

METODOLOGIA INFORMATICA PARA LA EDICION DE TEXTOS

Exponemos a continuación un conjunto de programas que se han creado para la edición de textos medievales, conjunto que está a disposición de los investigadores⁽¹⁾.

Transcripción

Recomendamos que, en la transcripción del texto, se sigan las normas del Seminario de Estudios Medievales de Madison y que se eviten todos los signos que no forman parte del teclado básico (es decir, el de una máquina de escribir con teclado inglés) como tildes y acentos, sustituyéndolos por signos básicos: >n´ = ñ<, >:´< = 'sigma', >a = á<, etc.. Téngase en cuenta que es mucho más sencilla la re conversión desde signos de este teclado, visibles en cualquier sistema, que desde unos hexadecimales peculiares de uno de ellos, que pueden sufrir transformaciones impredecibles en los procesos de copia o traslado de un sistema a otro.

Observada la norma de transcripción o, en su defecto, hechos explícitos los criterios de la misma, si difieren, para que se puedan hacer las transformaciones oportunas, debe tenerse en cuenta que el conjunto de programas de UNIFICACION TEXTUAL (UNITE) puede actuar sobre textos en verso cuyo formato puede disponer libremente el investigador, siempre que esa disposición sea seguida regularmente en todo el texto.

El paquete de ficheros UNITE

Este conjunto de ficheros se puede dividir en dos grupos, según realicen la entrada de datos desde la pantalla o desde el fichero DAT_UNIT.DATOS. Tanto uno como otro utilizan los ficheros llamados FTExi SCRIPT (1 < i < 7), en los que están almacenadas las versiones que se van a comparar.

Los ficheros FTExi SCRIPT A contienen las versiones que se van a comparar en el proceso de unificación automática. El máximo de versiones que se permiten

en el programa es 6, por tanto, existen seis ficheros FTEXi (FTEX1...FTEX6). Dentro de este límite se utilizarán aquellos que sean necesarios, según el número de textos que se quieran comparar. La prioridad de cada versión a la hora de la comparación depende del nombre del fichero en el que esté almacenada y va en orden descendente desde el FTEX1 hasta el FTEX6.

Reglas para el almacenamiento de las versiones

- * Las tres primeras líneas de los ficheros no serán tenidas en cuenta por el programa; pueden ser utilizadas por el usuario para identificar la versión que se va a almacenar en el fichero.
- * Los textos deben estar compuestos por conjuntos de versos a los que, por comodidad, llamaremos *estrofas*, cada una de las cuales puede tener un número variable de versos. Por ahora el máximo de versos está fijado en 5, el mínimo es 1.
- * Entre estrofa y estrofa debe haber una línea en blanco de separación y sólo una.
- * En el final del fichero puede existir como máximo una línea en blanco a partir de la última estrofa.

Formato de la estrofa

- * Cada estrofa debe tener un número asignado que figurará al menos en la línea correspondiente al primer verso de dicha estrofa.
- * La distancia desde el principio de la línea a la numeración, la longitud máxima de dicha numeración, así como la distancia desde ésta al principio del verso propiamente dicho son optativas, pero constantes. Es muy importante que estos valores se correspondan con los aportados al programa en DAT_UNIT o en la pantalla.
- * El verso propiamente dicho debe comenzar en la misma columna en todos los versos de la estrofa, es decir en aquella correspondiente a la suma de los parámetros: margen antes de la numeración, longitud máxima de ésta y distancia de dicha numeración al principio del verso.
- * Las palabras pertenecientes a un verso deben estar separadas por un único

NOTAS

espacio en blanco, ya que el fin del verso se detecta cuando se llega al fin de línea o a dos espacios blancos seguidos.

- * Si falta un verso completo en una estrofa, esta línea no se debe dejar en blanco, sino rellenarse con los caracteres:.-.-. (punto y guión).
- * No es necesario que las estrofas estén ordenadas según su numeración, ya que se puede utilizar para ordenarlas el módulo SORT_EST.
- * El programa UNITE compara los versos en el interior de la estrofa, aunque estén en orden distinto.

