

Evidentemente esto no plantea ningún problema para el corpus bíblico sefardí. Sin embargo, sí complica en algunos fragmentos la integración de la *Vulgata* latina o en el futuro la de los textos medievales que la usaron como fuente de su traducción (*General Estoria*, E6, E8, etc.). Nuestra propuesta para los corpus bíblicos es que se atengan a los estándares establecidos para el texto hebreo (seguido por todas las ediciones bíblicas modernas) ya que así se facilitará la consulta y la integración de cuantos textos, comentarios, o glosarios bíblicos se pongan a disposición de los estudiosos en cualquier variedad o en cualquier

tradicción de traducción. Hacer uso de una tabla intermedia de correspondencias para la *Vulgata* y para sus traducciones no parece una labor demasiado compleja y nos permitirá ofrecer la numeración de la *Vulgata* en los resultados cuando sea necesario.

3.4. Alineación de las versiones

Una vez marcadas de manera independiente las palabras hebreas (y sus formantes) y las palabras sefardíes, el siguiente proceso que se realiza en el corpus es el de la alineación de los textos a nivel de palabra¹³. Esta propuesta supone quizá la parte más novedosa y paradójicamente también la más tradicional de nuestro corpus.

Dado que el método tradicional de traducción nos asegura un número de palabras más o menos constante (recordemos: «palabra por palabra o por palabras») y que el orden de cada palabra dentro del versículo es fijo y predecible, hemos querido aprovecharlo para centrar la alineación de los textos paralelos no solo a nivel del versículo, sino llevar dicha alineación al nivel mínimo de la palabra.

Tal y como explicaba anteriormente, la traducción palabra por palabra es un método estricto. Fue seguido en las traducciones medievales judías, aunque nunca de una forma tan rigurosa como en las traducciones realizadas para judíos antes o después de la expulsión. Los traductores judíos medievales que trabajaban para señores cristianos sabían muy bien que debían adaptar la traducción tradicional comunitaria al desconocimiento de los lectores cristianos del texto hebreo subyacente. Las traducciones sefardíes son estrictas en el método de traducción porque presuponen que el lector conoce (o quiere conocer) el texto hebreo de la Biblia. El hecho de que la *Biblia de Ferrara* sea menos rigurosa en su sistema que el *Pentateuco de Constantinopla* (y también el hecho de que se imprimiera en letras latinas) se explica porque su destinatario desconoce, aunque desea conocer, el texto hebreo.

En el corpus, cada palabra sefardí se anota con el identificador único de la palabra hebrea a la cual traduce. Este identificador, como hemos visto, tiene la forma {libro-capítulo-versículo-palabra} y por lo tanto sigue el orden secuencial del texto hebreo, por lo que no se corresponde con ninguna numeración estándar, ya que éstas no asignan un número a cada palabra del texto bíblico hebreo sino que, al tratarse de concordancias o de diccionarios de lemas, lo asignan a la raíz, y por lo tanto no pueden utilizarse para la alineación de textos paralelos.

En la Tabla 4 podemos examinar la alineación del versículo 16 del primer capítulo del *Cantar de los cantares* en hebreo y en las versiones de la *Biblia de Ferrara* (ladino occidental), *Biblia de Asá* (ladino oriental) y la *Paráfrasis caldaica* de Ámsterdam (véase en la Tabla 12, al final de este trabajo, un ejemplo del mismo versículo alineado en 16 versiones diferentes).

¹³ Una descripción detallada de los procesos de alineación automatizados desarrollados para el corpus bíblico sefardí y su relación con el sistema de traducción tradicional sefardí, puede encontrarse en Pueyo Mena (2013).

Tabla 4: Alineación a nivel de palabra en tres versiones bíblicas (Cantar 1:16)

Hebreo	Ferrara	Asá	Paráfrasis
<hinejá> הִנֵּיָא	Hé {31-01-16-01} tú {31-01-16-01}	Hec {31-01-16-01} tú {31-01-16-01}	He {31-01-16-01} tú {31-01-16-01}
<yafé> יָפֵה	hermoso {31-01-16-02}	hermošo {31-01-16-02}	hermoso {31-01-16-02}
<dodí> דּוֹדִי	mi {31-01-16-03} querido {31-01-16-03}	mi {31-01-16-03} querido {31-01-16-03}	mi {31-01-16-03} querido {31-01-16-03}
<'af> אָף	también {31-01-16-04}	también {31-01-16-04}	también {31-01-16-04}
<na'im> נְעִיִם	suave {31-01-16-05}	hermošo {31-01-16-05}	suave {31-01-16-05}
<'af> אָף	también {31-01-16-06}	también {31-01-16-64}	también {31-01-16-06}
<'arsenu> עֲרִשְׁנוֹ	nuestro {31-01-16-07} lecho {31-01-16-07}	nuestro {31-01-16-07} lecho {31-01-16-07}	nuestro {31-01-16-07} lecho {31-01-16-07}
<ra'ananá> רַעֲנָנָה	florido {31-01-16-08}	reverdido {31-01-16-08}	reverdecido {31-01-16-08}

Como puede observarse algunas palabras en ladino comparten el mismo identificador, ya que traducen el mismo vocablo hebreo.

