

Aoidos: una herramienta para análisis automático de versos

Adiel MITTMANN

Universidade Federal de Santa Catarina

adiel@mittmann.net.br

Alckmar LUIZ DOS SANTOS

Universidade Federal de Santa Catarina

alckmar@gmail.com

Samanta ROSA MAIA

Universidade Federal de Santa Catarina

samantamaia9@gmail.com

Sumario

1. Introducción

2. El sistema *Aoidos*

3. Preparación y escansión de la obra poética de Gregório de Matos

4. *Aoidos* frente a Chociay

5. Algunas conclusiones

6. Notas

7. Bibliografía

1. Introducción

Escandir versos es una actividad que nosotros, humanos, después de algún entrenamiento y la adquisición de alguna experiencia, logramos realizar de forma bastante eficiente. En la mayoría de los casos, una mirada al poema ya nos da una buena idea de su patrón estrófico y métrico; la lectura del primer verso confirma o ajusta las sospechas que teníamos acerca del metro; al final de la primera estrofa ya identificamos los patrones rítmicos preferidos del poeta y tenemos expectativas de lo que viene por delante a partir de allí. Cuando llegamos al final del poema, sin embargo, ciertas preguntas son difíciles de responder: ¿qué variaciones rítmicas el poeta utilizó y con qué frecuencia?, por ejemplo, ¿cuántas sinalefas y sinérisis empleó el poeta? ¿Cómo comparar el poema o el poeta con otros con base en datos objetivos? Si estamos leyendo un solo soneto, podemos hacer algunas cuentas y responder a las dos primeras preguntas; la tercera se queda sin respuesta. Y si estamos leyendo *Os Lusíadas*, con sus 8816 versos —¿cómo vamos a hacer las cuentas? ¿Y si queremos analizar cuantitativamente los más de 30 000 versos de Gregório de Matos?

Aoidos (Mittmann; Wangenheim; Luiz Dos Santos, 2016; Mittmann, 2016) es una herramienta computacional capaz de escandir versos métricos en portugués. Su objetivo no es generar escansiones mejores que aquellas hechas manualmente por académicos; de hecho, es capaz de generar buenas escansiones —además de recoger información cuantitativa mientras lo hace— mucho más rápidamente que los lectores, aunque sean especialistas. Si alguien pudiera escandir cada uno de los 4184 versos de las *Cartas Chilenas* en apenas un segundo, aun así estaría trabajando 200 veces más lento que *Aoidos*. La escansión de un gran número de versos y la anotación de datos cuantitativos es agotadora para un humano, pero el ordenador no se cansa. Además, conforme los ordenadores se vuelven más rápidos, también crece la velocidad de tales programas. Así, además de posibles objetivos pedagógicos y educativos, la utilidad de una herramienta como esta es escandir versos y generar estadísticas cuando nosotros, los humanos, no tendríamos tiempo suficiente para hacerlo. Esa enorme ventaja es evidente sobre todo en lo que se refiere a análisis que toman también en cuenta elementos cuantitativos de los versos, de los poemas y de las obras de sus autores: ¿cuántos años nos llevaría, por ejemplo, hacer una simple comparación del uso de las sinalefas complejas en la obra del poeta brasileño Augusto dos Anjos, con respecto a algunos de sus contemporáneos? Ese tiempo se reduce drásticamente, sobre todo cuando ya tenemos las obras literarias digitalizadas y disponibles en formato texto. Por otro lado, este tipo de acercamiento es de suma importancia para excluir arbitrariedades o impresiones personales en la lectura. De hecho, no basta con decir que tal obra o tal autor utiliza este o aquel elemento de construcción del verso de manera más o menos frecuente; es necesario dar una base concreta a ese género de afirmación, lo que se hace fácil e inmediato con la utilización de herramientas como *Aoidos*.

Ahora bien, aplicaciones similares se propusieron para versos escritos en otros idiomas. En la década de 1990, Robey (1993) presentó un sistema que ayuda a hacer la escansión de los versos de la *Divina Comedia*. Para el español, Gervás (2000) creó

un sistema de escaneo de versos que se mostró capaz de acertar un 89% los tipos de versos de un corpus de 64 sonetos. Para el mismo idioma, Navarro-Colorado, Lafoz y Sánchez (2016) propusieron un sistema para escandir sonetos, acertando el 96% de los versos. Para el francés, hay el Métromètre, de Beaudouin (2000), que fue utilizado para escandir más de 80 000 versos del teatro francés; y hay también Anamètre, desarrollado por Delente y Renault (2015), que analiza versos en metros variados. Mamede, Trancoso, Araújo y Viana (2004) presentaron un sistema de análisis de versos para la lengua portuguesa, pero con un enfoque diferente de aquel de *Aoidos* y con un desempeño inferior: mientras éste acierta un 99% de los versos, la tasa de acierto del sistema de aquellos autores es un 82%.

Uno de los pocos investigadores que se dispusieron a escandir y hacer recuentos manuales de grandes cantidades de versos es Rogério Chociay, autor de la conocida *Teoria do Verso* (Chociay, 1974). Para caracterizar, por ejemplo, *O Uruguai* de Basílio da Gama, Chociay contó el número de ocurrencias de todos los esquemas rítmicos encontrados en los 1377 decasílabos de la obra, agrupándolos según los cantos en que aparecen y clasificándolos en heroicos y sáficos (Chociay, 1994). Para comparar, en el mismo artículo, ese poema de Basílio da Gama con obras de otros seis poetas, de Camões a Jorge de Lima, el autor de *Teoria do verso* hizo el mismo recuento de esquemas rítmicos para los 100 primeros versos de una obra escogida de cada poeta, motivado por la posibilidad de comparar las obras de los poetas de forma objetiva. Así, después de citar algunos críticos de *O Uruguai*, indaga:

No explican, sin embargo, los mencionados críticos, en qué consiste, objetivamente, esa "fluidez", esa "diferencia" de los versos de *O Uruguai*: ¿sería un modelo sui generis de decasílabo? o apenas un modo diferente de construir el verso?¹

En este artículo, vamos a analizar otro estudio realizado por Chociay, más ambicioso en la cantidad de versos: se trata de los datos que publicó en *Os Metros do Boca* (Chociay, 1993). En ese libro, el autor

explora la técnica versificatoria de Gregório de Matos y presenta una serie de cantidades acerca de los más de 30 000 versos compuestos por el poeta. Si *Aoidos* puede ser usado para producir análisis semejantes a las hechas manualmente por Chociay sobre los versos de Gregório, tendremos buenos indicios de que el mismo puede ser hecho para otros poetas. Es justamente lo que pretende este artículo: presentar paralelamente los resultados de Chociay y de *Aoidos* y compararlos, mostrando que, aunque no sean idénticos, son equivalentes, es decir, permiten prácticamente las mismas inferencias e hipótesis de lectura. La diferencia fundamental no reside en los resultados producidos por el humano y el ordenador, sino en la velocidad: a *Aoidos* le llevó 127 segundos para analizar todos los versos, mientras que Chociay tardó "muchos meses" en hacerlo (Chociay, 1993, 13).

El resto del artículo, por lo tanto, explica cómo la aplicación genera sus propuestas de escaneo a partir del texto escrito; a continuación discurre acerca de la preparación del corpus con los versos de Gregório de Matos; presenta, entonces, y compara los números de Chociay y de *Aoidos*, destacando semejanzas y diferencias.