Como paso previo a la ejecución de cualquiera de los dos grupos de programas, se debe utilizar el módulo SORT_EST MODULE.

Descripción del programa SORT_EST

Función

Este programa ordena de menor a mayor las estrofas almacenadas en un fichero. Si dos estrofas tienen el mismo número las deja tal y como están almacenadas.

Condiciones para su funcionamiento

El fichero cuyas estrofas se van a ordenar debe estar declarado como DESORD SCRIPT A; también debe haber declarados dos ficheros llamados FICH1 SCRIPT A y FICH2 SCRIPT A con el contenido que se quiera, éstos son utilizados como ficheros auxiliares por el programa. Una vez acabada la ejecución, los ficheros auxiliares desaparecen, mientras que el fichero DESORD SCRIPT A contiene las estrofas ordenadas.

El número de estrofa en el fichero DESORD debe estar situado en las cuatro primeras posiciones de la línea correspondiente.

Al principio del fichero no debe haber ninguna línea en blanco y al final sólo una.

Modo de empleo

Una vez declarados los ficheros auxiliares y DESORD, es conveniente asegurarse de que existen estrofas descolocadas antes de ejecutar el programa; para ello se puede utilizar la macro AVISA, que muestra en pantalla el número de estrofa descolocada. Para utilizar correctamente esta macro la línea vigente debe ser el primer verso de una estrofa (normalmente de la primera del fichero). Hay que decir que AVISA considera descolocadas aquellas estrofas que tienen el mismo

Incipit, VI (1986)

número. Si detectamos estrofas descolocadas, salimos del fichero editado y ejecutamos el módulo SORT_EST con la siguiente instrucción:

SORT_EST

Quando acaba la ejecución podemos comprobar si ésta ha sido correcta volviendo a editar el fichero DESORD y utilizando de nuevo la macro AVISA.

Ejecución con el fichero DAT_UNIT DATOS

El fichero DAT_UNIT DATOS se utiliza para que el usuario pueda controlar la ejecución del programa. En este fichero se pide una serie de valores y se indica en qué lugares se deben teclear las respuestas. Conviene recordar que DAT_UNIT sólo se puede modificar en los espacios reservados para respuestas.

Además de este fichero existen también UNITE1 PASCAL y el módulo generado a partir de él, UNITE1 MODULE. Integran también el grupo de ficheros los EXEC COMPUNI1 (compila en BATCH el programa UNITE1, su ejecución no es necesaria puesto que ya está generado UNITE1 MODULE) y EJECUNI1, que ejecuta en el BATCH el módulo de UNITE1.

El Libro de Alexandre

El texto que nos ha servido para comprobar la validez de la metodología e irla desarrollando es este texto de la primera mitad del siglo XIII, cuyas dificultades para conseguir una edición unificada son muchas. Para empezar, la existencia de dos textos, el de la Biblioteca Nacional de Madrid, procedente de la biblioteca de la casa de Osuna, por lo cual se llama manuscrito O, del siglo XIV, copiado en León por Fray Lorenzo de Astorga, y el de la Bibliothéque Nationale de Paris, del s. XV, aragonés, ms. P, que el copista atribuye a Gonzalo de Berceo, célebre poeta riojano del siglo XIII, autor de obras de carácter religioso, escritas en la misma estrofa de cuatro versos de catorce sílabas (tetrástrofos monorrimos) que el *Libro de Alexandre*, y que constituyen, como este libro y el *Libro de Apolonio*, los ejemplos más antiguos de la escuela poética denominada *mesa ter de clerecía*.

El *Libro de Alexandre* se ha conservado también en algunos fragmentos menores, de escaso número de versos, que testimonian la gran extensión antigua de las familias de manuscritos, pues no se relacionan muy directamente con ninguno de los dos manuscritos largos conservados, ninguno de los cuales, por su parte está tam poco completo.