3.5. Modelo de etiquetación: XML / SQL

En cuanto al modelo de etiquetación, los grandes corpus lingüísticos como CREA, CORDE, BNC, etc. han sentado un precedente en la utilización de SGML y sus derivados (como XML) para la marcación de sus textos. Ciertamente XML no es un lenguaje de marcación que pueda usarse durante las fases de transcripción, corrección o edición de los textos (como puede comprobarse en la Tabla 5), ni tampoco está específicamente diseñado para que los sistemas de recuperación de información hagan un uso óptimo de los documentos codificados en dicho formato. En lo que sí resulta fundamental el uso de XML y más concretamente el uso de las especificaciones XML de las instituciones de estandarización, como por ejemplo la TEI, es en el intercambio automatizado de información digital.

Tabla 5: Etiquetación de la palabra hebrea *bere'sit* en XML

<pre><?xml version="1.0" encoding="utf-8" ?> <bk> <ch> <v> <w> <m> <bibref isKQ="notKQ"> <book>1</book> <chapter>1</chapter> <verse>1</verse> <word>1</word> <part>a</part> </bibref> <hebtex>א</hebtex> <lemma>א</lemma> <homonym>a</homonym></pre>	<pre><lang>he</lang> <pos><pre></pre></pos> </m> <m> <bibref isKQ="notKQ"> <book>1</book> <chapter>1</chapter> <verse>1</verse> <word>1</word> <part>b</part> </bibref> <hebtex>ראשית</hebtex> <lemma>ראשית</lemma> <homonym>a</homonym> <lang>he</lang></pre>	<pre><pos> <n> <cn> <gender kind="f"/> <number value="s"/> <state type="a"/> </cn> </n> </pos> </m> </w> [...]</pre>
--	--	--

En nuestro caso, hemos decidido no marcar inicialmente los textos en dicho formato, sino generarlo de forma automatizada cuando sea necesario para la integración de nuestro corpus en otros sistemas. Una arquitectura fiable para realizar esta labor y para asegurar que las etiquetas XML están bien formadas es, en nuestra opinión, una base de datos con tablas y campos bien definidos y cuidadosamente relacionados. Para ello usamos bases de datos relacionales y, tanto en la creación como en las consultas del corpus, utilizamos el lenguaje estándar SQL¹⁴.

Las Tablas 6 a 11 muestran una representación de algunas de las tablas y campos SQL del corpus y ejemplifican cómo se almacena la información en la base de datos. Los campos de identificación única (bible_id, book_id, chapter_id, verse_id, word_id, morph_id, hword_id, etc.) son los que permiten mantener relacionados de forma paralela todos los textos, palabras y formantes del corpus.

Tabla 6: Ejemplo de creación de la tabla en la que se guardan los datos cabecera de cada versión bíblica

CREATE TABLE biblia (bible_id, collection, biblename, register);			
1	romanceamientos	Fazienda	...
9	romanceamientos	E3	...
108	ladinamientos	Asá	...
109	ladinamientos	Ḥayim	...
110	ladinamientos	Ḥéšec Šelomó	...
111	ladinamientos	Schauffler	...
1000	referencia	Hebreo	...
1001	referencia	Vulgata	...

Tabla 7: Ejemplo de creación de la tabla que recoge el texto bíblico hebreo y su transcripción normalizada

CREATE TABLE Hebreo (bible_id, book_id, book_order, chapter_id, verse_id, verse_text, verse_transcription, notes);							
1000	01	01	01001	01001001	בְּרֵאשִׁית בְּרָא אֱלֹהִים אֶת הַשָּׁמַיִם וְאֶת הָאָרֶץ:	Bere'shit bará' 'Elohim 'et hašamáyim ve'et ha'areš	...

¹⁴ Para una discusión de las ventajas de una arquitectura SQL relacional para la creación y explotación de corpus diacrónicos, véase Davies (2005, 2010).

4. EXPLOTACIÓN Y HERRAMIENTAS INFORMÁTICAS

Como acabamos de ver, en la implementación nos hemos centrado en la creación de una base de datos relacional que nos permite el uso informático o computacional instantáneo de los textos. Para ello hemos desarrollado y adaptado una serie de programas que automatizan las labores de segmentación, normalización, categorización, lematización, alineación y exportación a los formatos estándar de los que hemos hablado antes¹⁵.

Después de realizar una primera anotación automática de los textos, se procede a una revisión manual de las etiquetas asignadas. Para ello, hemos desarrollado un programa gráfico, con diferentes módulos de trabajo, que permiten la revisión detallada de los textos a nivel de palabra y de las propuestas de alineación. Hasta el momento se han incorporado los módulos de categorización, normalización, lematización, alineación y consulta. Uno de los objetivos fundamentales de nuestro grupo de investigación es la edición crítica y paralela de los textos medievales y sefardíes con sus fuentes originales y los comentarios relevantes, por lo que también se ha incorporado un módulo de presentación de triple acceso a los textos: facsímil, transcripción ajustada al original y edición crítica normalizada.

