



TEXTO DIGITAL

Revista de Literatura, Linguística e Artes

Análise comparativa entre as escansões manual e automática dos versos de Gregório de Matos

A comparative analysis between automatic and manual scansion of Gregório de Matos' verses

Adiel Mittmann^a; Samanta Rosa Maia^b; Alckmar Luiz dos Santos^c

^a Universidade Federal de Santa Catarina, Santa Catarina, Brasil - adiel@mittmann.net.br

^b Universidade Federal de Santa Catarina, Santa Catarina, Brasil - samantamaia9@gmail.com

^c Universidade Federal de Santa Catarina, Santa Catarina, Brasil - alckmar@gmail.com

Palavras-chave:

Escansão Automática.
Poesia. Gregório de Matos.

Resumo: Este artigo compara resultados obtidos a partir de duas escansões do mesmo *corpus*, a obra poética completa de Gregório de Matos, uma delas feita manualmente e publicada por Rogério Chociay em seu livro *Os Metros do Boca*, e outra produzida automaticamente pela ferramenta Aoidos. A comparação mostra que existem algumas discrepâncias entre os dois tipos de escansão: o Aoidos encontrou um número significativo de versos a mais, o que evidencia que Chociay pulou alguns poemas; por outro lado, Chociay escandiu corretamente e com propriedade os versos de Gregório, enquanto o Aoidos cometeu alguns erros que um ser humano não cometeria. De toda maneira, os resultados produzidos pela abordagem automática, de um lado, e os advindos do trabalho manual, de outro, produziram resultados compatíveis, o que ressalta a utilidade da ferramenta automática, já que ela escande em poucos minutos o que leva meses para analisar através do processo de leitura habitual. Essa maior presteza nos resultados libera o pesquisador ou estudioso para dedicar seu tempo a atividades menos cansativas e mais nobres, isto é, a análises mais detalhadas, podendo agora dispor, de modo imediato, de grandes quantidades de dados com que trabalhar.

Keywords:

Automatic Scansion.
Poetry. Gregório de Matos.

Abstract: This article compares results obtained from two scansion of the same corpus, the complete poetic works of Gregório de Matos. One scansion was manually established and published by Rogério Chociay in his book *Os Metros do Boca*, while the other one was automatically produced by the tool Aoidos. The comparison shows that there are some differences between these two types of scansion: Aoidos found significantly more verses in the corpus, which is evidence that Chociay skipped some poems; on the other hand, Chociay scanned Matos' verses correctly and with propriety, whereas Aoidos made some mistakes that humans would not make. In any case, both the automatic approach as well as the manual work produced compatible results, which highlights the usefulness of the automatic tool, since it scans in a few minutes what usually take months for a human to analyze in the course of the regular reading process. This speed-up in the production of results allows researchers and scholars to focus their time on less tiring and more refined activities, since they now have quick access to a great amount of data.



1 INTRODUÇÃO

Escandir versos é uma atividade que nós, humanos, após algum treinamento e a aquisição de alguma experiência, conseguimos realizar de forma bastante eficiente: na maioria dos casos, uma olhadela sobre o poema já nos dá uma boa ideia do seu padrão estrófico e métrico; a leitura do primeiro verso confirma ou ajusta as suspeitas que tínhamos acerca do metro; chegando ao fim da primeira estrofe, já identificamos os padrões rítmicos preferidos do poeta e temos expectativas do que vem pela frente a partir dali. Quando chegamos ao fim do poema, contudo, certas perguntas são difíceis de serem respondidas: Quais variações rítmicas o poeta utilizou e com qual frequência? Quantas sinalefas e sinéreses o poeta empregou? Como comparar o poema ou o poeta a outros com base em dados objetivos? Se estivermos lendo um único soneto, podemos fazer algumas contagens e responder às duas primeiras perguntas; a terceira mesmo assim fica sem resposta. E se estivermos lendo *Os Lusíadas*, com seus 8.816 versos — como vamos fazer as contagens? E se quisermos analisar quantitativamente os mais de 30 mil versos de Gregório de Matos?

O Aoidos (MITTMANN; WANGENHEIM; LUIZ DOS SANTOS, 2016; MITTMANN, 2016) é uma ferramenta computacional capaz de escandir versos métricos em português. Seu objetivo não é gerar escansões melhores do que aquelas feitas manualmente por estudiosos; seu objetivo é gerar boas escansões (e coletar informações quantitativas enquanto o faz) muito mais rápido que os estudiosos. Se um humano conseguisse escandir cada um dos 4.184 versos das *Cartas Chilenas* em apenas um segundo, ainda assim estaria trabalhando 200 vezes mais lentamente que o Aoidos. A escansão de um grande número de versos e a anotação de dados quantitativos é fatigante para um humano, mas o computador não cansa. Ademais, conforme os computadores tornam-se mais rápidos, também a velocidade do Aoidos aumenta. Se excetuarmos, portanto, possíveis objetivos pedagógicos e educacionais, a utilidade de uma ferramenta como essa é escandir versos e gerar estatísticas quando nós, humanos, não teríamos tempo suficiente para fazê-lo.