Esta complicada situación afecta a la caracterización lingüística. Emilio Alarcos, en 1948, defendió la tesis de que el original sería básicamente castellano, y que los rasgos dialectales leoneses (del manuscrito de Madrid) y aragoneses (del manuscrito de París) se deberían a los copistas. Hacia un tipo occidental peninsular, en todo caso, parecen inclinarse Joan Corominas y Yakov Malkiel, de acuerdo con la tendencia predominante de las letras castellanas, que reciben desde el oeste la influencia de mayor prestigio. La atribución a Berceo ha sido defendida recientemente por Dana Nelson. Esta tesis riojana, cuyas dificultades han sido puestas de manifiesto por Rafael Lapesa, en 1980, ha encontrado un apoyo subsidiario en la crítica de Raymond Willis a la edición de Nelson.

En lo que concierne a la fecha, tampoco hay acuerdo preciso, fuera de los límites muy generales de la primera mitad del siglo XIII. Tiene que ser anterior a 1250, fecha aproximada del *Poema de Fernán González*, en el que influye. Por el verso 860d, referido a Damietta, podría estar relacionado con sucesos de la quinta cruzada (1217), y la reconquista de esta ciudad en 1217. La estrofa 2522, en cambio, puede contener una alusión al rey de Sicilia y la cruzada de 1228. Ha de ser posterior, naturalmente, a 1182, fecha de la *Alexandreis*, poema latino de Gautier de Châtillon, de esa fecha, que en buena parte traduce.

Ejemplo: hacia una edición unificada del Libro de Alexandre

Pasos para la ejecución:

1. Preparar los ficheros FTEXi
2. Editar el fichero DAT_UNIT DATOS y modificar si es necesario las respuestas.
3. Ejecutar el programa. La ejecución se puede hacer de dos formas: en nuestra máquina virtual o utilizando el BATCH, esta última normalmente se utilizará para procesos con textos muy largos.
 - a. UNITEL para la primera
 - b. SUBMIT EJECUNIL (cl i di no sec 36 para el BATCH.
(las opciones del exec SUBMIT se pueden modificar)

El resultado de esta ejecución queda almacenado en el fichero FILE SALUNITF, que aparecerá en el FILEList, o en el READER si se ha ejecutado en el BATCH.

Ejemplo del proceso de unificación

A continuación presentamos una representación del proceso, según las suce-

sivas fases aplicadas para unificar la estrofa 51 del *Libro de Alexandre*, conservada en los manuscritos de París (1), de Madrid (2), en la versión impresa del *Victorial* de Gutierre Díez de Games (4) y en la versión manuscrita del mismo fragmento (5). Falta el número (3), que corresponde al fragmento de Medina celi, en el cual no está la estrofa que unificamos. La primera línea de cada ejemplo (que lleva la numeración 51 A/B/C/D) contiene el estado que la comparación ha alcanzado al terminar cada fase, mientras que las líneas sin numerar que la siguen contienen los elementos que todavía debe ir tratando el programa o, al final, los residuos no unificables, que el editor humano tendrá que considerar para completar la tarea. No se olvide que el proceso sigue las instrucciones de DAT_UNIT DATOS. En el primer intento tiene en cuenta el número de orden de las palabras en el verso. En lugar de separar el presunto clítico, a diferencia de programas nuestros anteriores como ALCOMCLI y UNION PASCAL, lo une si va separado, con lo que se evitan falsos cortes de >-l<, >-t<, etc. en posición final. El resto del proceso es de fácil deducción, creemos.

LIBRO DE ALEXANDRE
COMPARACION DE P, O, G' y G
estrofa 51, cuatro versiones

ETAPAS

1. Ordenar las estrofas con SORT_EST.
2. Preparar los ficheros de los textos.
3. Preparar la tabla de palabras vacías.
4. Editar el fichero DAT_UNI DATOS y cambiar, si es preciso, las variables.
5. Invocar el módulo UNITEL.

DAT_UNI DATOS

FICHERO DE ENTRADA DE DATOS

En aquellos apartados en los que se pidan datos que, por respuestas anteriores, no son necesarios, no hay que insertar nada. Este fichero sólo se puede modificar en los espacios reservados para las respuestas.