Mediante el uso de algunas imágenes tomadas del programa *Bibli@teca*, ejemplificamos en este punto algunas de las posibilidades actuales de explotación interna del corpus: desde las más obvias, como búsquedas y comparaciones, a las más específicas como el análisis del método de traducción o la asistencia en la elaboración de ediciones críticas, paralelas y globales.

4.1. Comparaciones

Gracias a la marcación estructural que hemos realizado a nivel de versículo pueden mostrarse en paralelo todas las versiones incorporadas. La Figura 12 muestra las versiones en presentación paleográfica (junto a las fuentes hebrea, latina o griega en su grafía original):

¹⁵ Para la etiquetación automatizada de los textos hispánicos se está llevando a cabo la adaptación al judeoespañol de los módulos de español (diccionarios, reglas morfológicas, corpus de entrenamiento, predicción estadística, etc.) de la herramienta de procesamiento lingüístico *Freeling*; véase Padró et al. (2010) y Sánchez-Marco et al. (2011).

**Figura 12: Bibli@teca. Módulo de comparaciones:
presentación paleográfica de la consulta**

Ediciones | Búsquedas y comparaciones | Alineación del hebreo y las versiones | Categor. y lemat. del hebreo | Categor., normal. y l

Búsqueda sencilla

Elegir versiones

Fazienda Ajuda Evora PC
 E6 Alba E7 Sal1540 Asa CantSalonica
 E8 BNM E19 PA Hayim
 GE E4 RAH CantAmsterdam HS
 E3 E5 Ferrara CantLiorna Sch
 Mostrar también fuentes y referencias

Elegir pasaje

Libro:
 Capítulo:
 Versículos:
 Buscar en las ediciones críticas

Pasaje	Asa	E6	Hebreo	Vulgata
Ca 1:2	2 Beáseme de besos de su boca, que mejores tus querencias más que vino.	2 {#r ¶ 1} Bese me con el beso de su boca. Ca meiores son las tus tetas que el uino.	יִשְׁקֵנִי מִנְּשִׁיקוֹת פִּיהוּ כִּי-טוֹבִים הֵדִיב מִיַּיִן:	osculetur me osculo oris sui quia meliora sunt ubera tua vino

La Figura 13 muestra la opción de recuperar el pasaje en modo de presentación crítica, junto a las fuentes en transcripción normalizada:

Figura 13: Bibli@teca. Módulo de comparaciones: presentación crítica de la consulta

Pasaje	Asa	E6	Hebreo	Vulgata
Ca 1:2	2 Beśáseme de beśos de su boca, que mejores tus querencias más que vino.	2 Béseme con el beso de su boca, ca mejores son las tus tetas que el vino.	yišāqēni mi-nēšiqôt pihû kî ṭôbîm dôdējā mi-yāyin	osculetur me osculo oris sui quia meliora sunt ubera tua vino

También es posible añadir a la comparación los comentarios y glosas correspondientes que se hayan incorporado hasta ese momento (Figura 14):

Figura 14: Bibli@teca. Módulo de comparaciones: glosas y comentarios

Búsqueda sencilla

Elegir versiones

Fazienda Ajuda Evora PC CantSalonica
 E6 Alba E7 Sal1540 Asa
 E8 BNM E19 PA Hayim
 GE E4 RAH CantAmsterdam HS
 E3 E5 Ferrara CantLiorna Sch
 Mostrar también fuentes y referencias

Elegir pasaje

Libro: Cantar de los cantares
 Capítulo: 1
 Versículos: 2-17
 Buscar en las ediciones críticas
 Mostrar pasaje

Griego	Targum	Rasi	Lira
φιλησάτω με ἀπὸ φίλημάτων στόματος αὐτοῦ ὅτι ἀγαθοὶ μαστοὶ σου ὑπὲρ οἴνον	yīšāqēnī: Dijo Šelomó el profeta: Bendicho su nombre de Adonay que dio a nos ley por mano de Mošé, el escribano grande, escrita sobre dos tablas de piedra y seš ódrenes de Mišná y Talmud con meldadura, y era hablán con nos faces con faces como varón que besa a su compañero de muchedumbre de querencia, que amó a nos más que a setenta pueblos.	[1:2] ¡Oh, que él me besara con los besos de su boca! Este canto es bello en su exilio y en su viudez: «Sólo que (literalmente: 'quien me diera') el rey Salomón quisiera besarme con los besos de su boca como antaño, porque hay lugares donde besan en el revés de la mano o en el hombro; pero yo deseo y anhelo que él se comporte conmigo como se comportaba al principio, como el novio con la novia. Besos a besos...	[1:2] Osculetur me, etc.] Expedito primo Salomonis libro in quo traduntur documenta quae faciunt ad illustrationem mentis, et secundo in quo inducimur ad contemptum affluentiae mobilis, hic qonsequenter incipit tertius inducens ad amorem supernae felicitatis; sicut dictum fuit plenius in principio primi libri, ubi posu quandam praefationem pro istis tribus libris. Appetitus suae felicitatis supernae...

