

Guía de anotación métrica

(Versión 0.2)

Borja Navarro Colorado, María Ribes Lafoz, Noelia Sánchez
Universidad de Alicante
borja@dlsi.ua.es

30 de noviembre de 2015

Índice general

Índice general	III
1 Introducción	3
2 Aspectos generales	5
2.1. Representación de la información métrica.	5
2.2. Proceso de anotación	7
3 Reglas generales	9
3.1. Reglas para determinar acentos de intensidad.	9
3.2. Reglas para determinar el tamaño del verso.	17
3.3. Sobre la corrección del texto.	18
4 Casos problemáticos	23
4.1. “Ya que”	23
4.2. “Cada”	24
4.3. “Tan”	25
4.4. Vocativos tónicos	25
4.5. Sinalefa y “h” aspirada	26
4.6. Ambigüedad en la sinalefa.	27
Bibliografía	29

Resumen

El objeto de esta guía es exponer las reglas generales para anotar el patrón métrico de un verso, así como dar cuenta de los principales problemas de la anotación y la solución que se debe tomar. La calidad de un corpus anotado a mano reside en la consistencia de la anotación: que ante los mismos fenómenos, anotadores distintos marquen lo mismo. Mediante esta guía se pretende dar la suficiente consistencia a la anotación del corpus, de tal manera que todos los anotadores anoten lo mismo ante los mismos o similares fenómenos.

La guía de anotación se crea para el Corpus de Sonetos del Siglo de Oro. El verso endecasílabo es, por tanto, el tipo de verso principal. Todas las indicaciones están pensadas para este tipo de verso, con independencia de que se pueda aplicar a otro tipo de metro u otra época.

Introducción

El Corpus de Sonetos del Siglo de Oro es un corpus representativo de toda la sonetística áurea castellana, que comprende los siglos XVI y XVII (los periodos llamados Renacimiento y Barroco, aproximadamente desde Garcilaso de la Vega hasta Sor Juana Inés de la Cruz). A diferencia de otros estudios tradicionales, este corpus pretende ser representativo de toda la producción sonetística de ambos siglos, no sólo de los poetas considerados canónicos.

La calidad literaria no es, en principio, criterio para incluir o rechazar a un autor. Los dos únicos criterios que se han tenido en cuenta para incluir a un autor en el corpus son: (i) que exista una edición digital de su obra y (ii) que se disponga de al menos 10 sonetos.

El corpus se presenta como recurso para el estudio de los rasgos generales del soneto en este periodo y su evolución. Este planteamiento se enmarca dentro de la línea de análisis de grandes corpus literarios propuesta y desarrollada por autores como [3] o [2]. En nuestro caso concreto, el objetivo es mostrar los aspectos rítmicos y métricos generales de todo el periodo. Para poder analizar estos aspectos con técnicas computacionales es necesario representar formalmente el patrón métrico de cada verso. En esta guía se expone tanto esa representación formal como los criterios para anotar el patrón métrico de cada uno.

Aspectos generales

2.1. Representación de la información métrica.

Se pretende representar la estructura métrica de un poema (soneto) mediante patrones métricos. La estructura métrica es la secuencia de sílabas átonas y tónicas que se producen dentro de un poema. El patrón métrico es la secuencia de sílabas tónicas y átonas que se producen dentro de un verso. Así, la estructura métrica de un soneto estará formada, en principio, por catorce patrones métricos.

Como norma general, sólo se representan de manera explícita aquellos elementos métricos que se puedan inferir del poema de manera más o menos objetiva. Por ello el patrón es métrico y no rítmico ni prosódico. Así, todos aquellos aspectos relacionados con el ritmo y la prosodia que puedan sufrir variaciones en la declamación (pausas potenciales, acentos secundarios, etc.) no serán representados. El patrón métrico consta, por tanto, únicamente de aquellos elementos métricos que deben ser respetados por la prosodia y que son por tanto obligatorios en el recitado del verso.

Según [6], la unidad métrica básica es el grupo métrico. Un grupo métrico es un conjunto de una o más sílabas alrededor de un acento de intensidad y separados del anterior o siguiente grupo métrico por una pausa potencial.

Los elementos métricos a representar en el patrón, por tanto, son dos: la sílaba y la pausa. Se consideran dos tipos de sílabas:

- Tónicas: aquellas que soportan un acento de intensidad primario. Se representan con el signo +.

- Átonas: aquellos que no soportan acento de intensidad primario. Se representan con el signo -.¹

En cuanto a la pausa que delimita un grupo métrico, ésta no es evidente: no hay rasgos objetivos que marquen su presencia. Las pausas que delimitan los grupos métricos son potenciales, pueden realizarse o no en la declamación según criterio del lector. Esto es especialmente evidente en el endecasílabo, donde pueden aparecer los grupos métricos seguidos sin solución de continuidad, como en 1,

(1) Si a vuestra voluntad yo soy de cera (---+++--)

o bien más claramente delimitados como en 2.