2. El sistema *Aoidos*

En resumen, lo que tenemos aquí es un programa de computadora capaz de escandir versos métricos en portugués. Mientras que un lector humano utiliza sus conocimientos lingüísticos y culturales para escandir versos, *Aoidos* parte del verso escrito y, a través de una serie de reglas y procedimientos, todos codificados de forma explícita y mecánica, produce una propuesta de escansión. Tomemos, por ejemplo, el primer cuarteto de un soneto de Gregorio de Matos:

Triste Bahia! Oh quão dessemelhante
Estás, e estou do nosso antigo estado!
Pobre te vejo a ti, tu a mi empenhado,

Rica te vejo eu já, tu a mi abundante.

Para este cuarteto, *Aoidos* produce la siguiente escansión:

1	2	3	4	5	6	7	8	9	10	11
Tris-	te	Bah-	ial	Oh	quãõ	de-	sse-	me-	lhan-	te
Es-	tá-	s, e es-	tou	do	no-	sso an-	ti-	go es-	ta-	do!
Po-	bre	te	ve-	jo a	ti,	tu a	mi em-	pe-	nha-	do,
Ri-	ca	te	ve-	jo eu	jã,	tu a	mi a-	bun-	dan-	te.

No todas las informaciones recogidas por el sistema son visibles en esa escansión. El sistema sabe, entre otros, que el "a" final de "Bahía", en el primer verso, fue elidido, y que hay una sinalefa entre "tú" y "a" en el tercer verso. A continuación, intentamos explicar, de forma bastante general, cómo *Aoidos* llega a una propuesta de escansión como esa, a partir del verso escrito.

De hecho, la función primordial de *Aoidos* es la escansión de versos explícita, concreta y directamente disponible, es decir, de versos que estén individualizados e identificados como tal para el sistema. No se pretende hacer la extracción de documentos de cualquier tipo (que pueden presentarse en varios formatos posibles: DOC, de Microsoft Word, PDF, Portable Document Format, típicamente visualizados en Acrobat Reader, HTML, Hypertext Markup Language, visualizados en navegadores Web, etc.). Disponemos de la habilidad de abrir un libro y, de forma rápida y eficaz, localizar visualmente dónde están los versos de los poemas y separar los que constituyen el poema en sí de aquellos que están en el epígrafe debajo del título. Para un

programa como *Aoidos*, sin embargo, estas tareas no son triviales, ya que no está dotado del sentido común necesario para interpretar todas las variaciones de formato (espaciamientos, tamaño de letras, uso de itálicas, etc.) que pueden aparecer en libros. Por ello, se debe informar de cuáles son y dónde están los versos, lo que explica la obligación de que los poemas sean estandarizados en su presentación: *Aoidos* solo es capaz de analizar poemas en el formato TEI (Text Encoding Initiative). En este tipo de archivo, los versos quedan explícitamente delimitados y agrupados, lo que facilita la interpretación de su contenido por herramientas computacionales. Un archivo en formato TEI se puede crear y modificar en editores simples como el bloc de notas de Windows. El cuarteto anterior, por ejemplo, queda así en el formato TEI:

```
<lg n="1">  
  <l n="1">Triste Bahia! Oh quão dessemelhante</l>  
  <l n="2">Estás, e estou do nosso antigo estado!</l>  
  <l n="3">Pobre te vejo a ti, tu a mi empenhado,</l>  
  <l n="4">Rica te vejo eu já, tu a mi abundante.</l>  
</lg>
```

Con esto, podemos añadir materialmente una serie de informaciones que, intuitivas e incluso inmediatas para nosotros, son ignoradas por ordenadores cuando no están explícitamente insertadas en el sistema: la división en estrofas y los tipos de cada una de ellas, el tipo de rima o de poema, los fenómenos de acomodación silábica utilizados y así sucesivamente. Por ello, la utilización de ese formato TEI para codificar obras poéticas, además de permitir que herramientas computacionales puedan procesar su contenido, deja abierta la posibilidad de que, con base en los mismos archivos, se creen versiones para ser leídas, por ejemplo, en la Web.

Para llegar a la escansión al final del proceso, *Aoidos* toma los poemas de un archivo TEI y produce una transcripción fonética para

cada verso, lo que se hace con base en reglas internas, de modo que los diccionarios no son necesarios y los neologismos se incorporan sin problemas. El objetivo de esa transcripción fonética no es producir una representación exacta de cómo los poemas deben ser leídos en voz alta; se trata solo de un artificio para representar los sonidos del portugués de forma más regular de lo que permite la ortografía y con vistas solo a la escansión. Tenemos tres etapas, todas realizadas sobre las palabras individuales de cada verso. En primer lugar, la sílaba tónica (o su ausencia) es determinada, sea por la presencia de diacríticos (como en "atônito"), sea por la consonante final que denuncia el acento ("mordaz"), sea por el análisis de los grafemas finales de la palabra ("rainha"). En seguida, el sistema hace una división silábica, convirtiendo palabras como "abstração" en "abs-tra-ção" y "vácuo" en "vá-cuo". Finalmente, un mapeo lleva cada grafema (o grupo de grafemas) a fonemas, llevando "distante" a /dis'tã.ti/ y "mocinhas" a /mo'si.ɲas/. El tercer verso de la cuadra tomada como ejemplo queda, por lo tanto, así²:

'pɔ.brɪ ti 'vɛ.ʒu a 'ti 'tu a 'mi ã.pe'ɲa.du

Esta transcripción fonética, hecha palabra por palabra y de forma totalmente determinista (en ella "negociante" nunca tiene 4 sílabas, siempre 5: "ne-go-ci-an-te"), sirve de base para producir alternativas de pronunciación para el verso como un todo. La idea es generar el mayor número posible de diferentes maneras de declamar el mismo verso, teniendo en cuenta los metaplasmos disponibles en la lengua y posiblemente explorados por el poeta. El verso mencionado, por ejemplo, si leído de acuerdo con la transcripción fonética inicial, contaría 13 sílabas poéticas. A partir de ella, el sistema encuentra la posibilidad de sinalefa entre "vejo" y "a" y produce una alternativa de lectura para el verso que incorpora esa sinalefa. También entre "tu" y "a" el sistema percibe que una sinalefa puede ser hecha, y registra una tercera alternativa. Una cuarta alternativa es realizar ambas sinalefas. Para ese verso, *Aoidos* encontró un total de 14 alternativas. No todas, sin embargo, son igualmente plausibles. El oído moderno dice que la sinalefa entre "vejo" y "a" es bastante deseable, pero

aquella entre "tu" y "a" es prácticamente forzada. *Aoidos*, al producir las diferentes alternativas, mantiene el registro de qué metaplasmos fueron empleados y el índice de "naturalidad" — que es evidentemente bastante subjetivo — de cada alternativa de pronunciación del verso. En el caso del ejemplo, la alternativa más "natural" encontrada por el sistema fue aquella en que el único metaplasmo empleado es la sinalefa de entre "vejo" y "a", quedando así la transcripción fonética correspondiente:

'pɔ.brɪ ti 'vɛ.ʒwa 'ti 'tu a 'mi ã.pe'ja.dʊ

Esta "naturalidad", evidentemente, no tiene en cuenta el contexto en que el verso aparece; el sistema aún no sabe que el soneto está escrito en decasílabos y que esa alternativa de 12 sílabas claramente no es la correcta. Tampoco el sistema tiene la sofisticación de percibir que está escrito "mi" y no "mim", lo que probablemente es una indicación al lector, por parte del poeta o del editor, de que se debe hacer la sinalefa entre "mi" y "empenhado". En resumen, *Aoidos* tiene reglas para varios metaplasmos, incluyendo sinalefas, sinéresis, síncope, epéntesis, sístoles, etc.