Introducir siempre el dato en el espacio indicado por - - ->

ESTABLECER SI SE DEBEN EJECUTAR LOS PROCESOS DE AYUDA A LA SELECCION

Ejecución de unión de palabras (s/n) - - -> s

Ejecución de separación de palabras (s/n) - - -> s

NOTAS

Ejecución de unifica varias posiciones de una palabra (s/n)? - - -> s
Ejecución de formación de palabras a partir de letras comunes
de otras (s/n) - - -> s

INTRODUCIR LOS SIGUIENTES DATOS:

Número de versiones originales: - - -> 6
Número de estrofa para comienzo de ejecución: - - -> 1
Número de fases de unificación (1 - 5): - - -> 5
Mínimo de palabras comunes (todas vacías) entre dos versos para no
tomarlos como descolocados: - - -> 3
Número de espacios desde el principio hasta la numeración de las
estrofas: - - -> 0
Longitud de la numeración: - - -> 4
Distancia desde el final de la numeración hasta el principio de la
estrofa: - - -> 7
Radio de acción del proceso union_de_palabras (número recomendado 4): - - -> 4
Radio de acción del proceso separacion_de_palabras (número recomendado 4): - - -> 4
Radio de acción de unifica_posiciones (se recomienda entre 4 y 6): - - -> 6
Radio de acción de letras_comunes (se recomienda 3): - - -> 3

---o0o---

Introducir el mínimo de versiones en las que debe coincidir una palabra para ser seleccionada según las versiones que se estén comparando.

Mínimo de versiones cuando comparemos 2 textos: - - -> 2
Mínimo de versiones cuando comparemos 3 textos: - - -> 2
Mínimo de versiones cuando comparemos 4 textos: - - -> 3
Mínimo de versiones cuando comparemos 5 textos: - - -> 3
Mínimo de versiones cuando comparemos 6 textos: - - -> 4

---o0o---

Introducir los códigos de las fases de unificación en el orden que se desee.
Las fases van numeradas del 1 al 15 y son las siguientes:

1 : Fase inactiva
2 : Elimina signos de puntuación
3 : Cambio de mayúsculas y vocales acentuadas
4 : Unificación de caracteres y grupos de caracteres hacia formas generales
5 : Simplificación de letras consecutivas iguales dentro de una palabra

El número que sigue a la + es el código de la fase. El orden de aplicación depende de la voluntad del usuario. Ejemplificamos con el orden normal, creciente, de 1 a 5, pero nada impediría un orden decreciente, del 5 al 1, o mixto (3,4,2,5,1). Téngase en cuenta que la fase 1 es útil sólo cuando se aplica en primer lugar.

Código de la 1 fase aplicable: - - -> 1
Código de la 2 fase aplicable: - - -> 2
Código de la 3 fase aplicable: - - -> 3
Código de la 4 fase aplicable: - - -> 4
Código de la 5 fase aplicable: - - -> 5

EL PROGRAMA UNITE1 PASCAL

UNITE1

PROCESO DE UNIFICACION DE LOS SIGUIENTES VERSOS

Enpec'ol Aristotiles com m o om e bien honrrado
Comec'o don Aristotil cuemo ombr e bie n letrado
Comenz'o Aristotiles, como ome bien lenguado,
comenc'o Aristotiles como onbre bien lenguado

Ejecución del primer intento de selección.