(2) Ir y quedarse y con quedar partirse (+-+--+--)

Esta misma secuencia de sílabas acentuadas y no acentuadas permite separar grupos métricos de manera diferente como en 3 (el símbolo “|” representa pausa potencial):

(3) Ir y quedarse | y con quedar partirse
 Ir y quedarse | y con quedar | partirse
 Ir | y quedarse | y con quedar | partirse

Dada esta variabilidad, en la anotación métrica del corpus no se va a tomar como **unidad métrica básica** el grupo métrico sino **el verso**. Un verso es una secuencia de grupos métricos (uno o más según el tamaño del verso) delimitado por una pausa [6]. A diferencia de grupo métrico, la pausa del verso es explícita en el sentido de que queda marcada de manera objetiva por el cambio de línea. Esta pausa queda establecida de manera consciente por el autor (incluso, o sobre todo, en verso libre). No sólo es unívoco (no da pie a ambigüedad), sino que es un elemento artístico que el poeta utiliza como marca de literariedad: tiene una clara función artística.

Por todo ello, un patrón métrico es una sucesión de sílabas tónicas y átonas, con un orden secuencial específico y delimitado por una pausa métrica (fin de verso). La cantidad de sílabas (límite del verso) y el orden específico de sílabas tónicas son la base del ritmo del verso: aquellos aspectos que la prosodia debe

¹No se tendrán en cuenta, por tanto, los acentos secundarios ni los acentos de apoyo. La determinación de estos, al no tener una representación objetiva en el poema, pueden sufrir variaciones en la prosodia o declamación. Por ello no se tendrán aquí en cuenta. Por otro lado, no se descarta utilizar sólo el número de posición de las sílabas tónicas.

respetar, y por tanto la base objetiva de la musicalidad del verso. Esta base métrica es lo que se quiere representar en el corpus.

A modo de ejemplo, presentamos los patrones métricos del poema “Cuando me paro a contemplar mi estado” de Garcilaso de la Vega.

```
<l met="---+---+--">Cuando me paro a contemplar mi estado,</l>
<l met="--+----+--">y a ver los pasos por do me ha traído,</l>
<l met="+---+---+--">hallo, según por do anduve perdido,</l>
<l met="--++-+-+--">que a mayor mal pudiera haber llegado;</l>

<l met="-----+--+--">mas cuando del camino esto olvidado,</l>
<l met="--+----+--">a tanto mal no sé por do he venido;</l>
<l met="+---+---+--">sé que me acabo, y más he yo sentido</l>
<l met="+---+---+--">ver acabar conmigo mi cuidado.</l>

<l met="+---+---+--">Yo acabaré, que me entregué sin arte</l>
<l met="---+---+--">a quien sabrá perderme y acabarme</l>
<l met="+---+---+--">si ella quisiere, y aun sabrá querello;</l>

<l met="-----+--+--">que, pues, mi voluntad puede matarme,</l>
<l met="--+---+---+--">la suya, que no es tanto de mi parte,</l>
<l met="--+---+---+--">pudiendo, qué hará sino hacello</l>
```

2.2. Proceso de anotación

Dado el tamaño del corpus, su anotación manual es inviable. Para poder tener el corpus completamente anotado en el menor tiempo y con la máxima calidad posibles, se va a seguir un proceso de anotación semi-automático.

El proceso de anotación consta de cuatro fases principales:

1. Anotación automática de todo el corpus mediante un sistema de escan-sión automático.²
2. Revisión y corrección manual de una parte seleccionada del corpus. Como mínimo un 10 % del corpus debe estar revisado a mano, pero cuanto más se revise, mejor. La parte revisada a mano queda como *Gold Standard*. Se determinarán y justificarán los casos erróneos y ambiguos.

²Me falta ponerle un nombre al sistema. ¿Sugerencias?

3. Definido el *Gold Standard*, se utilizará éste para mejorar el sistema con técnicas de aprendizaje automático. Con esta versión mejorada del sistema se anotará de nuevo la parte del corpus no validada a mano.
4. De esta parte re-anotada automáticamente, únicamente aquellos versos que presenten una ambigüedad potencial (según el análisis de errores de la fase 2) serán revisados a mano por un solo anotador (revisión transversal). Con esto finaliza la anotación del corpus.

Finalizado el proceso, se obtendrán dos recursos:

- El corpus completo anotado, donde los casos complejos y ambiguos han sido revisados y corregidos manualmente.
- Un sistema automático de escansión híbrido con el que anotar nuevos poemas.

La revisión y corrección manual consiste en comprobar que el patrón generado automáticamente se corresponde con el patrón métrico del verso. Si es correcto, se deja como está. Si es incorrecto, se modifica el patrón. En la medida de lo posible, tomad nota de aquellos casos más complejos o dudosos para poder mejorar la guía de anotación y el sistema.

Para que la anotación manual sea consistente (dos anotadores diferentes marquen lo mismo ante casos similares), se debería validar todo el corpus por al menos dos anotadores. Dado que esto no es viable, a efectos de revisión manual, el corpus se va a dividir en tres partes:

- Entrenamiento. Conjunto de sonetos para que el anotador se entrene y comprenda los problemas de la anotación. Esta revisión no formará parte de la anotación final del corpus. En principio, por importancia y complejidad, se utilizarán sonetos de Garcilaso de la Vega.
- Evaluación. Conjunto de sonetos que serán anotados en paralelo por varios anotadores. Estas anotaciones serán comparadas para detectar los casos de discrepancia y analizarlos. A partir de estos se determinarán las reglas de anotación específicas para casos complejos (sección 4).
- Resto de corpus, que quedará anotado por un solo anotador. Sólo los casos dudosos o complejos serán revisados por otro anotador para tomar una decisión consensuada.