Ferramentas similares foram propostas para versos escritos em outras línguas. Ainda na década de 90, Robey (1993) apresentou um sistema que auxilia na escansão dos versos da *Divina Comédia*. Para o espanhol, Gervás (2000) criou um sistema de escansão de versos que se mostrou capaz de acertar em 89% dos versos de um *corpus* de 64 sonetos. Para o mesmo idioma, Navarro-Colorado, Lafoz e Sánchez (2016) propuseram um sistema para escandir

sonetos, acertando em 96% dos versos. Para o francês, há o *Métromètre*, de Beaudouin (2000), que foi utilizado para escandir mais de 80 mil versos do teatro francês; e há também o *Anamètre*, desenvolvido por Delente e Renault (2015), que analisa versos em metros variados. Mamede, Trancoso, Araújo e Viana (2004) apresentaram um sistema de análise de versos para a língua portuguesa, mas com um foco diferente daquele do *Aoidos* e com um desempenho inferior: enquanto o *Aoidos* acerta em 99% dos versos, a taxa de acerto do sistema desses autores é de 82%.

Um dos poucos pesquisadores que se dispuseram a escandir e a fazer contagens manuais de grandes quantidades de versos é Rogério Chociay, autor do conhecido *Teoria do Verso* (CHOCIAY, 1974). Para caracterizar, por exemplo, *O Uruguai* de Basílio da Gama, Chociay contou o número de ocorrência de todos os esquemas rítmicos encontrados nos 1.377 decassílabos¹ da obra, agrupando-os ainda segundo os cantos em que aparecem e classificando-os em heroicos e sáficos (CHOCIAY, 1994). Para comparar, no mesmo artigo, esse poema de Basílio da Gama com obras de outros seis poetas, indo de Camões a Jorge de Lima, o autor de *Teoria do Verso* fez a mesma contagem de esquemas rítmicos para os 100 primeiros versos de uma obra escolhida de cada poeta, motivado pela possibilidade de comparar as obras dos poetas de forma objetiva. Assim, após citar alguns críticos de *O Uruguai*, ele indaga:

Não explicam, porém, os mencionados críticos, em que consiste, objetivamente, essa “fluidez”, essa “diferença” dos versos de *O Uruguai*: seria um modelo *sui generis* de decassílabo? ou apenas um modo diferente de construir o verso?

Abordaremos em mais detalhes, neste artigo, uma outra contagem realizada por Chociay, mais ambiciosa na quantidade de versos: trata-se dos números que o esticólogo publicou em *Os Metros do Boca* (CHOCIAY, 1993). Nesse livro, o autor explora a técnica versificatória de Gregório de Matos² e apresenta uma série de números a respeito dos mais de 30 mil versos compostos pelo poeta. Se o *Aoidos* puder ser usado para produzir análises semelhantes àquelas feitas manualmente por Chociay sobre os versos de Gregório, teremos bons indícios

¹ Este artigo sempre usa a contagem francesa de sílabas, ou seja, um decassílabo contém 10 sílabas poéticas porque a sua 10ª sílaba é a última acentuada no verso.

² Assim como Chociay em seu livro, também neste artigo nos referimos ao “Gregório de Matos” sem nos preocuparmos com questões de autoria, que não são o foco do artigo. Como disse Chociay (1993, p. 30), “Afinal, pode-se negar o poeta, sua originalidade e até mesmo o seu nome; nunca, porém, o poema”.

de que o mesmo pode ser feito para outros poetas. É justamente o que pretende este artigo: apresentar paralelamente os resultados de Chociay e do Aoidos e compará-los, mostrando que, embora não sejam idênticos, são equivalentes, ou seja, permitem praticamente as mesmas inferências e hipóteses de leitura. A diferença fundamental não reside nos resultados produzidos pelo humano e pelo computador, mas sim na velocidade: o Aoidos levou 127 segundos para analisar todos os versos, enquanto Chociay levou “muitos meses” para fazê-lo (CHOCIAY, 1993, p. 13).

O restante do artigo, por conseguinte, explica como o Aoidos gera suas propostas de escansão a partir do texto escrito; em seguida discorre acerca da preparação do *corpus* com os versos de Gregório de Matos; apresenta, então, e compara os números de Chociay e do Aoidos, destacando semelhanças e diferenças; por fim, o artigo conclui com uma mensagem de paz para robôs e humanos.