4.2. Búsquedas

Gracias al módulo de consultas, pueden realizarse búsquedas sencillas por vocablo, restringiendo su aparición a una versión en concreto o varias. La Figura 15 muestra los resultados obtenidos al buscar el término *que* simultáneamente en la versión medieval de E6 y en la sefardí de Asá.

**Figura 15: Bibli@teca
Módulo de búsqueda**

Pasaje	Asa	E6
	más que vino, derechdades te amaron.	ementando de las tus tetas más que del vino; amante los derechureros
Ca 1:6	6 Non escarnescadeš de mí, que yo denegrída, que me desfetijó el sol; hijos de mi madre erēcieron en mí, pusiéronme guardadera a las viñas, mi viña que a mí non guardí.	6 No asmedes que só negra, ca el sol me descoloró; los fijos de mi madre lidiaron contra mí, pusiéronme por guarda en las viñas e non guardé la mi viña.
Ca 1:7	7 Denucia a mí a el que amó mi alma, cómo pacerás, cómo yaçerás en las siestas, que ¿por qué seré como enyuelta sobre rebaños de tus compañeros?	7 Demuéstrame el que ama la mi alma, ó pasces, ó yazes a mediodía; que non vaya debalde empós las greyes de tus compañeros.

Aunque el objetivo final, para el que estamos desarrollando un nuevo módulo de la aplicación, es la posibilidad de realizar búsquedas complejas que tengan en cuenta no solo el texto sino toda la información gramatical y lexicográfica que ya contiene el corpus. El nuevo módulo posibilitará al menos las siguientes búsquedas adicionales:

- Expresiones regulares {en .* prin[a-zAz]ipio}
- Especificación de categorías gramaticales {cat=V}
- Lemas {lema=crear}

- d) Operadores lógicos {AND / OR / NOT}
- e) Operadores de proximidad {NEXT / WITHIN}
- f) Búsquedas anidadas {(principio AND (cielo* OR tierra))}
- h) La combinación de todos los anteriores y el filtrado de los resultados

4.3. Análisis de traducción

Uno de los objetivos finales de la creación de un corpus paralelo de versiones bíblicas es el de posibilitar un estudio sistemático de los métodos seguidos por los traductores de las diferentes versiones. La arquitectura del corpus, en particular su completa marcación gramatical, lematización y alineación a nivel de palabra, permite realizar valoraciones no solo cualitativas sino cuantitativas sobre las elecciones léxicas o gramaticales de los traductores. Los diversos módulos de *Bibli@teca* dan acceso a algunas de las consultas necesarias para llevar a cabo el análisis comparado de dichas elecciones.

Los módulos de categorización tanto del hebreo como de los textos hispánicos (mostrados en las Figuras 10 y 11) permiten realizar o enmendar manualmente el proceso de categorización. De igual modo hemos desarrollado un módulo que permite realizar, comparar o enmendar el proceso de alineación de todas las versiones hispánicas con el hebreo (Figura 16):

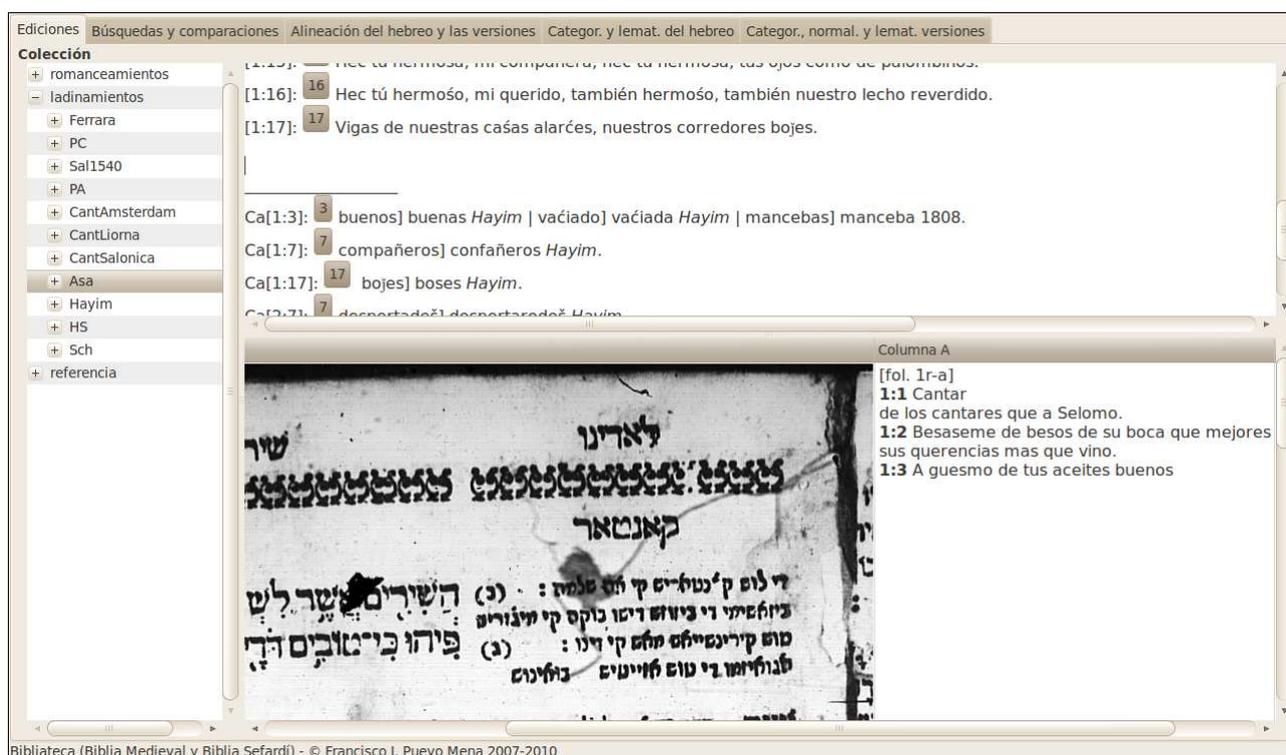
Figura 16: Bibli@teca
Módulo de alineación a nivel de palabra