En posesión de todas las alternativas de pronunciación de todos los versos, el sistema vuelve a su análisis del poema como un todo, a fin de determinar la cantidad de metros utilizados. Para hacerlo, se consideran cuáles son las alternativas de pronunciación más "naturales" de cada verso, así como el tamaño silábico mínimo y máximo del poema y de las estrofas. En el soneto del que fue tomado el cuarteto de ejemplo, el único tamaño silábico compatible con todos los versos es 10, o sea, todos los versos pueden ser escandidos en 10 sílabas, pero hay versos que no pueden ser escandidos en 9 o 11. Si hubiera más de un tamaño silábico compatible con todos los versos del poema, el sistema consideraría la "naturalidad" conjunta de los versos escandidos en diferentes números de sílabas. En total, *Aoidos* intenta cuadrar los poemas en una de cuatro categorías: a) todos los versos del poema tienen la misma métrica, o sea, el poema es isométrico; b) los versos del poema poseen dos métricas distintas, lo

que sería el caso de versos heroicos (10 sílabas) y sus quebrados (6 sílabas); c) las estrofas del poema, cuando se consideran aisladamente, contienen uno o dos metros; d) el poema (y sus estrofas) contiene más de dos metros.

Cuando el sistema ya ha decidido cuántos metros están presentes en el poema, llega a su fase final: encontrar, entre las alternativas de pronunciación de cada verso, aquella que mejor encaja en el contexto rítmico-acental en que aparece. La selección de la alternativa supuestamente (o posiblemente) más natural entre aquellas compatibles con el metro del poema puede no ser suficiente, ya que diferentes alternativas del mismo tamaño pueden presentar diferentes ritmos. Para el segundo verso del primer tercer del soneto de ejemplo, el sistema encontró dos escansiones decasilábicas:

Pe-	las	dro-i	ga-	s i-	nú-	teis,	que a-	be-	lhu-	da
Pelas	dro-	ga-	s i-	nú-	teis,	que	a-	be-	lhu-	da

En la primera alternativa, el sistema hizo la sinalefa entre "que" y "a", mientras que la segunda mantuvo allí el hiato e hizo un síncope en "pelas", como si la palabra estuviera escrita "p'las" en la edición. El esquema rítmico de las dos alternativas es diferente: la primera es un típico heroico 3-6-10, mientras que la segunda posee un rarísimo 2-5-10. En el caso de los esquemas rítmicos encontrados, el sistema puede considerar el contexto del poema en que el verso aparece: percibiendo que casi todos los versos pueden ser "naturalmente" escandidos con la 6ª tónica, el sistema escoge la alternativa 3-6-10 del verso de ejemplo. Opcionalmente, se puede emplear un catálogo de ritmos en que el usuario expresa, de manera estadística, cuáles son los patrones rítmicos más aceptables en la tradición poética en

que el poema se sitúa. Al no usar un catálogo de ritmos, el usuario permite que el sistema analice cada poema de forma bastante individual, sin nociones preconcebidas de cuáles ritmos un verso debe o no presentar; por otro lado, el empleo de un catálogo posibilitaría que fueran escandidos poemas pequeños, que tengan un número insuficiente de versos para que el sistema entienda cuáles son los ritmos preferidos por el poeta. En pocas palabras, el sistema puede solo concluir que Camões prefiere los esquemas 6-10 y 4-8-10 en *Os Lusíadas*, pero necesitará ayuda del usuario para escandir un cuarteto cualquiera tomado de forma aislada.

La exactitud de *Aoidos* fue evaluada en un experimento en que 1080 versos fueron escandidos por el sistema y manualmente conferidos. Estos versos fueron seleccionados de forma aleatoria de un corpus mayor, correspondiendo al 1% de un total de 107 438 versos. El corpus completo contenía obras poéticas de 10 autores, variando temporalmente de Camões a Delminda Silveira, pasando por Gregório de Matos, Gonçalves de Magalhães y Augusto dos Anjos. La tasa de acierto global del sistema fue del 99.0%, variando entre 97.5 y 100.0% conforma la obra poética. De Gregório de Matos fueron conferidos manualmente 358 versos, un 99.2% de los cuales estaban correctamente escandidos.

3. Preparación y escansión de la obra poética de Gregório de Matos

Para hacer una comparación entre los dos análisis producidos por Chociay y *Aoidos*, es necesario que ambos se basen en el mismo corpus. El profesor brasileño trabajó sobre la segunda edición de la obra completa de Gregório publicada por James Amado en 1990 (Chociay, 1993: 52), que no presenta diferencias con la de 1992, que utilizamos en *Aoidos*. En esa tercera edición, en dos volúmenes, no hay ninguna mención en cuanto a diferencias con

relación a la edición anterior. Entre la publicación de las dos primeras ediciones se pasaron veintiún años; la tercera surgió apenas dos años después de la segunda, lo que evidencia que no hubo una gran revisión. Además, los versos cuya localización en los volúmenes y en las páginas de la 2ª edición es dada por Chociay aparecen exactamente en los mismos volúmenes y páginas de la 3ª. Creemos, por lo tanto, que poco o nada se pierde cuando utilizamos la 3ª edición y Chociay, la 2ª. Cuando nos referimos, en adelante, a la "edición de James Amado", lo estamos haciendo de forma genérica a la 2ª o a la 3ª.

El proceso de establecimiento del corpus en formato TEI se inició con la adaptación de la edición accesible a la Literatura Digital (Matos, 2017), que fue creada a través de la digitalización de la 3ª edición en papel de James Amado (Matos, 1992). Los 40 archivos HTML que componen esta edición digital se han convertido en TEI de forma semiautomática: los programas de ordenador, desarrollados específicamente para esa tarea, fueron utilizados por nosotros para convertir gradualmente cada uno de los archivos HTML en TEI. De forma general, cada línea encontrada en el cuerpo de los poemas fue vista como un verso y codificada con el elemento <l> del TEI. Las estrofas, codificadas con el elemento <lg>, se identificaron a través de espacios verticales en blanco. Sin embargo, se consideraron algunos casos especiales. Por ejemplo, varios poemas (en la edición digital e impresa) contienen punteados que coordinan dos versos en una misma línea; en el corpus, tanto el verso a la izquierda como a la derecha del punteado ganó su propio elemento, a excepción de un soneto, con *incipit* "Na oração, que desaterra aterra", en que las rimas internas habían sido destacadas con un punteado, que fue entonces ignorado.