Estado inicial del Verso_Unificado

51 A Aristotiles como bien
/1enpesol/ /14ome/ /1onrado/
/2comeso/ /2don/ /2aristotil/ /2cuemo/ /2omure/ /2bien/ /2letrado/
/45comenso/ /45lenguado/
 /5onure/

Ejecución de Unifica_Posiciones

51 A Aristotiles como bien
/1enpesol/ /14ome/ /1onrado/
/2comeso/ /2don/ /2aristotil/ /2cuemo/ /2omure/ /2letrado/
/45comenso/ /45lenguado/
 /5onure/

Ejecución de Letras_Comunes Fase= 1

51 A come_so Aristotiles como bien
/1enpesol/ /14ome/ /1onrado/
 /2don/ /2aristotil/ /2cuemo/ /2omure/ /2letrado/
 /45lenguado/
 /5onure/

NOTAS

Ejecución de Letras_Comunes Fase= 2

51 A come_so Aristotiles como om__e bien
/1enpesol/ /1onrado/
/2don/ /2aristotil//2cuemo/ /2letrado/
/45lenguado/
/5onure/

Ejecución de Letras_Comunes Fase= 3

51 A come_so Aristotiles como om__e bien le__ado
/1enpesol/ /1onrado/
/2don/ /2aristotil//2cuemo/
/5onure/

PROCESO DE UNIFICACION DE LOS SIGUIENTES VERSOS

fijo dixol en buena hedat sodes vos llegado

fijo a bona edat sodes embiado

é dixo: fijo, á buena edad eres llegado

e dixo fijo a buena hedad heres llegado

Ejecución del primer intento de selección.

Estado inicial del Verso_Unificado

51 B llegado
/12fijo//1dixol//1en//1buena//1edat//1sodes//1uos/
/2a/ /2uona//2edat//2sodes//2enuiado/
/45e/ /45dixo//45fijo//45a//45buena//45edad//45eres/

Ejecución de Unifica_Posiciones

51 B fijo a buena eda llegado
/1dixol//1en/ /12edat//12sodes//1uos/
/2uona/ /2enuiado/
/45e//45dixo/ /45edad//45eres/

Ejecución de Letras_Comunes Fase= 1

51 B fijo dixo__e__a buena eda__ llegado
/12sodes//1uos/
/2uona/ /2enuiado/
/45eres/

PROCESO DE UNIFICACION DE LOS SIGUIENTES VERSOS

pora seyer om e bueno tu lo as aguisado
de seer ombre bueno tenedes lo bien guisado
de seer ome bueno: fazlo bien á guisado,
de ser hombre bueno faslo bien aguisado

Ejecución del primer intento de selección.

Estado inicial del Verso_Unificado

51 C de ser bueno
/1pora//1seier//14ome//1tu//12lo//1as//1aguisado/
 /2omure/ /2tenedes/ /2bien//24guisado/
 /45faslo//45bien//4a/
 /5onure/ /5aguisado/

Ejecución de Unifica_Posiciones

51 C de ser bueno bien
/1pora//1seier//14ome//1tu//12lo//1as//15aguisado/
 /2omure/ /2tenedes/ /24guisado/
 /45faslo/ /4a/
 /5onure/

Ejecución de Une_Palabras Fase= 1

51 C de ser bueno bien aguisado
/1pora//1seier//14ome//1tu//12lo//1as/
 /2omure/ /2tenedes/
 /45faslo/
 /5onure/

Ejecución de Separa_Palabras Fase= 1

51 C de ser bueno lo bien aguisado
/1pora//1seier//14ome//1tu/ /1as/
 /2omure/ /2tenedes/
 /45(fas)/
 /5onure/

Ejecución de Letras_Comunes Fase= 2

51 C de ser om_e bueno _as lo bien aguisado
/1pora//1seier/ /1tu/
 /2tenedes/
 /5onure/

NOTAS

PROCESO DE UNIFICACION DE LOS SIGUIENTES VERSOS

sy leuar lo qujeres com m o lo as enpec'ado
si leuar lo q ui sierdes cuemo auedes co m pec'ado
si llevarlo quisieres como has comenzado
sy llevarlo qujsjeres como as comenc'ado

Ejecución del primer intento de selección.

Estado inicial del Verso_Unificado

51 D si
/12leuar//12lo//lquieres//lcomo//1lo//las//1enpesado/
/2quisierdes//2cuemo//2auedes//2conpesado/
/45leuarlo//45quisieres//45como//45as//45comensado/

Ejecución de Unifica_Posiciones.