En las siguientes secciones se especificarán las reglas de anotación generales y específicas.

Reglas generales

3.1. Reglas para determinar acentos de intensidad.

En un contexto textual, no todas las palabras cuentan con sílaba tónica y, por tanto, no se consideran palabras acentuadas a efectos métricos. En esta sección se indicarán qué tipo de palabras se consideran acentuadas a efectos métricos y en qué casos. Estas reglas generales de anotación métrica están basadas en [5] (págs. 22–26).

Palabras acentuadas

A efectos métricos, se consideran palabras acentuadas y tienen, por tanto, al menos un acento de intensidad (una sílaba tónica):

1. Los sustantivos (“el *gato*, la *mesa*”).

Ejemplo:

- (4) Marchitará la **rosa** el **viento** helado,
 $\langle met = - - - + - + - + - + - \rangle$
 (GarcilasodelaVega23.xml)

Escrito está en mi **alma** vuestro **gesto**,
 $\langle met = - + - + - + - - - + - \rangle$
 (GarcilasodelaVega5.xml)

Los sustantivos vienen marcados por el sistema de escansión automático mediante la etiqueta categorial (*PoS: Part of Speech*) NC (nombre común) o NP (nombre propio)¹. En principio esta categoría no presenta ambigüedad, salvo errores del analizador sintáctico.

2. El adjetivo (“el gato *negro*, la casa *gris*, la *triste* viuda”).

(5) Si para refrenar este deseo
loco, imposible, vano, temeroso
 $\langle met = + - - + - + - - - + - \rangle$
 (GarcilasodelaVega12.xml)

Los adjetivos vienen marcados con la etiqueta de adjetivo calificativo: AQ. En principio tampoco hay ningún problema de ambigüedad con esta categoría.

3. El pronombre tónico en función de sujeto o complemento con preposición (“*Tú* sabes poco, *él* y *nosotros* jugaremos, para *mí* y para *tí*”, etc.).

Ejemplos:

(6) **Yo** no nací sino para quereros;
 $\langle met = + + - + - - - - - + - \rangle$
 (GarcilasoDeLaVega_5.xml)

luego que sobre **mí** fueron mostrados
 $\langle met = + - - - - + + - - + - \rangle$
 (GarcilasoDeLaVega_20.xml)

Viene marcado por la etiqueta categorial PP (Pronombre personal).

4. Los adjetivos y pronombres indefinidos, bien aparezcan apocopados o no (“*Algún* hombre, viene *alguno*, *ningún* otro caso, *algún* fatigado”, etc.).

(7) Si **alguna** parte queda por ventura
 $\langle met = - + - + - + - - - + - \rangle$

¹El formalismo de representación son las etiquetas EAGLES. Estas etiquetas no aparecen en la anotación final del corpus, pero son utilizadas por el sistema de escansión automático para decidir si la sílaba tónica de una palabra debe formar parte del patrón métrico como tónica o como átona. Lo dejo aquí marcado porque es una fuente importante de error, dada la ambigüedad categorial de muchas palabras.

Del sueño, si hay **alguno**, aquella parte

$\langle met = - + - + - + - + - + - \rangle$

(GarcilasodelaVega17.xml)

Vienen marcados por las etiquetas DI (determinante indefinido) para el caso de los adjetivos y la etiqueta PI para el caso de los pronombres.

5. Los pronombres posesivos. (“la culpa es *mía*; este lápiz es *tuyo*; el gato no es *vuestro*, es *nuestro*,” etc.)

(8) Juntas estáis en la memoria **mía**,

$\langle met = + - - + - - - + - + - \rangle$

(GarcilasodelaVega10.xml)

Queda representado por la etiqueta PX.

Con los posesivos hay que llevar cuidado, ya que el determinante (o adjetivo) posesivo es átono. No confundir los pronombres con los determinantes:

(9) Escrito está en mi alma **vuestro** gesto,

$\langle met = - + - + - + - - - + - \rangle$

(GarcilasodelaVega5.xml)

6. Los demostrativos, tanto pronombres como adjetivos (“Quiero *este* libro, prefiero *aquél*”)

(10) De **aquella** vista pura y excelente

$\langle met = - + - + - + - - - + - \rangle$

(GarcilasodelaVega8.xml)

Están marcados con la etiqueta DD.

7. Los numerales cardinales y ordinales (“*Dos* casas, *mil* casas, viene el *primero*”)

Si el número es compuesto, sólo se acentúa el último numeral. Así,

cuarenta y seis = ---+

Quedan marcados con la etiqueta MC y MO (cardinal y ordinal respectivamente).

Ejemplo:

- (11) cuánto corta *una* espada en *un* rendido.

$\langle met = + - + + - + - + - + - \rangle$

(GarcilasoDeLaVega_02.xml)

8. El verbo en todas sus formas, tanto principal como auxiliar (“el gato *come*, se *casan* hoy, el pájaro *es* negro, Pepe *ha comido*”).

- (12) **hiere** y **enciende** el alma temerosa,

$\langle met = + - - + - + - - - + - \rangle$

(GarcilasodelaVega35.xml)

que con él **ha venido** a conformarse?

$\langle met = - - + + - + - - - + - \rangle$

(GarcilasodelaVega27.xml)

Marcados con la etiqueta categorial V.