2 O SISTEMA AOIDOS

O Aoidos³ é um programa de computador capaz de escandir versos métricos em português. Enquanto nós, humanos, utilizamos nossos conhecimentos linguísticos e culturais para escandir versos, o Aoidos parte do verso escrito e, através de uma série de regras e procedimentos, todos codificados de forma explícita e mecânica, produz uma proposta de escansão. Tomemos, por exemplo, a primeira quadra de um soneto de Gregório 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 esta quadra, o Aoidos produz a seguinte escansão:

1	2	3	4	5	6	7	8	9	10	11
Tris-	te	Bah-	ia!	Oh	quão	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	já,	tu a	mi a-	bun-	dan-	te.

³ Uma versão *online* do Aoidos está disponível no endereço <<http://aoidos.ufsc.br/>>.

Nem todas as informações coletadas pelo sistema estão visíveis nessa escansão. O sistema sabe, entre outros, que o “a” final de “Bahia”, no primeiro verso, foi elidido, e que há uma sinalefa entre “tu” e “a” no terceiro verso. A seguir, tentamos explicar, de forma bastante geral, como o Aoidos chega a uma proposta de escansão, como essa, a partir do verso escrito.

A função primordial do Aoidos é a escansão de versos, e não a extração deles a partir de documentos em formatos como DOC (do Microsoft Word), PDF (*Portable Document Format*, tipicamente visualizados no Acrobat Reader) ou HTML (*Hypertext Markup Language*, visualizados em navegadores Web). Nós dispomos da habilidade de abrir um livro e, de forma rápida e eficaz, localizar visualmente onde estão os versos dos poemas, separar os versos que constituem o poema em si daqueles versos que estão na epígrafe abaixo do título e entender que “D. João” deve ser lido como “Dom João”. Para um programa como o Aoidos, entretanto, essas tarefas não são triviais, pois ele não é dotado do senso comum necessário para interpretar todas as variações de formatação (espaçamentos, tamanho da fonte, uso de itálicos, etc.) que podem aparecer em livros. Por este motivo, o Aoidos lê poemas apenas no formato TEI (*Text Encoding Initiative*). Neste tipo de arquivo os versos ficam explicitamente delimitados e agrupados, o que facilita a interpretação de seu conteúdo por ferramentas computacionais. Um arquivo no formato TEI pode ser criado e modificado em editores simples como o Notepad. A quadra anterior, por exemplo, fica assim no 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>
```

A utilização do formato TEI para codificar obras poéticas, além de permitir que ferramentas computacionais tenham acesso ao seu conteúdo, deixa aberta a possibilidade de que com base nos mesmos arquivos, sejam criadas versões para serem lidas, por exemplo, na Web.

O Aoidos toma os poemas de um arquivo TEI e produz uma transcrição fonética para cada verso. A transcrição fonética é feita com base em regras, de forma que dicionários não são necessários e neologismos são naturalmente processados. O objetivo desta transcrição fonética não é produzir uma representação exata de como os poemas devem ser lidos em voz alta; trata-se apenas de um artifício para representar os sons do português de forma mais

regular do que a ortografia permite e com vistas apenas à escansão. A transcrição é feita através de três etapas, todas realizadas sobre as palavras individuais de cada verso. Primeiramente a sílaba tônica (ou a sua ausência) é determinada, seja pela presença de diacríticos (como em “atônito”), pela consoante final que denuncia o acento (“mordaz”) ou pela análise dos grafemas finais da palavra (“rainha”). Em seguida o sistema faz uma divisão silábica, convertendo palavras como “abstração” em “abs-tra-ção” e “vácuo” em “vá-cuo”. Finalmente, um mapeamento leva cada grafema (ou grupo deles) a fonemas, levando “distante” a /dis'tã.ti/ e “mocinhas” a /mo'si.ɲas/. O terceiro verso da quadra de exemplo fica, portanto, assim⁴:

'pɔ.brɪ tɪ 'vɛ.ʒo a 'ti 'tu a 'mi ẽ.pe'ɲa.du