The screenshot shows the 'Alineación del hebreo y las versiones' module. It features a navigation bar with tabs: 'Ediciones', 'Búsquedas y comparaciones', 'Alineación del hebreo y las versiones', 'Categor. y lemat. del hebreo', 'Categor., normal. y lemat. versiones', and 'Otros'. The main content area is titled 'Módulo de coordinación con el hebreo' and is divided into three sections: 'Elegir versiones', 'Elegir pasaje', and 'Ayuda'. The 'Elegir versiones' section has checkboxes for various editions: Fazienda, Ajuda, Evora, PC, CantSalonica, E6, Alba, E7, Sai1540, Asa, E8, BNM, E19, PA, Hayim, GE, E4, RAH, CantAmsterdam, HS, E3, E5, Ferrara, CantLiorna, and Sch. The 'Elegir pasaje' section has dropdown menus for 'Libro' (Cantar de los cantares), 'Capítulo' (1), and 'Versículos' (2), along with a 'Mostrar pasaje' button. The 'Ayuda' section provides instructions on how to use the module. Below these sections, a table displays the alignment of the Hebrew word 'ישקני' (yishqēnî) with its Spanish translation 'Beśáseme' and other related terms like 'de', 'boca', 'mejores', 'tus', 'más', 'que', and 'vino'. The table includes columns for the Hebrew word, its transliteration, and its meaning in Spanish. A detailed linguistic note for the Hebrew word 'ישקני' (yishqēnî) is provided at the bottom, explaining its usage in various biblical contexts and its relationship to other words like 'שָׁקַן' (shakan) and 'שָׁקַן' (shakan).

4.4. Ediciones electrónicas y ediciones críticas

Para finalizar este recorrido panorámico sobre el corpus bíblico sefardí me gustaría destacar que nuestro grupo de investigación se especializa en la edición crítica de textos sefardíes, por lo que resulta fundamental el poder realizar, a partir de los materiales incorporados al corpus, ediciones críticas de los textos bíblicos. El módulo de triple presentación de *Bibli@teca* nos permite realizar una edición electrónica de los textos (Figura 18):

Figura 18: Bibli@teca
Módulo de edición electrónica (triple presentación)



La Figura 18 muestra la división del programa en tres paneles diferentes. El panel izquierdo permite navegar cómodamente por la colección textual del corpus y elegir el pasaje que deseamos consultar (por versículo [*Cantar* 1:1-3] o por hoja del original del volumen [1ra]). En la columna izquierda del panel inferior se presenta la reproducción facsimilar del original, lo cual permitirá a los investigadores discutir la validez de las transcripciones que se proponen. En la columna derecha del mismo panel y de forma paralela al facsímil se muestra la transcripción seudo-paleográfica (en el sentido de que se mantienen los cortes de línea y la separación y unión de las palabras, pero no la grafía aljamiada del original, que en cualquier caso puede consultarse en el facsímil). Finalmente, en el panel superior se muestra la edición crítica normalizada del pasaje seleccionado y el aparato crítico y las notas que se hayan preparado hasta el momento.

Más allá de la edición electrónica, la alineación a nivel de palabra del corpus permite la colección exhaustiva de las variantes de traducción empleadas por los diversos traductores a lo largo de la historia de las traducciones judías al castellano y al ladino. Ejemplos del

aprovechamiento de dicha colección de variantes de traducción pueden encontrarse en Pueyo Mena (2011, 2012), en las que tomando como texto base un fragmento del glosario bíblico *Ḥēšec Šelomó* (Gé caps. 1-10) y de la traducción de Abraham Asá del libro de *Rut*, se presentan, comparan y discuten las diferentes opciones adoptadas por los traductores.