9. Los adverbios (“come *poco*, juega *mal*”)

Marcados con la etiqueta RG. Ejemplo:

- (13) virtud, **siempre** triunfante, **siempre** hermosa

$\langle met = - + + - - + - + - + - \rangle$

(Quevedo159.xml)

Los adverbios acabado en “-mente” tienen doble acentuación. Por ejemplo:

- (14) que solo le matara **fácilmente**.

$\langle met = - + - - - + - + - + - \rangle$

(LopeDeVega800.xml)

10. Las formas interrogativas “qué, cuál, quién, dónde, cuándo, cuánto, cómo” (“*Qué* quieres, *cómo* va la vida”)

No es necesario etiqueta para localizarlas, pues es una lista cerrada y pequeña de palabras que se detectan directamente. Están representadas con la etiqueta PT.

11. Las siguientes conjunciones se consideran tónicas:

- las disyuntivas **ora**, **ya**, **bien**;

- la consecutiva **así**;
- la temporal **apenas**;
- las compuestas adversativas **no obstante**, **con todo**, **fuera de**,
- las compuestas consecutivas **en efecto**, **por tanto**, **por consiguiente**, **así que**,
- las compuestas temporales **aún no**, **no bien**, **luego que**, **después que**, **en tanto que** (pero es átono en cuanto, en cuanto que),
- las compuestas condicionales **a no ser que**, **dado que**, **con tal que**,
- las compuestas concesivas **por más que**, **a pesar de que**, **mal que**, etc.

Ejemplo:

- (15) Y **así**, se quedan tristes en la puerta
 $\langle met = - + - + - + - - - + - \rangle$
 (GarcilasoDeLaVega_22.xml)

Palabras inacentuadas

Siguiendo a Quilis [5] (págs 23 y 24), se consideran, a efectos métricos, palabras sin acento de intensidad en la estructura métrica:

1. El artículo determinado (*el alma*, *la casa*, etc.)

Ejemplo:

- (16) dejad un rato **la** labor, alzando
 $\langle met = - + + + - - - + - + - \rangle$
 (GarcilasodelaVega_11.xml)

2. La preposición

Ejemplo:

- (17) juzgando **por** tu cielo, **en** cuya esfera
 $\langle met = - + - - - + - - - + - \rangle$
 (Quevedo_153.xml)

3. Las conjunciones:

- copulativas **y**, **e**, **ni**;
- disyuntivas **o**, **u**;
- el **que** en todas sus formas átonas (excepto pronombre interrogativo o exclamativo): copulativo, disyuntivo, determinativo, etc.;
- adversativas **pero**, **sino**, **mas**, **aunque**;
- causales **pues**, **porque**, **como**, **pues que**, **puesto que**, **supuesto que**;
- consecutivas **pues**, **luego**, **conque**;
- condicionales **si**, **cuando**;
- concesivas **aunque**, **aun**, **cuando**.

Ejemplo:

- (18) en él puse la vista; **mas** detiene
 $\langle met = - + + - - + - - - + - \rangle$
 (GarcilasoDeLaVega_22.xml)

4. Los términos de tratamiento (“*don* José, *santo* Tomás”, etc.)

Ejemplo:

- (19) a **don** Carlos no aclame y no le llore
 $\langle met = - - + - + + - + - + - \rangle$
 (Quevedo_23.xml)

5. El primer elemento de los compuestos (“*María* José, *dos* mil”), incluso “palabras compuestas en las que aún se sienten sus componentes” (pág. 24) como “*tragaluz*”.6. Los pronombres átonos: funcionan como complemento, incluyendo el **se** reflexivo.

Ejemplos:

- (20) sola la noche **los** ostenta oscura
 $\langle met = + - - + - - - + - + - \rangle$

(Quevedo_81.xml)

por descansar en su dolor, **te** llama
 $\langle met = - - - + - - - + - + - \rangle$
 (Quevedo_160.xml)

7. Determinantes posesivos, apocopados o no (“*mi* madre, *nuestra* casa”, etc.).

Ejemplo:

(21) Para agotar **sus** luces la hermosura
 $\langle met = - - - + - + - - - + - \rangle$
 (Quevedo_81.xml)

8. Las siguientes formas, siempre y cuando **no** funcionen como interrogativas ni exclamativas: **que**, **cual**, **quien**, **donde**, **cuando**, **cuanto**, **como**.

Ejemplos:

(22) **cuando** el vencido os dedicara altares
 $\langle met = - - - + - - - + - + - \rangle$
 (Quevedo_90.xml)

de **cuanto** bien, cuitado, yo tenía;
 $\langle met = - - - + - + - + - + - \rangle$
 (GarcilasoDeLaVega_03.xml)

que con todas las Indias enamora
 $\langle met = - - + - - + - - - + - \rangle$
 (Quevedo_90.xml)

Quien no teme alcanzar lo que desea
 $\langle met = - + + - - + - - - + - \rangle$
 (Quevedo_100.xml)

9. En vocativos y “expresiones exclamativas cortas de cariño o reproche son inacentuados los elementos que acompañan al núcleo” (pág. 25). Así, la métrica de las siguientes expresiones es diferente según sea vocativo/exclamativo o no:

- “no puedo, *buen* hombre” (+ + - - + -) vs.
 “no es un *buen* hombre” (+ - + + -)
- “el hermano Juan” (- - - - +) vs.
 “el hermano de Juan” (- - + - - +)

Para este último caso véase, sin embargo, el apartado 4.4.