Ya codificado en el formato TEI, el corpus pasó por una revisión inicial. La cantidad de versos en las estrofas y la cantidad de estrofas en los poemas fue verificada. Se percibió, por ejemplo, que el poema "Pois no que toca a guardar" contenía un número sospechoso de versos: 71. El cotejo entre la edición digital y la impresa permitió

constatar que un verso desapareció en el proceso de digitalización. El poema, con el verso restablecido, pasó a contar con 72 versos (18 cuartetos). En total, solo 4 versos fueron insertados en la edición digital porque constaban en la impresa; pero muchos casos sospechosos no pudieron resolverse, ya que los versos faltan en la propia edición impresa. En esa revisión también se localizaron los poemas que Gregório de Matos escribió en español y que fueron especialmente marcados con descriptores TEI, para que *Aoidos* no los intentase escandir.

Después de la revisión inicial, los poemas de Gregório fueron sometidos al análisis del programa. El sistema apuntó entonces una serie de versos que no pudo escandir. Se examinaron, uno a uno, para que errores en la edición digital (es decir, discrepancias con la edición impresa) fueran corregidos y para que abreviaturas y extranjerismos fueran debidamente marcados en el corpus. En ese paso se encontraron 240 versos con errores en la edición digital. Muchos de ellos eran fallos en la etapa de OCR (Optical Character Recognition) de la digitalización: así, 130 versos eran incorrectos porque decían "lhe" (con la letra "i" mayúscula al principio) en vez de "lhe" (con la letra "L" minúscula). También hubo casos de palabras que perdieron el diacrítico en la digitalización, como "roido" en lugar de "roído". Se han extendido las abreviaturas (aquí consideradas de forma amplia como cualquier fragmento textual no escrito por extenso, como son típicamente los casos de "D.", en lugar de "Dom", y "Fr.", en lugar de "Frei"), así como las asociaciones de letras, como "ff" ("efes"). Los extranjerismos encontrados son expresiones en español (como en el verso "*estos me lleban dizia*") o en latín ("e dizendo o *jube domine*"). Es importante señalar que tanto los errores de la edición digital como las marcas de abreviaturas y extranjerismos solo fueron percibidos cuando el sistema no fue capaz de escandir un verso, lo que ya nos revelaba una primera utilidad hasta entonces insospechada de *Aoidos*, es decir, la detección de problemas de edición. Así, se percibió que faltaba un diacrítico en "roido" porque el sistema no consiguió escandir el verso "Saí desta canalha tão roido" en el contexto decasílabico en que aparece. Otros problemas

similares: la abreviatura "Fr." no puede ser leída por el sistema como una palabra, lo que es señalado durante su operación; y la palabra "domine", si se lee como si fuera portugués, es grave, pero en latín ella es esdrújula, lectura esta que es necesaria para escandir el verso con siete sílabas y no con ocho. Por fin, quedaron 45 versos que, aunque fueron manualmente cotejados con la edición impresa y cuidadosamente examinados, desafían la escansión. El siguiente verso, por ejemplo, aparece en un contexto que requiere siete sílabas:

Terra, que não parece

En la reciente edición del Códice Asensio-Cunha (Matos, 2014, 193), ese mismo verso viene así:

Terra, que não se parece

Se trataría, tal vez, de un error presente en la propia edición de James Amado. Sin embargo, en la versificación medieval, se encuentran muchos ejemplos de hiato en palabras que, habitualmente, serían consideradas como monosílabos, lo que implicaría aceptar la pronunciación de "não", originariamente con una sola sílaba, como "nã-o", ahora con dos sílabas. También el hecho de que se trataba a lo mejor de poemas cantados, podría justificar esa separación de "não" en dos sílabas. De todos modos, como no se pretendía hacer una edición crítica de esos poemas de Gregório, no se hizo ningún intento de establecer las versiones supuestamente correctas de esos 45 versos, que fueron solo marcados con inescandibles en el corpus.

Algunos versos fueron excluidos o ignorados en las escansiones y recuentos de que trata este artículo. En primer lugar, los poemas del llamado "coronista resuçitado" fueron completamente descartados, por no haber sido escritos por Gregorio, conforme queda claro al verificar que la autoría de ellos es atribuida al final a Tomás Pinto Brandão. Varios otros casos (y no solo con respecto a Pinto Brandão)

fueron censados y analizados por Francisco Topa (1999). Aun los epígrafes y los motes, a pesar de permanecer en el corpus, fueron desconsiderados para propósitos de escansión y no se incluyen en ningún tipo de estadística.

Es importante decir que Chociay también ha utilizado en su trabajo la edición de la obra poética de Gregório publicada por la Academia Brasileña de Letras: "Aunque hemos hecho el recuento basado en la segunda edición de James Amado, incluimos versos que no figuran en esta edición (fueron descubiertos en el cotejo con la edición de la Academia)". Chociay afirma que la edición de la Academia, "a pesar de ser la primera que se pretendió" completa "y del inestimable servicio que prestó para el conocimiento de la poesía gregoriana, se afilia a una tradición de falso moralismo" (Chociay, 1993, 24) y que "debe ser consultada con reservas y cautela" (Chociay, 1993, 25). Incluso con sus reservas con relación a la edición de la Academia, Chociay la utilizó porque la de James Amado ofrece "soluciones ingenuas" (Chociay, 1993, 30) en diversos casos.

4. Aidos frente a Chociay

En el libro *Os Metros do Boca* (Chociay, 1993), se discute la obra poética de Gregório de Matos bajo varias perspectivas: el autor evalúa las ediciones de la poesía completa del poeta en la época, hace comentarios acerca de su técnica versificatoria y, de particular interés al presente artículo, cuantifica los tipos de versos que utilizó el poeta. En la p. 52, presenta una tabla resumiendo los tipos de versos y las frecuencias absolutas y relativas de ellos en la obra del Boca del Infierno, que es aquí íntegramente reproducida en la Tabla 1. Sobre esos números, Chociay comenta que, por incluir versos que no constan en la edición de James Amado, "y aún aceptando un pequeño margen de error, la cifra mencionada, 33 123, no tiene pretensión de exactitud absoluta."

No todos los tipos de verso presentados en la Tabla 1 corresponden de forma biunívoca a tamaños silábicos. Así, un verso cuya escansión acuse 6 sílabas puede, por ejemplo, constar en la tabla bajo el tipo "hexasílabo", "heroico quebrado" o "verso de seguidilla"; los versos agrupados bajo el rótulo "redondillo roto", por otro lado, pueden ser escandidos en 3 o 4 sílabas. Una consecuencia de esta no correspondencia directa entre tipos y tamaños es la imposibilidad de producir una tabla similar a partir de la escansión automática. A partir de los resultados de la herramienta, es fácil calcular cuántos versos fueron escandidos con 6 sílabas, pero no se puede saber inmediatamente cuáles de ellos son hexasílabos y cuáles son heroicos quebrados sin considerar el contexto en que aparecen. *Aidos*, por lo tanto, fue construido con algunas reglas de uso específico que miran más allá del verso, en el intento de imitar el esquema clasificatorio adoptado por Chociay. Las reglas son cuatro: una para los redondillos rotos, otra para los quebrados de ovillejo (como los denomina Chociay, o simplemente ovillejo), una tercera para los hexasílabos, heroicos quebrados y versos de seguidilla, y una última para los decasílabos.