51 D si como as
/12leuar//12lo//lquieres//1lo/ /1enpesado/
/2quisierdes//2cuemo//2auedes//2conpesado/
/45leuarlo//45quisieres//45comensado/

Ejecución de Une_Palabras Fase= 1

51 D si leuarlo como as
/lquieres//1lo/ /1enpesado/
/2quisierdes//2cuemo//2auedes//2conpesado/
/45quisieres/ /45comensado/

Ejecución de Letras_Comunes Fase= 1

51 D si leuarlo quisier_es como as
/lquieres//1lo/ /1enpesado/
/2cuemo//2auedes//2conpesado/
/45comensado/

Ejecución de Letras_Comunes Fase= 2

51 D si leuarlo quisier_es como as co_nsado
/lquieres//1lo/ /1enpesado/
/2cuemo//2auedes/

ESTROFA UNIFICADA COMPLETA

- 51 A come_{so} Aristotiles como om_e bien le_{ado}
/lenpesol/ /lonrado/
/2don/ /2aristotil//2cuemo/
/5onure/
- 51 B fijo dixo_e a buena eda_{llegado}
/12sodes//luos/
/2uona/ /2enuiado/
/45eres/
- 51 C de ser om_e bueno_{as} lo bien aguisado
/lpora//lseier/ /ltu/
/2tenedes/
/5onure/
- 51 D si leuarlo quisier_{es} como as co_{nsado}
/lquieres//llo/ /lenpesado/
/2cuemo//2auedes/

IMPRESION DE LA ESTROFA UNIFICADA

- 51⁽²⁾ Come[n]ço Aristotil como onbre bien le[tr]lado
fiiio dixo-[l]a buena edat [eres] llegado
de seer onbre bueno [flaz-lo bien aguisado
si lleuar-lo quisieres como as començado

51a O "Começo", P "Enpeçol"; P, G' y G "Aristotiles", amétrico; cfr. como; P "omé", G "ome"; O "letrado", P "honrrado"; b O falta "dixol"; P "en"; cfr. -ue-; O y P "sodes", tratamiento incongruente con el resto del discurso; O "embiado"; c P "pora seyer", G "ser"; O "ombre", P "omé", G' "honbre", G "ome"; O "tenedes lo", P "tu lo as", G' "faslo", G' "fazlo"; "bien" falta en P; O "guisado", G "á guisado"; d O y P "leuar lo"; O "quisierdes cuemo", P "quieres como lo"; O "auedes", G "as comenzado".

NOTAS

NOTAS

1 Para ello basta con entrar en contacto con el autor, bien por correo ordinario: Ap. 46348, E-28080 Madrid, España, bien por correo electrónico: MARCOS at EEARN.EARNET (= BITNET), o MARCOS at EMDUAMI1.EARNET (= BITNET). La fórmula de cesión dependerá del tipo de colaboración que se pueda establecer. Los programas funcionan, de momento, en ordenadores IBM 43xx y superiores, en el sistema VM/CMS, pero está previsto ofrecerlos, más adelante, en versiones que se ejecuten en IBM/PC o compatibles, escritas en TURBOPASCAL. Como procesadores de textos se usan GML (Generalized Mark-up Language) y SCRIPT/VS. Este conjunto de programas ha sido realizado por Juan de Dios Godoy, en colaboración con el autor, gracias a una beca de investigación de IBM España para la realización del proyecto de investigación de la Cátedra de Lingüística General de la Universidad Autónoma de Madrid sobre tratamiento filológico informatizado de textos españoles, en el Centro de Investigación UAM-IBM.

2 Las estrofas 51-55 se conservan también en los fragmentos del *Victorial*, manuscrito (G') e impreso (G). No es posible ni aconsejable diferenciar tipográficamente cuando las coincidencias se dan con uno o con los dos fragmentos, que se remontan a un original común, del que se han diferenciado en el proceso de copia. Usamos la **negrita** cuando uno de los manuscritos largos coincide con un fragmento y la *cursiva* cuando coinciden los dos manuscritos largos entre sí.