Essa transcrição fonética, feita palavra a palavra e de forma 100% determinística (nela “negociante” nunca tem 4 sílabas, sempre 5: “ne-go-ci-an-te”), serve de base para a derivação de alternativas de pronúncia para o verso como um todo. A ideia é gerar o maior número possível de diferentes maneiras de pronunciar o mesmo verso levando-se em consideração os metaplasmos disponíveis na língua e explorados pelo poeta. O verso mencionado, por exemplo, se lido de acordo com a transcrição fonética inicial, contaria 13 sílabas poéticas. A partir dela, o sistema encontra a possibilidade de sinalefa entre “vejo” e “a” e produz uma alternativa de pronúncia para o verso que incorpora essa sinalefa. Também entre “tu” e “a” o sistema percebe que uma sinalefa pode ser feita, e registra uma terceira alternativa. Uma quarta alternativa é realizar ambas as sinalefas. Para esse verso o Aoidos encontrou um total de 14 alternativas. Nem todas, contudo, são igualmente plausíveis. O ouvido moderno diz que a sinalefa entre “vejo” e “a” é bastante desejável, mas aquela entre “tu” e “a” é praticamente forçada. O Aoidos, ao produzir as diferentes alternativas, mantém o registro de quais metaplasmos foram empregados e o índice de “naturalidade”, evidentemente bastante subjetivo, de cada alternativa de pronúncia do verso. No caso do exemplo, a alternativa mais “natural” encontrada pelo sistema foi aquela em que o único metaplasmo empregado é a sinalefa entre “vejo” e “a”, ficando assim a transcrição fonética correspondente:

'pɔ.brɪ tɪ 'vɛ.ʒwa 'ti 'tu a 'mi ẽ.pe'ɲa.du

⁴ O fato de a transcrição fonética do sistema mostrar /'vɛ.ʒo/ e não /'vɛ.ʒo/ é testemunho de que o sistema está preocupado somente com escansão. A pronúncia correta de “vejo” é importante, por exemplo, para analisar rimas; mas, para a escansão, a vogal ser aberta ou fechada influencia em pouco ou nada a escansão.

Evidentemente que esta “naturalidade” não leva em consideração o contexto em que o verso aparece; o sistema ainda não sabe que o soneto está escrito em decassílabos e que essa alternativa de 12 sílabas claramente não é a correta. Tampouco o sistema tem a sofisticação de perceber que está escrito “mi” e não “mim”, o que provavelmente é uma indicação ao leitor, por parte do poeta ou do editor, de que a sinalefa entre “mi” e “empenhado” deve ser feita. O Aoidos possui regras para vários metaplasmos, incluindo sinalefas, sinéreses, síncope, epênteses, sístoles, etc.

De posse de todas as alternativas de pronúncia de todos os versos, o Aoidos olha para o poema como um todo para determinar a quantidade de metros utilizados. Para fazê-lo, o sistema considera quais são as alternativas de pronúncia mais “naturais” de cada verso, bem como o tamanho silábico mínimo e máximo do poema e das estrofes. Considerando o soneto de onde a quadra de exemplo foi tirada, o único tamanho silábico compatível com todos os versos é 10, ou seja, todos os versos podem ser escandidos em 10 sílabas, mas há versos que não podem ser escandidos em 9 ou 11. Caso houvesse mais de um tamanho silábico compatível com todos os versos do poema, o sistema consideraria a “naturalidade” conjunta dos versos escandidos em diferentes números de sílabas. Ao todo, o Aoidos tenta enquadrar os poemas em uma de quatro categorias: **a)** todos os versos do poema têm o mesmo comprimento, ou seja, o poema é isométrico; **b)** os versos do poema possuem exatamente dois comprimentos distintos, o que seria o caso de versos heroicos (10 sílabas) e seus quebrados (tipicamente 6 sílabas); **c)** as estrofes do poema, quando consideradas isoladamente, contêm um ou dois metros; **d)** o poema (e suas estrofes) contém mais de duas medidas.

Quando o sistema já decidiu quantos metros estão presentes no poema, ele chega à sua fase final: encontrar, entre as alternativas de pronúncia de cada verso, aquela que melhor se encaixa no contexto rítmico-acental em que ele aparece. Selecionar a alternativa pretensamente (ou possivelmente) mais natural entre aquelas compatíveis com o metro do poema pode não ser suficiente, já que diferentes alternativas de mesmo tamanho podem apresentar diferentes ritmos. Para o segundo verso do primeiro terceto do soneto de exemplo, o sistema encontrou duas escansões decassilábicas:

1	2	3	4	5	6	7	8	9	10	11
Pe-	las	dro-	ga-	s i-	nú-	teis,	que a-	be-	lhu-	da
Pelas	dro-	ga-	s i-	nú-	teis,	que	a-	be-	lhu-	da