Palabras ambiguas

Consideramos palabra ambiguas aquéllas que, teniendo la misma forma, pueden ser tónicas o átonas según la función que asuman en cada contexto.

1. **Luego**: tónica con función temporal, átona con función consecutiva.
2. **Aun**: tónica como adverbial temporal (cuando equivale a *todavía*), átona resto de casos.

Ejemplos:

(23) si fue cuerpo o deidad, **aún** hoy lo dudo.

$\langle met = - + + - - + + + - + - \rangle$

(Quevedo_51.xml)

tan solo, que **aun** de vos me guardo en esto.

$\langle met = - + - - - + - + - + - \rangle$

(GarcilasoDeLaVega_05.xml)

3. **Mientras**: tónica como adverbio, átona como conjunción.
4. **Medio**: tónica como adjetivo, átona en formas compuestas (**medio** dormido, medio día)
5. **Más**: tónica como adverbio, átona como nexos de relación (“cuatro **más** dos”).
6. **Menos**: tónica como adverbio, átona como nexos de relación (“cuatro **menos** dos”).

En caso de duda, documentar la fonética de la palabra para determinar su carácter tónico o átono. Si aún así hay dudas, seleccionar siempre la opción **ÁTONA**. De esta manera se entiende como acento secundario y, por tanto, no se debe representar.

3.2. Reglas para determinar el tamaño del verso.

En ocasiones a un mismo verso se le puede asignar un patrón métrico u otro según las sinalefas o sinéresis que se realice. Esta ambigüedad se produce cuando un verso tiene más posibles sinalefas o sinéresis de las necesarias para establecer el cómputo de once sílabas.

Por ejemplo, si se cuentan las sílabas del siguiente ejemplo (24) sin realizar ninguna sinalefa, el verso constaría de doce sílabas.

- (24) cuando el padre Hebrero nos enseña
cuan-do-el-pa-dre-He-bre-ro-nos-en-se-ña

Para que tenga once sílabas es necesario realizar una única sinalefa. Sin embargo, hay dos posibles sinalefas: **cuan-do-el** y **padre-Hebrero**. Según se realice una u otra, el patrón métrico resultante es diferente: en el primer caso habría acento en tercera, y en el segundo habría acento en cuarta (ver 25).

- (25) cuan-doel-pa-dre-He-bre-ro-nos-en-se-ña
< met = - - + - - + - - - + - >

cuan-do-el-pa-dreHe-bre-ro-nos-en-se-ña
< met = - - - + - + - - - + - >

Otro ejemplo con diferentes patrones asignables al mismo verso según la realización o no de diversas sinalefas: 26.

- (26) hiere y enciende el alma temerosa,
hie-rey-en-cien-deel-al-ma-te-me-ro-sa,
< met = + - - + - + - - - + - >

hie-reyen-cien-de-el-al-ma-te-me-ro-sa,
< met = + - + - - + - - - + - >
(GarcilasodelaVega35.xml)

Ante estos casos de ambigüedad, se aplican dos reglas:

1. Seleccionar siempre la sinalefa considerada menos forzada fonológicamente.

Así, se considera menos forzada la unión de:

- dos sílabas átonas, frente a dos sílabas tónicas o tónica más átona; Por ejemplo, el siguiente verso (27) es mejor el primer patrón, donde se hace sinalefa entre... a la segunda, donde se hace sinalefa en ..., donde habría que juntar una sílaba átona con única tónica.

(27) y en torcidas raíces se volvían.

y_ **en**-tor-ci-das-ra-íces-se-vol-ví-an.

< met = - - + - - + - - - + - >

y-en-tor-ci-das-**ra**_íces-se-vol-ví-an.

< met = - - - + - + - - - + - >

(GarcilasoDeLaVega_13.xml)

- sílaba tónica y átona (o al revés) frente a dos tónicas;
 - dos sílabas frente a la unión de tres o más sílabas en una única sílaba;
 - dos sílabas entre las que no haya posible pausa potencial, frente a sílabas entre las que sí la pueda haber (los signos de puntuación pueden ayudar a establecer esas posibles pausas potenciales. Recordad, en todo caso, el ejemplo 3).
 - ... [Me falta completar esto. *Don't Panic!*]
2. Seleccionar el patrón resultante más común, el más estable o menos forzado.² Así, se preferirá un patrón con acento en sexta a uno con acento en séptima, pues es más común aquél que éste. De la misma manera, se preferirá un patrón sin acento anti-rítmico (dos sílabas acentuadas seguidas) a un patrón con acento anti-rítmico.

3.3. Sobre la corrección del texto.

No es objetivo de este proyecto hacer la edición crítica de los sonetos del Siglo de Oro. En todo momento asumimos un texto ya fijado y editado. La edición base que se utiliza es la de la Biblioteca Virtual Miguel de Cervantes (salvo que para determinado autor se haya localizado una edición mejor y que ésta se pueda distribuir libremente). En el encabezado de cada fichero TEI-XML se indica de manera explícita la edición del texto utilizada (etiqueta < sourceDesc >: *source description*). En principio el texto no se modifica, salvo errores de transcripción o digitalización evidentes.