En el futuro esperamos que el corpus y las herramientas asociadas al mismo puedan servir para la realización de una edición global de la *Biblia hispanojudía medieval y sefardí*, en la línea de proyectos como el emprendido por el *Corpus Biblicum Catalanicum* para las traducciones catalanas de la Biblia (véase, por ejemplo, Riera i Sans et al., 2004).

En este sentido, se está realizando un trabajo experimental de edición paralela del *Cantar de los cantares*, mediante la elaboración de un subcorpus integral en el que se han incorporado las versiones hispanojudías medievales, las versiones sefardíes y algunos comentarios relevantes.

La elección del *Cantar de los cantares* no es aleatoria: al tratarse del libro con mayor número de versiones (tanto en español medieval como en ladino), y con mayor difusión de cuantos conforman la Biblia hebrea, su elección nos permite valorar las posibilidades de explotación e integración del corpus, así como el grado de complejidad y rendimiento de su arquitectura. En la Figura 19 se recoge una página representativa de dicha edición.

Figura 19: Edición experimental del Cantar de los cantares (1:1-2): Versiones medievales, sefardíes y comentarios

Table with 12 columns: E1, GE1, Vulgata, TM, E2, BNM, Alfo, ES, Ferrara, Ferrar, HSec Selami, Asi-Hayim, Schuffner, Ven-Sal, Liorna, Parafraiss. Each column contains a small snippet of text from a different manuscript version.

Glora de Moé Arrogé... Nicolás de Lira... Targum Asi... Rashi... This section contains the beginning of the text from various editions, including the Vulgate, medieval Hebrew, and commentaries by Nicolás de Lira, Targum Asi, and Rashi.

Quodam die... This section continues the text from the Vulgate and other medieval editions, starting with the phrase 'Quodam die'.

Et in diebus illis... This section continues the text, starting with 'Et in diebus illis', and includes a large block of text in Hebrew script with a Latin translation.

Et in diebus illis... This section continues the text, starting with 'Et in diebus illis', and includes a large block of text in Hebrew script with a Latin translation.

Notes... This section contains a list of notes and references for the text, including manuscript identifiers like 'LANGE H. POS. P. TYPE. C. ASPECT.' and 'MORPHO 220001004 - MORPH 1 - 1:1-2'.

Como puede observarse en la imagen, pretendemos ofrecer de forma paralela la edición crítica de las dieciséis versiones medievales y sefardíes que se han incorporado de momento al corpus. Cada versión lleva su propio aparato crítico, que tiene en cuenta todos los manuscritos y ediciones que ofrecen variantes del mismo texto. En la franja central de la página, se presentan algunos comentarios bíblicos que pueden ofrecer explicaciones exegéticas sobre las elecciones léxicas de los traductores. En la franja inferior de la página se recogen de forma alineada las lecturas hispánicas, agrupadas bajo la palabra hebrea correspondiente, tal y como se muestra en el ejemplo de la Tabla 12, correspondiente al versículo 1:16.

Tabla 12: Alineación a nivel de palabra de todas las versiones bíblicas (*Cantar 1:16*)