Versos	Número de veces	%
Disílabo	15	0.04
Redondillo quebrado	251	0.76
Quebrado de ovillejo	66	0.20
Tetrasílabo	95	0.29
Pentasilabo	317	0.96
Hexasílabo	168	0.51
Heroico quebrado	123	0.37
Verso de seguidilla	52	0.16
Heptasílabo	28 028	84.1
Eneasílabo	96	0.30
Decasílabo italiano	3798	11.46
Decasílabo 5-10	16	0.05
Decasílabo 4-7-10	10	0.03
Endecasílabo	88	0.26
Totales generales	33 123	100

Tabla 1. Tipos de versos y su frecuencia en la obra poética de Gregório de Matos, conforme originalmente publicado por Chociay (1993)

Para Chociay, un redondillo roto es "tradicionalmente el verso de tres sílabas, siempre que se usa en asociación con el redondillo mayor en diferentes tipos de estrofa o sistema estrófico" (Chociay, 1993, 53). De hecho, tanto el trisílabo, como el tetrasílabo pueden funcionar como quebrados de redondillo; además, en la versificación española, los quebrados de ovillejo, empleados originalmente por Cervantes, alternaban versos redondillos con disílabos. En el caso de Gregório, Chociay argumenta que, a pesar de presentar a menudo 3 sílabas, la escansión de esos versos quebrados podría resultar, cuando se consideran aisladamente, en 3 o 4 sílabas poéticas. El profesor argumenta que los fenómenos de sinafía y compensación (Chociay, 1993, 45-47) explicarían buena parte de los tetrasílabos aparentes.

Las dos estrofas siguientes contienen un caso de cada fenómeno, como muestra Chociay (1993, 54):

É mau meteres o calvo
entre tanta pentelheira,
e sair co'a cabeleira
encrespadinha?

Que mau é Mariquitinha,
quando está com seus lundus
fazer-vos com quatro cus
o rebolado?

En la primera estrofa, ocurre una sinalefa entre la sílaba final de "cabeleira" y la inicial de "encrespadinha", de manera que para el cuarto verso sobra solo el fragmento "-crespadinha", con 3 sílabas. En la segunda estrofa, la palabra aguda "cus" del tercer verso permite que el artículo definido "o" del cuarto verso ocupe la posición final del tercero, dejando un total de 3 sílabas para el cuarto verso. Casos hay, sin embargo, en que ningún fenómeno es capaz de reducir a 3 la cantidad de sílabas poéticas del último verso de la estrofa; Chociay llama a la sílaba excedente de "perdida". Un ejemplo es la siguiente estrofa, en la que uno tendría dificultades para escandir el cuarto verso en menos de 4 sílabas:

Não sei, para que é nascer
neste Brasil empestado
um homem branco, e honrado
sem outra raça.

Ahora bien, en ese caso, dos posibilidades no son tenidas en cuenta por Chociay. La primera es la eclipsis³ proveniente de prosodia más popular (cuyos ejemplos, como atestigua el propio Chociay, abundan en Gregorio), que daría "se'outra raça", llegando al patrón trisílabo. Una segunda posibilidad sería ver, en esas aparentes anomalías en la métrica, simplemente el efecto de ser el verso cantado, como ya se

comentó más arriba: el compás musical podría ser mantenido, pero no necesariamente la isometría silábica (esa es la causa posible de buena parte de los pretendidos pies quebrados de diversas composiciones presentes en los cancioneros medievales).

En lo que se refiere al programa, *Aoidos* no tiene en cuenta, en su etapa actual, procesos de acomodación que no respeten el límite del verso. De esta forma, no es capaz de identificar casos de sinafía o de compensación. La regla que intenta detectar los redondillos quebrados, por lo tanto, adopta una estrategia más simple: si una estrofa tiene al menos dos versos heptasilábicos, entonces todos los versos con 3 o 4 sílabas de aquella estrofa son considerados redondillos quebrados. Los versos de 3 o 4 sílabas que no se encuentran en tales estrofas no son afectados por la heurística, como tampoco lo son los heptasilábicos.

La segunda regla tiene la función de identificar los quebrados de ovillejo. Chociay afirma:

Según los tratadistas españoles, el ovillejo es un sistema que presenta una primera semiestrofa formada por tres heptasílabos interrogativos, a cada uno de los cuales sigue una palabra-respuesta que repite la rima del verso respectivo; la segunda semiestrofa consiste en un redondillo ABBA que resume el sentido de los versos anteriores y cuyo último verso recoge las tres palabras-respuesta de tales versos. (Chociay, 1993, 99)

Un ovillo muy conocido de Gregorio — y que sigue precisamente la receta dada por Chociay — es el siguiente:

Que falta nesta cidade? Verdade
Que mais por sua desonra? Honra
Falta mais que se lhe ponha..... Vergonha.

O demo a viver se exponha,
por mais que a fama a exalta,
numa cidade, onde falta

Verdade, Honra, Vergonha.

La heurística para detectar la presencia de los quebrados de ovillejo funciona, por lo tanto, de la siguiente manera. Para que un poema sea compuesto de ovillejos, es necesario que: a) el poema contenga un número par de estrofas (ya que cada ovillejo posee dos); b) las estrofas impares contengan seis versos (que corresponden a las tres preguntas y a las tres respuestas); y c) las estrofas pares contengan cuatro versos (tres resumiendo el sentido de los anteriores y uno recogiendo las palabras-respuesta). Ahora bien, los versos de las palabras-respuesta son justamente los quebrados de ovillejo que la heurística quiere identificar, de manera que ella considera que los versos pares de las estrofas impares de tales poemas son quebrados de ovillejo, independientemente del número de sílabas que ellos puedan contener. Los demás versos de los ovillejos no reciben ningún tratamiento especial.

La tercera regla engloba todos los versos que *Aoidos* escandió en 6 sílabas y algunos que el sistema juzgó tener 4 sílabas; esos versos pueden figurar en la tabla de Chociay como hexasílabos, heroicos quebrados o versos de seguidilla. La categoría de los hexasílabos engloba los versos de 6 sílabas que aparecen en grupos — típicamente en estrofas isosilábicas; los heroicos quebrados son los versos de 6 sílabas que aparecen junto a los decasílabos. Los versos de seguidilla pueden tener 4 o 6 sílabas; En el caso de que el modelo de seguidilla adoptado por Gregório es el de una estrofa de cuatro versos, siendo los impares hexasílabos y los pares tetrasílabos, sin rima en los versos impares y con rima asonante en los pares (Chociay, 1993, 73). La heurística, por lo tanto, trabaja de ese modo: si un verso de 6 sílabas se encuentra en una estrofa en la que por lo menos un 75 % de los versos contiene 6 sílabas, entonces ese verso es clasificado como hexasílabo; un verso de 4 o 6 sílabas en una estrofa que contenga por lo menos 2 versos de cuatro sílabas y 2 de seis sílabas es clasificado como verso de seguidilla; cualquier verso que no sea clasificado ni como hexasílabo, ni como de seguidilla es entonces clasificado como quebrado de heroico. Los

versos de 4 sílabas que no han sido clasificados como de seguidilla ya no reciben ningún tratamiento por parte de esa heurística; los versos de 4 sílabas que ya fueron clasificados como redondillos quebrados o quebrados de ovillejos no sobrevivieron hasta esa heurística.