Na primeira alternativa, o sistema fez a sinalefa entre “que” e “a”, ao passo que a segunda manteve ali o hiato e fez uma síncope em “pelas”, como se a palavra estivesse escrita “p’las” na edição. O esquema rítmico das duas alternativas é diferente: a primeira é um típico heroico com esquema 3-6-10, enquanto a segunda possui um questionável esquema 2-5-10. Diante dos esquemas rítmicos encontrados, o sistema pode considerar o contexto do poema em que o verso aparece; percebendo que quase todos os versos podem ser “naturalmente” escandidos com a 6ª tônica, o sistema escolhe então a alternativa 3-6-10 do verso de exemplo. Opcionalmente, o sistema pode empregar um catálogo de ritmos em que o usuário expressa, de maneira estatística, quais são os padrões rítmicos mais aceitáveis na tradição poética em que o poema se situa. Ao não usar um catálogo de ritmos, o usuário permite que o sistema analise cada poema de forma bastante individual, sem noções preconcebidas de quais ritmos um verso deve ou não possuir; por outro lado, o emprego de um catálogo possibilita que sejam escandidos poemas pequenos, que tenham número insuficiente de versos para que o sistema entenda quais são os ritmos preferidos pelo poeta. Em poucas palavras, o sistema pode sozinho concluir que Camões prefere os esquemas 6-10 e 4-8-10 em *Os Lusíadas*, mas necessitará auxílio do usuário para escandir uma quadra qualquer tomada de forma isolada.

A acurácia do Aoidos foi avaliada num experimento em que 1.080 versos foram escandidos pelo sistema e manualmente conferidos. Esses versos foram selecionados de forma aleatória de um *corpus* maior, correspondendo a 1% do total de 107.438 versos. O *corpus* completo continha obras poéticas de 10 autores, variando temporalmente de Camões a Delminda Silveira, passando por Gregório de Matos, Gonçalves de Magalhães e Augusto dos Anjos. A taxa de acerto global do sistema foi de 99,0%, variando entre 97,5 e 100,0% conforme a obra poética. De Gregório de Matos, foram conferidos manualmente 358 versos, 99,2% dos quais estavam corretamente escandidos.

3 PREPARAÇÃO E ESCANSÃO DA OBRA POÉTICA DE GREGÓRIO DE MATOS

Para fazer uma comparação entre as análises produzidas por Chociay e pelo Aoidos, é preciso que o humano e o robô trabalhem sobre o mesmo *corpus*. Chociay trabalhou sobre a 2ª edição da obra completa de Gregório publicada por James Amado (CHOCIAY, 1993, p. 52), mas a versão digital de que dispúnhamos fora criada a partir da 3ª edição de James Amado. Apesar de haver grandes diferenças entre a 1ª edição de James Amado, de 1969, e a 2ª, de 1990, como o próprio Chociay destaca (CHOCIAY, 1993, pp. 25-29), a 2ª e a 3ª edição, esta

última de 1992, são similares ou idênticas. Na 3ª edição não há nenhuma menção quanto a diferenças com relação à edição anterior; com efeito, os dois volumes contêm menções somente à 2ª edição — exceto na capa e na ficha catalográfica. Enquanto 21 anos se passaram entre a publicação das duas primeiras edições, a terceira surgiu apenas dois anos após a segunda, o que evidencia que nenhuma grande revisão aconteceu. Ademais, os versos cuja localização nos volumes e nas páginas da 2ª edição é dada por Chociay aparecem exatamente nos mesmos volumes e páginas da 3ª. Acreditamos, portanto, que pouco ou nada se perde quando utilizamos a 3ª edição e Chociay, a 2ª. Quando nos referirmos, doravante, à “edição de James Amado”, estamos fazendo-o de forma genérica à 2ª ou 3ª.

O processo de estabelecimento do *corpus* em formato TEI iniciou com a adaptação da edição acessível na Literatura Digital (MATOS, 2017), que foi criada através da digitalização da 3ª edição em papel de James Amado (MATOS, 1992). Os 40 arquivos HTML que compõem esta edição digital foram convertidos em TEI de forma semiautomática: programas de computador, desenvolvidos especificamente para esta tarefa, foram comandados por um humano para gradualmente converter cada um dos arquivos HTML em TEI. De forma geral, cada linha encontrada no corpo dos poemas foi vista como um verso e codificada com o elemento <l> do TEI. Estrofes, codificadas com o elemento <l g>, foram identificadas através de espaços verticais em branco. Alguns casos especiais, contudo, foram considerados. Por exemplo, vários poemas (na edição digital e impressa) contêm pontilhados que coordenam dois versos numa mesma linha; no nosso *corpus*, tanto o verso à esquerda quanto à direita do pontilhado ganhou seu próprio elemento <l> — à exceção de um soneto, com *incipit*⁵ “Na oração, que desaterra aterra”, em que as rimas internas haviam sido destacadas com um pontilhado, que foi então ignorado.

Já codificado no formato *TEI*, o *corpus* passou por uma revisão inicial. A quantidade de versos nas estrofes e a quantidade de estrofes nos poemas foi verificada; assim, por exemplo, percebeu-se que o poema “Pois no que toca a guardar” continha um número suspeito de versos: 71. O cotejamento entre a edição digital e a impressa permitiu constatar que um verso desapareceu no processo de digitalização. O poema, com o verso restabelecido, passou a contar com 72 versos (18 quadras). No total, apenas 4 versos foram inseridos na edição digital porque constavam da impressa; vários casos suspeitos, todavia, não puderam ser