²Esta regla no es muy objetiva, pues depende del anotador decidir qué es y qué nos es estable. Pensaré cómo formalizarla mejor.

Si embargo, cotejando el texto con *editio princeps*³ y con ediciones moderna se han detectado que determinados versos no tienen una lectura única, y las diferentes lecturas suponen patrones métricos diferentes.

Sin ánimo de exhaustividad, si se detectara esta situación, se anotarán las diferentes lecturas con el patrón métrico correspondiente a cada una. Cada variante deberá estar apoyada, o bien por una edición original o bien por una edición crítica moderna. No se va a crear el *stemma codicum*, pues no es una edición crítica. El objetivo es simplemente marcar que en ese verso hay varias lecturas y, con ellos, patrones métricos diferentes. En todo caso, estas anotaciones pueden servir para hacer la edición crítica digital en un futuro.

Así ocurre, por ejemplo, con el verso 5 del poema 8 de Garcilaso de la Vega “De aquella vista pura y excelente”. En la edición base de la Biblioteca Virtual Miguel de Cervantes se lee:

«Encuéntanse al camino fácilmente,»
(- + - - - + - + - + -)

Sin embargo, se ha detectado que en diversas ediciones modernas⁴, ese verso es diferente. En éstas se lee:

«Éntranse en el camino fácilmente,»
(+ - - - - + - + - + -)

Ambas lecturas suponen cambio en el patrón métrico final, por tanto es un problema textual pertinente para el corpus. Al cotejar la *editio princeps* (1543)⁵ aparece un verso erróneo por tener 12 sílabas:

«Encuéntrese en el camino fácilmente,»
(- + - - - - + - + - + -)

Este error evidente fue enmendado tanto en la edición de Herrera como en la edición de El Brocense. El primero propuso «Encuéntanse al camino fácilmente,», que es la variante seguida por Navarro Tomás y la que hay en la

³Consultadas a través de la Biblioteca Digital Hispánica <http://www.bne.es/es/Catalogos/BibliotecaDigitalHispanica/Inicio/>

⁴En este caso se han cotejado las ediciones de Rivers (Castalia), Alcina (Austral) y Prieto de Paula (Castalia). También se ha cotejado al edición de Navarro Tomás (Espasa), que tiene la misma lectura que la edición digital de la Biblioteca Virtual Miguel de Cervantes.

⁵<http://bdh.bne.es/bnearch/CompleteSearch.do?field=autor&fechaFhasta=1543&text=Garcilaso+de+la+vega&fechaFdesde=1496&showYearItems=&exact=on&textH=&advanced=false&completeText=&pageSize=1&pageSizeAbrv=10&pageNumber=2>

edición de la Biblioteca Virtual Miguel de Cervantes. Las ediciones modernas, sin embargo, prefieren la lectura que propone la edición de El Brocense: «Éntranse en el camino fácilmente,».

Dado que ambas lecturas son propuestas posteriores y no podemos seleccionar ninguna de las dos como “correcta”, se opta por incorporar ambas al corpus.

Para marcar las variantes, hay que introducir nuevas etiquetas tanto en el encabezado TEI como en el cuerpo del texto.

- Encabezado TEI (< *teiHeader* >):

En el encabezado hay que introducir las fuentes donde se han detectado las diferentes lecturas. Para ello se utiliza la etiqueta < *listWit* >, que debe aparecer dentro de la descripción de la fuente (< *sourceDesc* >). Dentro de ésta, cada fuente queda marcada con la etiqueta < *witness* >, donde se debe indicar la información bibliográfica básica para describir la fuente. Junto a esto, se debe dar un identificador que permita hacer referencia a ella. Este identificador se marca con el atributo @*xml* : *id*.

Para el verso de Garcilaso, vamos a considerar sólo las ediciones de Herrera (id="GDLV_1") y de El Brocense (id="GDLV_2"). Las etiquetas del encabezado quedarían así:

```
<sourceDesc>
...
<listWit>
<witness xml:id="GDLV_1"><title>Obras de Garcilasso de la
Vega con anotaciones de Fernando de Herrera. Alonso de la
Barrera</title>, <date>1580</date>.
</witness>

<witness xml:id="GDLV_2"><title>Obras de Garcilasso de la
Vega con las anotaciones por el Mro. Francisco Sánchez
Brocense. Juan de la Cuesta</title>, <date>1612</date>
</witness>
</listWit>
</sourceDesc>
```

- Cuerpo del poema (< *body* >):

Una vez que en el encabezado se han indicado las fuentes de las variantes, en el cuerpo del texto se introducen esas variantes, siempre introduciendo

el identificador (id) que permita saber a qué fuente pertenece cada lectura o variante.

Para ello, en el propio texto, se utiliza la etiqueta `< app >` (*apparatus entry*). Dentro de ésta se introducen todas las variantes. De éstas, hay que seleccionar una como preferida, cuyo patrón métrico quedará marcado en la etiqueta `@met`. Esta lectura preferida se marca con la etiqueta `< lem >` (*lemma*). El resto de versiones o lecturas se marcan con la etiqueta `< rdg >` (*reading*). Mediante el atributo `@wit` se indica en ambos casos el identificado de la fuente tal y como se haya indicado en el encabezado.