La última regla tiene por objeto examinar los decasílabos y encajarlos en una de las tres categorías definidas por Chociay: la de los decasílabos italianos, la de los versos con esquema 5-10 y la de los versos con esquema 4-7-10. En la nomenclatura de Chociay, los italianos son los decasílabos de la nueva medida, los heroicos y los sáficos; los otros, más excepcionales en la obra de Gregório de Matos, son representantes de una tradición ibérica más antigua y no aparecen en el mismo contexto que los italianos. La heurística trabaja, por lo tanto, de la siguiente manera: si un decasílabo está acentuado en la 6ª sílaba o entonces simultáneamente en la 4ª y en la 8ª, es italiano; si la 5ª sílaba son tónicas, tiene esquema 5-10; si la 4ª y la 7ª están acentuadas, tiene esquema 4-7-10; si no se ajusta a ninguno de estos casos, se considera italiano. Es importante aclarar que el decasílabo con acento rítmico en la quinta sílaba no solo es raro, como impropio (o es raro justamente por ser inapropiado). La tendencia es que, en ese caso, se pierda la elisión rítmica entre los dos hemistiquios, que dejan de ser grupos silábicos dentro de un verso único, para convertirse en dos versos independientes de cinco sílabas cada uno. En el caso de los poemas de Gregório de Matos, *Aoidos* encontró, en los 4141 decasílabos analizados, aproximadamente un 0.7 % de versos con acento en la quinta sílaba. Un análisis más detenido podría tal vez mostrar que se trataba del antiguo verso doble medieval, el bipentasílabo. Ya el verso decasílabo con acento en la cuarta y en la séptima sílabas fue ampliamente empleado en la literatura provenzal de la Edad Media, siendo por ello justamente denominado decasílabo provenzal.

Aoidos, utilizado con esas cuatro heurísticas, produjo los resultados exhibidos en la Tabla 2. La diferencia total, de 1272 versos, con respecto al análisis realizada por Chociay, es bastante significativa y

merece ser examinada.

Categoría	Aoidos	Chociay	Diferencia	%
Disílabo	15	15	0	0,0
Trisílabo	3		3	100,0
Redondillo quebrado	256	251	5	2,0
Quebrado de <i>ovillejo</i>	75	66	9	12,0
Tetrasílabo	155	95	60	38,7
Pentasílabo	319	317	2	0,6
Hexasílabo	168	168	0	0,0
Heroico quebrado	194	123	71	36,6
Verso de seguidilla	48	52	-4	-8,3
Heptasílabo	28 747	28 028	719	2,5
Octosílabo	21		21	100,0
Eneasílabo	107	96	11	10,3
Decasílabo italiano	4141	3798	343	8,3
Decasílabo 5-10	30	16	14	46,7
Decasílabo 4-7-10	14	10	4	28,6
Endecasílabo	100	88	12	12,0
Dodecasílabo	1		1	100,0
Erro	1		1	100,0
	34 395	33 123	1272	3,7

Tabla 2. Comparación de los resultados obtenidos por *Aoidos* y por Chociay. Los porcentajes fueron computados dividiendo la diferencia por el valor encontrado por el programa

Comencemos con los poemas en español. El sistema ciertamente los incluyó en la cuenta, pero no está claro si Chociay lo hizo; además, los propios números de los versos en español no coinciden. Afirma el profesor que Gregório escribió 21 poemas en ese idioma, con un total de 584 versos, incluidos 472 redondillos y 112 decasílabos (Chociay, 1993, 51). Con los poemas codificados en formato TEI, obtuvimos el mismo número de poemas, 21, pero llegamos a un número diferente de versos en español, 555, o sea, 29 menos que Chociay. Es difícil

especular sobre las causas exactas de esa divergencia numérica, pero podemos demostrar que los números de Chociay no son correctos. En su estudio, afirma que son 112 los decasílabos escritos en español; la Tabla 3 muestra las páginas, en la edición de James Amado, en la que aparecen sonetos decasilábicos en español: son nueve sonetos, que totalizan, por lo tanto, $9 \times 14 = 126$ decasílabos, 14 más que los enumerados por Chociay. Si ajustamos su cuenta para tener en cuenta este número de decasílabos, el total de versos en español saltaría de 584 a 598. Los recuentos de versos hechos por *Aidos* no tienen en cuenta los versos de los motes, ya que se repiten a lo largo de los poemas; si, por un instante, dejamos de lado esa restricción y contamos todos los versos en español, incluyendo los versos de los motes, *Aidos* encuentra 597 versos, lo que es muy próximo a la cantidad ajustada de 598. Existe, sin embargo, un poema en español, con *incipit* " Coraçon: siente tu anhelo ", con estructura similar a varios otros escritos por Gregório: un mote seguido de cuatro décimas. En la tercera décima de ese poema, sin embargo, falta un verso, lo que se puede constatar a las pp. 422-423 de la edición de James Amado. Chociay posiblemente consultó la edición de la Academia Brasileña de Letras y por este motivo tuvo acceso al verso que no consta en la edición de James Amado; ese verso, si está presente en nuestro corpus, elevaría nuestra cuenta a los 598 versos (ajustados) de Chociay. Si el razonamiento es correcto, Chociay se saltó un soneto en español cuando contó versos pero no lo saltó cuando contó poemas, ya que, sin ningún ajuste, *Aidos* encontró los mismos 21 poemas que Chociay.

76	76	158	242	521	521	529	767	1180
----	----	-----	-----	-----	-----	-----	-----	------

Tabla 3. Números de las páginas en que los sonetos decasilábicos en español escritos por Gregório de Matos aparecen en la edición de James Amado

No está claro, sin embargo, como ya se ha dicho, si los poemas en español fueron incorporados en la cuenta de Chociay junto con los versos en portugués. Por un lado, tales poemas son parte indisoluble de la obra poética de Gregório de Matos y nos parece muy natural que debamos contarlos. Por otro lado, Chociay afirma:

Para tener una idea de esa poca importancia del castellano, basta recordar que son más de veintiocho mil los redondillos escritos por Gregorio en portugués (Chociay, 1993, 51).

Ahora bien, si la tabla producida por el investigador indica que hay 28 028 redondillos y si afirma también que hay más de 28 000 redondillos en portugués, quedaría margen para apenas 27 redondillos en español; él mismo dice, sin embargo, que hay 472 redondillos en ese idioma. La conclusión lógica, aunque implícita, es que la tabla no incluye versos en español. Chociay, sin embargo, ciertamente contó algunos versos en español: aquellos que aparecen en los poemas escritos mayoritariamente en portugués. La confirmación es que Chociay cuenta 88 endecasílabos, oriundos de los poemas "Sendo pois o alterar da moeda" y "Marinículas todos os dias" (Chociay, 1993, 65). Los dos poemas suman $43 + 1 = 44$ cuartetos que alternan versos de 9 y 11 sílabas, de forma que hay $44 \times 2 = 88$ endecasílabos. Sin embargo, el segundo poema mencionado contiene estos dos versos (no consecutivos) escritos enteramente en español:

Los ordeñadores se alquilan aquí
Pudiera cogerlos en un mal Latim

Chociay incluye, por lo tanto, estos dos versos en español en la cuenta de los endecasílabos, y probablemente hizo lo mismo con muchos otros versos en aquel idioma que aparecen junto a versos en portugués.

Podríamos disminuir la diferencia entre los totales de versos de 1272 a 717 si excluyéramos del recuento de *Aoidos* los 555 versos en español, siguiendo la lógica implícita de que Chociay no los contó. Si consideráramos, sin embargo, toda la lógica implícita del profesor, deberíamos también incluir los versos de los motes, como Chociay aparentemente hizo en los poemas en español, lo que agregaría 307 a nuestra cuenta, parcialmente compensando la sustracción de 555 y manteniendo la diferencia por encima de 1000 versos.