⁵ Usaremos sempre o *incipit*, mais compacto que os títulos, para nos referirmos aos poemas de Gregório.

resolvidos, já que os versos faltam na própria edição impressa. Nesta revisão, através de palavras-chave, foram também localizados os poemas que Gregório de Matos escreveu em castelhano; tais poemas foram marcados no *corpus* para que o Aoidos não os tentasse escandir.

Após a revisão inicial, os poemas de Gregório de Matos foram submetidos à análise do Aoidos. O sistema apontou, então, uma série de versos que não pôde escandir. Ele foram, então, examinados, um a um, para que erros na edição digital (ou seja, discrepâncias com a edição impressa) fossem corrigidos e para que abreviações e estrangeirismos fossem devidamente marcados no *corpus*. Nessa etapa foram encontrados 240 versos com erros na edição digital. Muitos destes eram resquícios da etapa de OCR (*Optical Character Recognition*) da digitalização: assim, 130 versos estavam incorretos porque diziam “Ihe” (com a letra “i” maiúscula no início) em vez de “Ihe” (com a letra “l” minúscula). Houve também casos de palavras que perderam o diacrítico na digitalização, como “roido” em vez de “roído”. As abreviações, aqui consideradas de forma ampla como quaisquer fragmentos textuais não escritos por extenso, são tipicamente casos de “D.” (“Dom”) e “Fr.” (“Frei”), mas há também nomes de letras, como “ff” (“efes”). Os estrangeirismos encontrados são expressões em castelhano (como no verso “estos me lleban dizia”⁶) ou em latim (“e dizendo o jube domine”). É importante salientar que tanto os erros da edição digital quanto as marcações de abreviações e estrangeirismos só foram detectados quando o sistema não foi capaz de escandir um verso. Assim, percebeu-se que faltava um diacrítico em “roído” porque o sistema não conseguiu escandir o verso “Saí desta canalha tão roido” no contexto decassilábico em que aparece; a abreviação “Fr.” não pode ser lida pelo sistema como uma palavra, o que é sinalizado pelo sistema durante sua operação; e a palavra “domine”, se lida como se fosse português, é paroxítona, mas em latim ela é paroxítona, leitura esta que é necessária para escandir o verso em sete sílabas, e não oito. Por fim, restaram 45 versos que, mesmo sendo manualmente cotejados com a edição impressa e cuidadosamente examinados, desafiam a escansão. O verso seguinte, por exemplo, aparece num contexto que exige sete sílabas:

Terra, que não parece

⁶ A edição de James Amado não atualizou a ortografia dos fragmentos em castelhano: “Nos poemas em língua espanhola mantivemos o modelo dos códices.” (1992, p. 9).

Na recente edição do Códice Asensio-Cunha (MATOS, 2014, p. 193), esse mesmo verso vem assim:

Terra, que não se parece

Tratar-se-ia, portanto, de um “erro” presente na própria edição de James Amado. Não foi feita nenhuma tentativa de se estabelecer as versões “corretas” destes 45 versos; foram apenas marcados como inescandíveis no *corpus*.

Alguns versos foram excluídos do *corpus* ou ignorados nas escansões e contagens de que trata este artigo. Primeiramente, os dois poemas do chamado “coronista resuçitado” foram completamente descartados, por terem sido escritos por Tomás Pinto Brandão (MATOS, 1992, p. 1245). Ainda as epígrafes e os motes, apesar de permanecerem no *corpus*, foram desconsiderados para propósitos de escansão, e não são incluídos em nenhum tipo de contagem.

Cabe, por fim, observar que Chociay considerou em seu trabalho também a edição da Academia Brasileira de Letras da obra poética do Boca do Inferno: “Embora tenhamos feito a contagem com base na segunda edição de James Amado, incluímos versos que não constam nessa edição (foram descobertos no cotejo com a edição da Academia)”. Chociay afirma que a edição da Academia, “apesar de ser a primeira que se pretendeu ‘completa’ e do inestimável serviço que prestou para o conhecimento da poesia gregoriana, filia-se a uma tradição de falso moralismo” (CHOCIAY, 1993, p. 24) e que “deve ser consultada com reservas e cautela” (CHOCIAY, 1993, p. 25). Mesmo com suas reservas com relação à edição da Academia, Chociay a utilizou porque a de James Amado oferece “soluções ingênuas” (CHOCIAY, 1993, p.30) em diversos casos.

4 AOIDOS VS. CHOCIAY