Así, el verso anterior quedaría anotado de la siguiente manera:

```
<l n="5" met="-+---+-+--">
<app>
<lem wit="GDLV_1">Encuéntanse al camino fácilmente,</lem>
<rdg wit="GDLV_2">Éntranse en el camino fácilmente,</rdg>
</app>
</l>
```

En este caso, de ambas lecturas se prefiere, a efectos métricos, la primera (Herrera). Por ello el patrón de ésta es el que queda marcado en el atributo `@met`.

Ante posibles errores que se consideren graves:

1. cotejar (en la medida de lo posible) con la *editio princeps* y seleccionar su lectura. Es preferible que en el corpus haya errores de la *princeps* que errores producidos por modernizaciones posteriores. Tomad nota del cambio realizado y la fuente.

Las ediciones *princeps* se encuentran en la Biblioteca Digital Hispánica:

<http://www.bne.es/en/Catalogos/BibliotecaDigitalHispanica/Inicio/>

2. cotejar (en la medida de lo posible) ediciones modernas por si en ese verso hubiera un problema de edición crítica. Decidir la lectura crítica que se considere mejor, y tomar nota del cambio.

En ambos casos, hay que dejar constancia del cambio y de las fuentes utilizadas (que deben ser citadas convenientemente).

Y en todo caso, estas modificaciones deben ser una excepción, sólo para errores muy evidentes. En la medida de lo posible, se debe intentar modificar el texto lo menos posible.

Casos problemáticos

En esta sección se expondrán casos problemáticos que requieren criterios específicos de anotación.

4.1. “Ya que”

Según Quilis, la conjunción “ya que” es conjunción compuesta temporal o concesiva [5]. Si bien en su origen la conjunción “ya que” tenía un sentido temporal [1], en el Siglo de Oro ese sentido temporal ya ha cambiado por el uso causal que tiene en el actualidad. Por ello, como norma general, toda conjunción “ya que” será considerada causal y por tanto átona (—). Únicamente si se localizara algún caso de uso claramente temporal, se anotaría como tónica (+—), pero debe ser un caso muy claro.

Caso de conjunción “ya que” concesiva puede ser el siguiente ejemplo 28:

- (28) El que puede caer, si él se derriba,
ya que no se conserva, se previene
 $< met = + - + - - + - - - + - >$
 contra el semblante de la suerte esquiva
 (Quevedo340.xml)

La conjunción “ya que” del siguiente ejemplo 29 podría considerarse temporal. Sin embargo, el valor causal está presente.

- (29) Ya que pasó mi verde primavera
 $< met = - - - + - + - - - + - >$
 (Quevedo158.xml)

Otro ejemplo, esta vez de una obra en prosa, lo encontramos en 30:

- (30) Y *ya que* la argentada Aurora enseñaba sus plateados cabellos por el sacro Oriente, la contentísima Berta le dijo al Emperador y Monarca del mundo
(Antonio de Eslava, *Noches de Invierno*, 1609)¹

Si bien en ambos casos el contexto da una interpretación temporal al “ya que”, éste sigue teniendo valor causal, por lo que debería anotarse como átono. Compárese con la siguiente colocación “ya que” claramente temporal (31)

- (31) ya es llegada la hora deseada; tiempo es *ya que* nos veamos, amado mío y Señor mío
(María de San Francisco, c. 1614-1622)²

4.2. “Cada”

“Cada” es determinante indefinido. Según Navarro Tomás [4], los indefinidos son tónicos.

Navarro Tomás, sin embargo, tras afirmar diferencias dialectales, concluye que es más predominante el uso átono de «cada» que su uso distributivo [4]:

“El distributivo «cada» presenta diferencias regionales, predominando la forma inacentuada: «**cada** día está más alto»”.

Por ahora lo dejamos tónico por similitud con su clase (indefinidos). De todas formas, queda como tema pendiente para estudiar con detalle más adelante.

Ejemplo:

- (32) y yéndome alejando **cada** día,
< met = - + - - - + - + - + - >
(GarcilasoDeLaVega_03.xml)

¹Tomado del corpus CORDE (<http://corpus.rae.es/cordenet.html>), 20 de septiembre de 2015.

²Tomado del corpus CORDE (<http://corpus.rae.es/cordenet.html>), 20 de septiembre de 2015.

4.3. “Tan”

Siguiendo a Navarro Tomás [4] (pág. 191), si bien «tanto» y «tanta» y sus plurales son formas tónicas, «tan», como indefinido que es, es forma átona siempre («*Tan* alto como las nubes»).

Ejemplo:

- (33) **Tan** cansada de haberse levantado
 $< met = - - + - - + - - - + - >$
 (GarcilasoDeLaVega_04.xml)

do sé que he de morir **tan** apretado,
 $< met = - + + - - + - - - + - >$
 (GarcilasoDeLaVega_02.xml)

4.4. Vocativos tónicos

Siguiendo a Navarro Tomás [4] y a Quilis [5] se ha establecido que las expresiones exclamativas cortas de cariño o reproche son inacentuados los elementos que acompañan al núcleo. Navarro Tomás deja establecido que:

En toda locución breve de carácter vocativo y en expresiones cortas de cariño o reproche, pierde de ordinario su acento cualquier nombre, sustantivo o adjetivo, que ocupe el principio de la invocación: «¡*buen* hombre!», «¡*mala* lengua!», «¡*gran* pícaro!», «¡*Dios* mío!», «¡*cara* de rosa!». Fuera de este caso, se acentúan siempre estas mismas palabras.