Incluso ignorando el problema de los versos en español, se sabe que Chociay no contó algunos versos. Él afirma que Gregório de Matos compuso 196 sonetos, pero un análisis automático produce la cuenta de 217 sonetos. Una conferencia manual fue hecha para averiguar que el número de sonetos es, como mínimo, 217; las páginas en que cada uno de ellos aparece figuran en la Tabla 4. Incluso si descontamos los 9 sonetos en español, la diferencia todavía existe: son $217-9-196 = 12$ sonetos que Chociay no contó, es decir, 168 versos.

33	33	66	67	67	68	68	69	69	76
76	77	77	78	80	81	81	84	85	88
123	123	124	125	125	126	129	133	137	137
138	146	157	157	158	158	159	162	163	167
175	175	177	186	187	188	188	189	197	198
198	199	199	200	204	206	210	215	228	229
239	242	242	243	275	318	319	319	320	320
321	322	333	347	370	402	403	404	406	406
407	407	410	411	411	412	415	415	420	424
425	426	426	431	432	432	447	507	507	510
511	511	514	517	518	519	520	521	521	527
529	537	543	543	544	544	546	555	557	558
560	568	591	627	627	634	639	639	640	641
641	648	649	649	650	650	657	662	662	678
713	719	725	747	747	748	749	749	751	752
760	761	766	767	768	792	802	823	838	840
843	851	852	860	860	868	868	877	878	879
879	880	880	883	888	899	900	901	910	924
930	987	1045	1056	1071	1075	1112	1113	1113	1117
1120	1120	1129	1130	1137	1138	1163	1170	1180	1191
1192	1195	1202	1202	1203	1203	1204	1204	1205	1206
1206	1209	1211	1215	1216	1216	1222			

Tabla 4. Números de las páginas en que los sonetos decasilábicos, en portugués y en español, aparecen en la edición de James Amado

Aidos encontró 75 quebrados de ovillejo, mientras que Chociay apunta la existencia de solo 66 en la obra de Gregorio de Matos. Sin

embargo, Chociay afirma también que cuatro fueron los poemas escritos en ovillejos (Chociay, 1993, 99-103) y cita las páginas de la edición de James Amado en que esos poemas aparecen: 55-56, 56-58, 830 y 1051-1052. Al consultar la edición impresa y hacer un recuento manual de los quebrados de ovillejo de esos cuatro poemas, percibimos que contienen 18, 27, 9 y 21 versos quebrados de ese tipo, respectivamente. La cuenta correcta, por lo tanto, es la del sistema: $18 + 27 + 9 + 21 = 75$. Los 9 versos más encontrados por el sistema corresponden justamente al número de quebrados del poema de la página 830, lo que es evidencia de que ese poema lo saltó Chociay en su cuenta.

El recuento automático de *Aoidos* reveló la presencia de 155 tetrasílabos, pero Chociay contó solo 95. Dos poemas bastan para que la suma de 95 versos tetrasílabos sea superada: uno con *incipit* "Vá de retrato" (págs. 183-186 de James Amado) y otro, "Vá de aparelho" (págs. 846-849). El primero tiene 84 tetrasílabos y el segundo, 57, totalizando 141 versos de ese tipo. El propio Chociay menciona los dos poemas a la página 76 de su libro, incluso con referencia a las páginas de la edición de James Amado. La evidencia, por lo tanto, de que Chociay se saltó por lo menos $141 - 95 = 46$ tetrasílabos es bastante clara. Descubrir exactamente qué versos fueron contados por Chociay para llegar a la suma de 95 y el origen de la diferencia entre la cuenta de él y la de *Aoidos* es más difícil, pero claro está que la cantidad de tetrasílabos es superior a aquella apuntada por Chociay.

El único tipo de verso en que la balanza se vuelve hacia el lado de Chociay es el verso de seguidilla; *Aoidos* contó 48, y Chociay, 52. En la página 58 de su libro, afirma el investigador que Gregório utilizó ese verso en dos poemas; el primero de ellos, "Retratar ao bizarro", fue correctamente contabilizado por el sistema, resultando en los 48 versos apuntados. El sistema, sin embargo, no logró identificar la siguiente estrofa del segundo poema indicado por Chociay como una cuadra de versos de seguidilla:

Não Anica, te escondas,
aparece sempre,
que o ser bem parecida
disso depende.

Aidos no fue capaz de hacerlo porque no pudo escandir el segundo verso del cuarteto en cuatro sílabas. No está claro cómo Chociay escande el verso. El sistema, al no clasificar estos versos como de seguidilla, los clasificó en otras categorías. En un recuento manual, podríamos considerar la posibilidad de una aféresis en “aparece”, lo que resultaría en un verso, ahora sí, con tres sílabas: “pa-re-ce-sem(pre)”. Sin embargo, ello implicaría una regla que no forma parte todavía de las heurísticas de *Aidos*.

Entre todas las discrepancias mencionadas, nos gustaría resaltar la concordancia que emerge cuando los números son considerados en su globalidad. Considere uno, por ejemplo, los dos gráficos que aparecen en la Figura 1; el gráfico a la izquierda representa los números de *Aidos* y aquel a la derecha, los de Chociay. Las proporciones de los tipos de versos son extremadamente similares en ambos gráficos. En los dos, salta a los ojos que hay un predominio masivo de los heptasílabos; que los decasílabos italianos participan como coadyuvantes; y que los demás tipos de versos son minoría absoluta. Cuando nos distanciamos un poco de los detalles, lo que vemos es que tanto los versos olvidados por Chociay (algo entre 700 y 1.300) como los fallos de escaneo por el sistema (aproximadamente 288 versos) no nos impiden caracterizar al poeta.