Rogério Chociay, em seu livro *Os Metros do Boca* (CHOCIAY, 1993), discute a obra poética de Gregório de Matos sob vários ângulos: o autor avalia as edições da poesia completa de Gregório à época, tece comentários acerca de sua técnica versificatória e, de particular interesse ao presente artigo, quantifica os tipos de versos que o poeta utilizou. À p. 52, ele apresenta uma tabela sumarizando os tipos de versos e as frequências absolutas e relativas

deles na obra do Boca, que é aqui integralmente reproduzida na Tabela 1. Sobre os números da tabela, Chociay comenta ainda que, por incluir versos que não constam na edição de James Amado, “e ainda aceitando uma pequena margem de erro, os números acima, 33.123, não têm pretensão de exatidão absoluta.”

Tabela 1. Tipos de versos e sua frequência na obra poética de Gregório de Matos, conforme originalmente publicado por Chociay (1993).

Versos	Número de vezes	%
Dissílabo	15	0,04
Redondilho quebrado	251	0,76
Quebrado de <i>ovillejo</i>	66	0,20
Tetrassílabo	95	0,29
Pentassílabo (redondilho menor)	317	0,96
Hexassílabo	168	0,51
Heroico quebrado	123	0,37
Verso de seguidilha	52	0,16
Heptassílabo (redondilho maior)	28.028	84,61
Eneassílabo	96	0,30
Decassílabo italiano	3.798	11,46
Decassílabo 5-10	16	0,05
Decassílabo 4-7-10	10	0,03
Hendecassílabo	88	0,26
Totais gerais	33.123	100

Não são todos os tipos de verso apresentados na Tabela 1 que correspondem de forma biunívoca a tamanhos silábicos. Assim, um verso cuja escansão acuse 6 sílabas pode, por exemplo, constar da tabela sob o tipo “hexassílabo”, “heroico quebrado” ou “verso de seguidilha”; os versos agrupados sob o rótulo “redondilho quebrado”, por outro lado, podem ser escandidos em 3 ou 4 sílabas. Uma consequência dessa não correspondência direta entre tipos e tamanhos é a impossibilidade de produzirmos uma tabela semelhante a partir da escansão automática. A partir dos resultados da ferramenta, é fácil calcular quantos versos foram escandidos com 6 sílabas, mas não se pode saber imediatamente quais deles são hexassílabos e quais são heroicos quebrados sem considerar o contexto em que eles aparecem. O Aoidos, portanto, foi estendido com algumas heurísticas, ou seja, algumas regras de uso específico, que olham para além do verso, na tentativa de imitar o esquema classificatório adotado por Chociay. As heurísticas são quatro: uma para os redondilhos quebrados, outra

para os quebrados de *ovillejo*, uma terceira para os hexassílabos, heroicos quebrados e versos de seguidilha, e uma última para os decassílabos.

Um redondilho quebrado é “tradicionalmente o verso de três sílabas, sempre que usado em associação com o redondilho maior em diferentes tipos de estrofe ou sistema estróficos.” (CHOCIAY, 1993, p. 53). Apesar das 3 sílabas tradicionais, a escansão destes versos quebrados pode resultar, quando considerados em isolamento, em 3 ou 4 sílabas poéticas. Chociay argumenta que os fenômenos de sinafia e compensação (CHOCIAY, 1993, pp. 45-47) explicam boa parte dos tetrassílabos aparentes. A sinafia é a sinalefa entre a sílaba átona final de um verso e a primeira do seguinte; a compensação é a absorção da átona inicial de um verso pelo precedente, que neste caso é obrigatoriamente agudo. As duas estrofes seguintes contêm um caso de cada fenômeno, conforme exemplifica Chociay (1993, p. 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?

Na primeira estrofe, ocorre uma sinalefa entre a sílaba final de “cabeleira” e a inicial de “encrespadinha”, de maneira que para o quarto verso sobra apenas o fragmento “-crespadinha”, com 3 sílabas. Na segunda estrofe, a palavra oxítone “cus” do terceiro verso permite que o artigo definido “o” do quarto verso ocupe a posição final do terceiro, deixando um total de 3 sílabas para o quarto verso. Casos há, contudo, em que nenhum fenômeno é capaz de reduzir a 3 a quantidade de sílabas poéticas do último verso da estrofe; Chociay chama a sílaba excedente de “perdida”. Um exemplo é a seguinte estrofe, em que o quarto verso não pode ser escandido em menos de 4 sílabas:

Terra tão grosseira, e crassa,
que a ninguém se tem respeito,
salvo quem mostra algum jeito
de ser Mulato.

O Aoidos não leva em conta, atualmente, processos de acomodação que não respeitem o limite do verso, de forma que ele não é capaz de identificar casos de sináfia ou compensação. A heurística que tenta detectar os redondilhos quebrados, portanto, adota uma estratégia mais simples: se uma estrofe tem no mínimo dois versos heptassilábicos, então todos os versos com 3 ou 4 sílabas daquela estrofe são considerados redondilhos quebrados. Versos de 3 ou 4 sílabas que não se encontram em tais estrofes não são afetados pela heurística, tampouco o são os heptassilábicos.