^{16]} hinejá yafé dodí 'af na'ím 'af-'arsenu ra'ananá הַנֶּהָדָה יָפֵה דוֹדִי אֶף נָעִים אֶף-אַרְשֵׁנוּ רַעֲנָנָה:	
<hinejá> הַנֶּהָדָה	O qué <i>E6</i> • Ahé <i>GE3</i> • Hete <i>E3</i> • Hete <i>BN/E4</i> • Cuán <i>E5</i> • Ahé que tú <i>Arr.</i> • Hé tú <i>Ferr.</i> • He aquí que tú eres <i>Oso</i> • \emptyset <i>HS</i> • Hec tú <i>Asá</i> • Hec tú <i>Hayim</i> • Hec tú <i>Sch</i> • Hec tú eres <i>Misión</i> • Hec tú <i>Sal.</i> • Heg tú <i>Liorn.</i> • He tú <i>Paráfr.</i>
<yafé> יָפֵה	fermoso eres <i>E6</i> • fermoso eres tú <i>GE3</i> • fermoso <i>E3</i> • fermoso <i>BN/E4</i> • fermoso eres <i>E5</i> • eres fermoso <i>Arr.</i> • hermoso <i>Ferr.</i> • hermoso <i>Oso</i> • \emptyset <i>HS</i> • hermošo <i>Asá</i> • hermošo <i>Hayim</i> • hermošo <i>Sch</i> • hermošo <i>Misión</i> • hermošo <i>Sal.</i> • hermošo <i>Liorn.</i> • hermoso <i>Paráfr.</i>
<dodí> דוֹדִי	mio amigo <i>E6</i> • mio amado <i>GE3</i> • mi enamorado <i>E3</i> • mi amigo <i>BN/E4</i> • el mi querido <i>E5</i> • el mi querido <i>Arr.</i> • mi querido <i>Ferr.</i> • o amado mío <i>Oso</i> • \emptyset <i>HS</i> • mi querido <i>Asá</i> • mi querido <i>Hayim</i> • mi querido <i>Sch</i> • querido mío <i>Misión</i> • mi querido <i>Sal.</i> • mi querido <i>Liorn.</i> • mi querido <i>Paráfr.</i>
<'af> אֶף	e qué <i>E6</i> • e <i>GE3</i> • e aún <i>E3</i> • e aún <i>BN/E4</i> • otrosí <i>E5</i> • e también <i>Arr.</i> • también <i>Ferr.</i> • también <i>Oso</i> • \emptyset <i>HS</i> • también <i>Asá</i> • también <i>Hayim</i> • también <i>Sch</i> • también <i>Misión</i> • también <i>Sal.</i> • también <i>Liorn.</i> • también <i>Paráfr.</i>
<na'ím> נָעִים	apuesto <i>E6</i> • fermoso <i>GE3</i> • donoso <i>E3</i> • excelente <i>BN/E4</i> • loçano <i>E5</i> • gracioso <i>Arr.</i> • suave <i>Ferr.</i> • suave <i>Oso</i> • gracioso <i>HS</i> • hermošo <i>Asá</i> • hermošo <i>Hayim</i> • gracioso <i>Sch</i> • gracioso <i>Misión</i> • hermošo <i>Sal.</i> • suave <i>Liorn.</i> • suave <i>Paráfr.</i>
<'af> אֶף	\emptyset <i>E6</i> • \emptyset <i>GE3</i> • aún <i>E3</i> • e aún <i>BN/E4</i> • eso mismo <i>E5</i> • e <i>Arr.</i> • también <i>Ferr.</i> • también <i>Oso</i> • \emptyset <i>HS</i> • también <i>Asá</i> • también <i>Hayim</i> • también <i>Sch</i> • también <i>Misión</i> • también <i>Sal.</i> • también <i>Liorn.</i> • también <i>Paráfr.</i>
<'arsenu> אֶרְשֵׁנוּ	El nuestro lecho <i>E6</i> • el nuestro lecho <i>GE3</i> • nuestra cama <i>E3</i> • nuestro lecho <i>BN/E4</i> • la nuestra cama <i>E5</i> • el nuestro lecho <i>Arr.</i> • nuestro lecho <i>Ferr.</i> • nuestro lecho <i>Oso</i> • nuestro lecho <i>HS</i> • nuestro lecho <i>Asá</i> • nuestro lecho <i>Hayim</i> • nuestro lecho <i>Sch</i> • nuestro lecho <i>Misión</i> • nuestro lecho <i>Sal.</i> • nuestro lecho <i>Liorn.</i> • nuestro lecho <i>Paráfr.</i>
<ra'ananá> רַעֲנָנָה	es florido <i>E6</i> • es florido <i>GE3</i> • es florida <i>E3</i> • reverdido <i>BN/E4</i> • muy adobada <i>E5</i> • está reverdescido <i>Arr.</i> • florido <i>Ferr.</i> • florido <i>Oso</i> • reverdido <i>HS</i> • reverdido <i>Asá</i> • reverdido <i>Hayim</i> • reverdido <i>Sch</i> • reverdido <i>Misión</i> • reverdido <i>Sal.</i> • reverdecido <i>Liorn.</i> • reverdecido <i>Paráfr.</i>