Para especificar qué es “corta”, a todos los efectos se consideran cortas las palabras monosílabas. Las bisílabas, salvo que su carácter átono sea evidente, se considerarán tónicas. Así, queda inacentuado los siguientes vocativos (ejemplo 34)

- (34) ¡Oh **crudo** nieto, que das vida al padre
 $< met = - - - + - - + + - + - >$
 (GarcilasoDeLaVega_31.xml)

(¡oh, **bella** Clori!) tus pisadas santas
 $< met = - - - + - - - + - + - >$

(Gongora_1.xml)

doña Pérez, saber ciertas verdades:

< *met* = - - + - - + + - - + - >

(Quevedo_170.xml)

pero son claramente tónicos el siguiente vocativo bisílabo (ejemplo ??):

(35) ¡Oh **grande** error! Pues cuando de ejemplares

< *met* = - + - + - - - - + - >

(Quevedo_20.xml)

Queda acentuado el mismo caso en los ejemplos 36 y 37, ya que el adjetivo que acompaña al núcleo del vocativo es tetrasílabo y trisílabo respectivamente.

(36) **Clarísimo** Marqués, en quien derrama

< *met* = - + - - - + - - - + - >

(GarcilasoDeLaVega_21.xml)

(37) ¡Oh **celoso** temor!, ¿a quién pareces?

< *met* = - - + - - + - + - + - >

(GarcilasoDeLaVega_31.xml)

4.5. Sinalefa y “h” aspirada

Se debe evitar, en la medida de lo posible, hacer sinalefas donde un “h” podría ser aspirada. Véase ejemplo 38.

(38) de vuestra hermosura el duro encuentro

< *met* = - + - - - + - + - + - >

(GarcilasoDeLaVega_22.xml)

En este ejemplo hay dos posibles sinalefas: “vues-**tra**_her-mo-su-ra” y “her-mo-su-**ra**_el”. Para establecer once sílabas, sólo una debe ser realizada. Entre las dos, se prefiere la segunda (“ra_el”) ya que en la primera la “h” es aspirada.

Lo mismo ocurre en los siguientes ejemplos (39):

- (39) No sé ya *qué hacer***me en** mal tamaño.

$\langle met = + + + + - + - + - + - \rangle$

(GarcilasoDeLaVega_09.xml)

los blancos pies en tierra **se hincaban**

$\langle met = - + - + - + - - - + - \rangle$

(GarcilasoDeLaVega_13.xml)

que considere el daño que haciendo

$\langle met = - - - + - + - - - + - \rangle$

(GarcilasoDeLaVega_14.xml)

Otro ejemplo de sinalefa que no se realiza por presencia de “h” aspirada lo tenemos en el siguiente verso (40)

- (40) **me hacen** descuidar de mi remedio.

$\langle met = - + - - - + - - - + - \rangle$

(GarcilasoDeLaVega_06.xml)

4.6. Ambigüedad en la sinalefa.

Hay versos donde la realización de sinalefas es ambigua, pues según se realicen unas u otras, el patrón resultante será diferente. En estos casos, como se comentó anteriormente, buscar siempre el patrón más común.

En el ejemplo 41 se prefiere ese patrón y no la alternativa con acentos en 4, 7 y 10 ($- - - + - - + - - + -$) porque es más común un endecasílabo con acento en sexta que en séptima. Además de que es más estable el metro seleccionado (3,6,10), se evita hacer sinéresis en “ia” de “italiano”, que algunas ediciones lo escriben con diéresis (“italiano”).

- (41) y el antiguo valor italiano,

$\langle met = - - + - - + - - - + - \rangle$

(GarcilasoDeLaVega_35.xml)

El ejemplo 42 se prefiere este patrón al alternativo en 4, 5, 7 y 10 ($- - - + - + - - + -$) por ser más común.

- (42) como de un solo hijo deseado;

$\langle met = - - + + - + - - - + - \rangle$

(GarcilasoDeLaVega_31.xml)

Finalmente, en el ejemplo 43 el patrón seleccionado es más estable y más común que la alternativa en 1, 4, 6, 8 y 10 (+ - - + - + - + - + -)

(43) su amo, y no le hallaba, y esto siente;
 $< met = - + - + - + - + - + - >$
 (GarcilasoDeLaVega_36.xml)

En todo caso, no forzar el análisis para que el patrón se ajuste a un patrón común. Aplicar esta regla en casos de ambigüedad evidente.

Bibliografía

- [1] Emilio Alarcos Llorach. *Gramática de la Lengua Española*. Espasa, Madrid, 1994-2000. 23
- [2] Matthew L. Jockers. *Macroanalysis. Digital Media and Literary History*. University of Illinois Press, Illinois, 2013. 3
- [3] Franco Moretti. *La literatura vista desde lejos*. Marbot ediciones, Barcelona, 2007. 3
- [4] Tomás Navarro Tomás. *Manual de pronunciación española*. Consejo Superior de Investigaciones Científicas, Madrid, 1974. 24, 25
- [5] Antonio Quilis. *Métrica española*. Ariel, Barcelona, 1984. 9, 13, 23, 25
- [6] Elena Varela Merino, Pablo Moíno Sánchez, and Pablo Jauralde Pou. *Manual de métrica española*. Castalia, Madrid, 2005. 5, 6