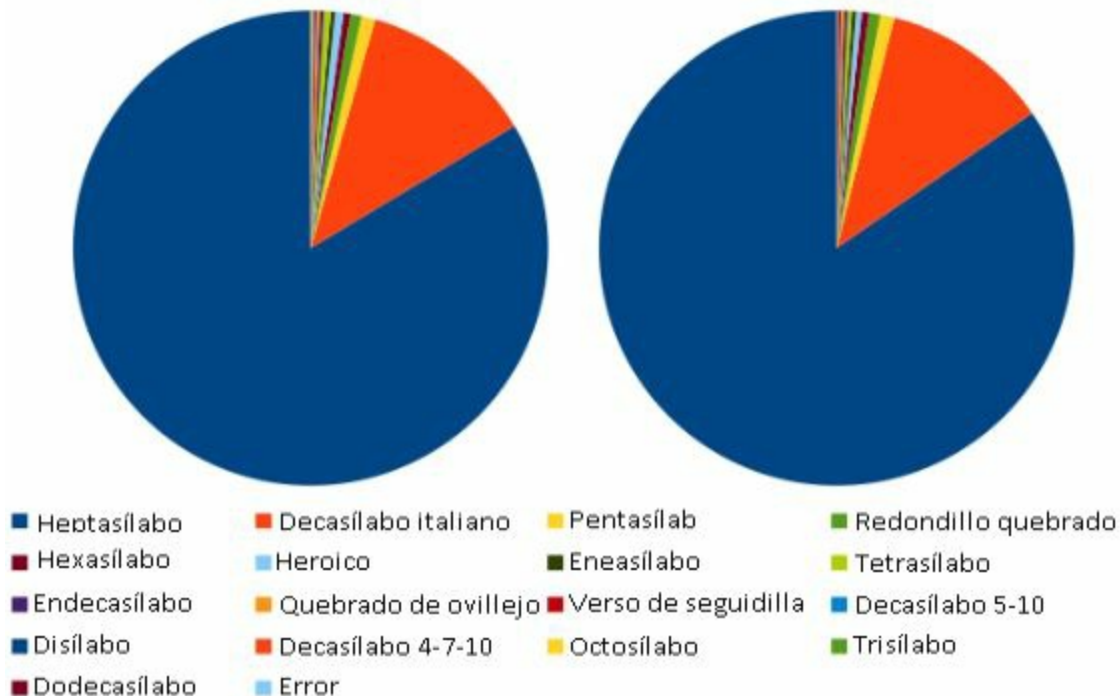


Figura 1. Proporción de los diversos tipos de versos, de acuerdo con *Aoidos* (a la izquierda) y *Chociay* (a la derecha)

5. Algunas conclusiones

Los resultados cuantitativos producidos por *Aoidos* son similares, pero no idénticos, a los relatados por *Chociay*. Por un lado, este ciertamente tiene una sofisticación muy superior a la del sistema. Al encontrar un verso en que parece faltar una sílaba, *Chociay* puede consultar ediciones alternativas para encontrar una versión más coherente y la sílaba que faltaba; el sistema, en el mejor de los casos, producirá una escansión incorrecta y, en la peor, producirá un error. *Chociay* puede emplear su amplia experiencia para definir un esquema de clasificación de versos que sea adecuado a la obra de Gregório de Matos, por ejemplo separando los decasílabos en más de un tipo; *Aoidos*, para clasificar los versos, necesita a un humano que diga cuáles son los tipos y cómo identificarlos. El libro de *Chociay*, *Os Metros do Boca*, va, en efecto, mucho más allá del recuento de versos: el autor explora profundamente la obra del poeta y

proporciona una rica experiencia versificatoria al lector; el sistema produce solo números.

Por otro lado, los resultados que resultan de la utilización de *Aoidos* son plenamente compatibles con aquellos presentados por Chociay. Las interpretaciones y conclusiones que se hacen con base en los números de Chociay son las mismas que se hacen cuando se toman los datos del sistema. Si los resultados son equivalentes, una ventaja de la utilización del sistema es el ahorro de tiempo, ya que *Aoidos* genera sus resultados en cuestión de pocos minutos. Otra ventaja, bastante evidente en la comparación hecha en este artículo, es que el sistema no comete ciertos tipos de errores; en particular, *Aoidos* no salta versos. Si consideramos que el sistema produce escansiones equivocadas a menos del 1 % de los versos, concluimos que el número de versos saltados por Chociay es mayor que los escandidos incorrectamente por *Aoidos* (y enfatizar esto es importante).

A pesar de que la insistente comparación entre *Aoidos* y Chociay en este artículo oponga los ordenadores a los humanos, ello solo tiene sentido cuando se quiere demostrar que los ordenadores hacen un trabajo comparable al de los humanos. En la práctica, lo que se desea es que investigadores y estudiosos puedan utilizar *Aoidos* como una herramienta auxiliar; la intención ciertamente no es sustituir al humano. Al emplear un sistema automático de escaneo, el investigador evita la tarea ardua de contar las sílabas de miles de versos para clasificarlos y puede así ocupar su tiempo con actividades más nobles, como la interpretación crítica de la obra literaria o, si se ciernen específicamente a los problemas de escansión, podrá buscar soluciones que el sistema informático no tiene todavía capacidad de proponer por sí mismo.

6. Notas

1. Nosotros hemos traducido todas las citas. Cabe ya aclarar

que estamos utilizando el sistema grave para contar las sílabas de los versos, tal como se utiliza en Brasil y Portugal, y no lo sistema agudo español. (Volver)

2. El hecho de que la transcripción fonética del sistema muestre /vɛ.ʒʊ/ y no /ve.ʒʊ/ demuestra que el sistema está preocupado solamente por la escansión. La pronunciación correcta de "vejo" es importante, por ejemplo, para analizar rimas; sin embargo, para la escansión, si la vocal es abierta o cerrada influye en poco o nada. (Volver)

3. Elisión de la "m" final de una palabra delante de la inicial vocálica de la palabra siguiente. (Volver)

7. Bibliografía

Beaudouin, V. (2000). *Rythme et rime de l'alexandrin classique: Étude empirique des 80 000 vers du théâtre de Corneille et Racine*. Tesis (Doctorado), École des Hautes Études en Sciences Sociales.

Chociay, R. (1974). *Teoria do Verso*. São Paulo: McGraw-Hill do Brasil.

_____ (1993). *Os Metros do Boca: Teoria do Verso em Gregório de Matos*. São Paulo: UNESP.

_____ (1994). "Identidade Formal do Decassílabo em 'O Uruguai'" en *Revista de Letras*, v. 34, p. 229–243.

Delente, É.; Renault, R. (2015). "Traitement automatique des formes métriques des textes versifiés" en *22ème conférence sur le traitement automatique des langues naturelles*, pp. 432–438.

Gervás, P. A. (2000). "Logic Programming Application for the Analysis of Spanish Verse" en *1st international conference on computational*

logic, pp. 1330–1344.

Mamede, N.; I. Trancoso; P. Araújo, y C. Viana (2004). “An Electronic Assistant for Poetry Writing” en *9th ibero-american conference on artificial intelligence*, pp. 286–294.

Matos, G. de. (2014). *Gregório de Matos: Poemas Atribuídos, Códice Asensio-Cunha*. Hansen, J. A. y M. Moreira (eds.). Belo Horizonte: Autêntica Editora, vol. 3.

_____ (1992). *Obra Poética*. Amado, J. (ed.). 3ª ed. edição. Rio de Janeiro: Record.

_____ (2016). *Obras Poéticas de Gregório de Matos Guerra*. Florianópolis: Literatura Digital. Disponible en Web: <https://goo.gl/zwJ01s>. [Consulta: 15 de diciembre de 2016].

Mittmann, A. (2016). *Escansão Automático de Versos em Português*. Tesis (Doctorado), Universidade Federal de Santa Catarina.

_____ ; Wangenheim, A. von, & Luiz dos Santos, A. (2016). “Aoidos: A System for the Automatic Scansion of Poetry Written in Portuguese” en *17th international conference on intelligent text processing and computational linguistics*.

Navarro-Colorado, B.; M. R. Lafoz, y N. Sánchez (2016). “Metrical Annotation of a Large Corpus of Spanish Sonnets: Representation, Scansion and Evaluation” en *9th international conference on language resources and evaluation*.

Robey, D. (1993). “Scanning Dante’s the Divine Comedy. A Computer-based Approach” en *Literary and Linguistic Computing*, vol. 8, nº 2, pp. 81–84.

Topa, F. (1999). *Edição crítica da obra de Gregório de Matos*. Tesis (Doctorado), Universidade do Porto. Disponible en Web: <https://repositorio-aberto.up.pt/bitstream/10216/73576/2/11057.pdf>.

[Consulta: 5 de septiembre de 2017].