Obras citadas

- Cantera, Francisco & Iglesias, Manuel. *Sagrada Biblia. Versión crítica sobre los textos hebreo, arameo y griego*. 3ª ed., Madrid: Biblioteca de Autores Cristianos, 2000.
- Davies, Mark. «The advantage of using relational databases for large corpora: speed, advanced queries, and unlimited annotation». *International Journal of Corpus Linguistics* 10 (2005): 301-328.
- . «Creating Useful Historical Corpora: A Comparison of CORDE, the Corpus del Español, and the Corpus do Português». En Andrés Enrique-Arias (ed.), *Diacronía de las lenguas iberorromances: nuevas perspectivas desde la lingüística de corpus*. Frankfurt / Madrid: Vervuert / Iberoamericana, , 2010: 137-166.
- Enrique-Arias, Andrés (dir.). *Corpus Biblia Medieval*, 2010-. En línea en <http://www.bibliamedieval.es>.
- García Moreno, Aitor (dir.). *Corpus Histórico Judeoespañol (CORHIJE)*, 2012-. En línea en <http://www.recursos.esefardic.es/corhije/>.
- García Moreno, Aitor & Pueyo Mena, F. Javier. «Towards a representative corpus of the history of Judeo-Spanish (the CORHIJE)». En Ana Stulic-Etchevers & Soufiane Rouissi (eds.), *Le numérique au service du patrimoine culturel séfardé*. Bordeaux: Presses Universitaires de Bordeaux, 2014 (en prensa).
- Hassán, Jacob M. «Transcripción normalizada de textos judeoespañoles». *Estudios Sefardíes* 1 (1978): 147-150.
- . «Sistemas gráficos del español sefardí». En M. Ariza & A. Salvador & A. Viudas (eds.), *Actas del I Congreso Internacional de Historia de la Lengua Española*. Madrid: Arco Libro, 1988: vol. I, 127-137.
- . «El español sefardí (judeoespañol, ladino)» (1) (y 2). [Ensayo «La lengua española hoy» XXIII]. *Boletín Informativo Fundación Juan March* 243 (oct. 1994): 3-16; 244 (nov. 1994): 3-18.
- . «El español sefardí (judeoespañol, ladino)» [versión revisada]. En M. Seco & G. Salvador (coords.), *La lengua española, hoy*. Madrid: Fundación Juan March, 1995: 117-140.
- . «El estudio del ladino: entre la tradición española y la tradición israelí». En Shmuel Refael (ed., con la col. de Judith Dishon), *Ladinar* 4 (= Actas del Primer encuentro académico programático: El español saluda a judeoespañol ..., 19-21.12.2004), 2006: 43-55.
- . «Sistemas gráficos del español sefardí». En Elena Romero (ed.), I. M. Hassán & Ricardo Izquierdo Benito (coords.), *Sefardíes: Literatura y lengua de una nación dispersa*, XV Curso de Cultura Hispanojudía y Sefardí de la Universidad de Castilla-La Mancha. Cuenca: Ed. de la Univ. de Castilla-La Mancha, 2008: 119-144.
- . «Ortografía latinada hispánica (por definición) del ladino (español sefardí o judeoespañol) para el siglo XXI. Comunicación presentada en el VI Congreso de la *European Association for Jewish Studies*, 1998 (inédito).
- Kohring, Heinrich. «“Kale enladinar komo uzamos a avlar”: La technique du ladino chez Abraham Asá (1743) et Yehuda Alkalai (1839)». En W. Busse & M.-C. Varol Bornes (eds.), *Hommage à Haim Vidal Sephiha*. Berna: Peter Lang, 1996: 315-335.
- Lazar, Moshe. «Tirgumé haMicrá beladino [‘Las traducciones de la Biblia en ladino’]». *Sefunot* 8 (1964): 347-375.
- . «Apéndice: Ladinamientos aljamiados de la Biblia». En Jacob M. Hassán & Ángel

- Berenguer (eds.), *Actas del Simposio Internacional sobre la Biblia de Ferrara*. Madrid: CSIC, 1994: 373-442.
- Padró, Ll. & Collado, M. & Reese S. & Lloberes M. & Castellón, I. FreeLing 2.1: «Five Years of Open-Source Language Processing Tools». En *Proceedings of 7th Language Resources and Evaluation Conference (LREC 2010)*. La Valletta, Malta, 2010. En línea en <<http://nlp.lsi.upc.edu/freeling>>.
- Pueyo Mena, F. Javier. «Biblias romanceadas y en ladino». En Elena Romero (ed.) & Iacob M. Hassán & Ricardo Izquierdo Benito (coords.), *Sefardíes: Literatura y lengua de una nación dispersa*. Cuenca: Univ. de Castilla-La Mancha, 2008: 193-263.
- . «Séfer Héšec Šelomó: edición de *Génesis / Berešit* (cps. 1-10)». En Elena Romero (ed.), *Estudios sefardíes dedicados a la memoria de Iacob M. Hassán (z"l)*. Madrid: CSIC (PES núm. 13), 2011: 433-478.
- . «“Tu pueblo mi pueblo y tu Dio mi Dio”: el libro de *Rut* en la Biblia de Abraham Asá», *eHumanista* 20 (2012): 263-295.
- . «Méthode traditionnelle de traduction et alignement automatisé des sources pour le corpus biblique séfaraide». En Soufiane Rouissi & Ana Stulic-Etchevers (eds.), *Recensement, analyse et traitement numérique des sources écrites pour les études séfaraides*. Bordeaux: Presses Universitaires de Bordeaux, Collection de la Maison des Pays Ibériques, Série Littéralité, 2013: 343-365.
- . «Problemas de representación del ladino bíblico y de sus fuentes en el *Corpus Bíblico Sefardí*». En Ana Stulic-Etchevers & Soufiane Rouissi (eds.), *Le numérique au service du patrimoine culturel séfaraide*. Bordeaux: Presses universitaires de Bordeaux, 2014 (en prensa).
- Riera i Sans, Jaume & Casanellas i Bassols, Pere & Puig i Tàrrrech, Armand. *Biblia del segle XIV: Èxode. Levític*. Barcelona: Associació Bíblica de Catalunya & Publicacions de l'Abadia de Montserrat, 2004.
- Sánchez-Marco, C. & Boleda, G. & Padró, Ll. «Extending the tool, or how to annotate historical language varieties». En *Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities. LaTeCH (24 June, 2011, Portland, Oregon): Proceedings of the Workshop*. The Association for Computational Linguistics, Omnipress Inc., 2011: 1-9.