A segunda heurística tem a função de identificar os quebrados de *ovillejo*. Chociay ensina (ênfases dele):

Segundo tratadistas espanhóis, o *ovillejo* é um sistema que apresenta uma primeira semiestrofe formada por três heptassílabos *interrogativos*, a cada um dos quais segue *uma palavra-resposta* que repete a rima do respectivo verso; a segunda semiestrofe consiste em uma redondilha *abba* que resume o sentido dos versos anteriores e cujo último verso *recolhe* as três palavras-resposta de tais versos. (CHOCIAY, 1993, p. 99)

Um *ovillejo* bastante conhecido de Gregório — e que segue precisamente a receita dada por Chociay — é o seguinte:

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.

A heurística para detectar a presença dos quebrados de *ovillejo* funciona, portanto, do seguinte modo. Para que um poema seja composto de *ovillejos*, é necessário que: **a)** o poema contenha um número par de estrofes (já que cada *ovillejo* possui duas); **b)** as estrofes ímpares contenham seis versos (que correspondem às três perguntas e às três respostas); e **c)** as estrofes pares contenham quatro versos (três resumindo o sentido dos anteriores e um coletando as palavras-resposta). Ora, os versos das palavras-resposta são justamente os quebrados de *ovillejo* que a heurística quer identificar, de maneira que ela considera que os

versos pares das estrofes ímpares de tais poemas são quebrados de *ovillejo*, independentemente do número de sílabas que eles possam conter. Os demais versos dos *ovillejos* não recebem nenhum tratamento especial.

A terceira heurística enquadra todos os versos que o Aoidos escandiu em 6 sílabas e alguns que o sistema julgou ter 4 sílabas; tais versos podem figurar na tabela de Chociay como hexassílabos, heroicos quebrados ou versos de seguidilha. A categoria dos hexassílabos engloba os versos de 6 sílabas que aparecem em grupos — tipicamente em estrofes isossilábicas; os heroicos quebrados são os versos de 6 sílabas que aparecem junto dos decassílabos. Já os versos de seguidilha podem ter 4 ou 6 sílabas; Chociay diz que o modelo de seguidilha adotado por Gregório é aquele de “uma estrofe de quatro versos, sendo os ímpares *hexassílabos* e os pares *tetrassílabos*, sem rima nos versos ímpares e com rima toante nos pares” (CHOCIAY, 1993, p. 73). A heurística, portanto, trabalha do seguinte modo: se um verso de 6 sílabas encontra-se em uma estrofe em que no mínimo 75% dos versos contém 6 sílabas, então este verso é classificado como hexassílabo; um verso de 4 ou 6 sílabas em uma estrofe que contenha pelo menos 2 versos de quatro sílabas e 2 de seis sílabas é classificado como verso de seguidilha; qualquer verso que não seja classificado nem como hexassílabo, nem como de seguidilha é então classificado como quebrado de heroico. Versos de 4 sílabas que não tenham sido classificados como de seguidilha não recebem mais nenhum tratamento por parte desta heurística; versos de 4 sílabas que já foram classificados como redondilhos quebrados ou quebrados de *ovillejos* não sobreviveram até esta heurística.

A última heurística visa a examinar os decassílabos e a encaixá-los em uma das três categorias definidas por Chociay: a dos decassílabos italianos, a dos versos com esquema 5-10 e a dos versos com esquema 4-7-10. Na nomenclatura de Chociay, os italianos são os decassílabos da nova medida, os heroicos e os sáficos; os outros, mais excepcionais na obra de Gregório de Matos, são representantes de uma tradição ibérica mais antiga e não aparecem no mesmo contexto que os italianos. A heurística trabalha, portanto, do seguinte modo: se um decassílabo é acentuado na 6ª sílaba ou então simultaneamente na 4ª e na 8ª, ele é italiano; se a 5ª sílaba é acentuada, ele tem esquema 5-10; se a 4ª e a 7ª são acentuadas, ele tem esquema 4-7-10; se ele não se encaixa em nenhum desses casos, ele é considerado italiano.

O Aoidos, estendido com essas quatro heurísticas, produziu os resultados exibidos na Tabela 2. A diferença total, de 1.272 versos, é bastante significativa e merece ser examinada.

Tabela 2. Comparação dos resultados obtidos pelo Aoidos e por Chociay. Os percentuais foram computados dividindo-se a diferença pelo valor encontrado pelo Aoidos.

Categoria	Aoidos	Chociay	Diferença	%
Dissílabo	15	15	0	0,0
Trissílabo	3		3	100,0
Redondilho quebrado	256	251	5	2,0
Quebrado de <i>ovillejo</i>	75	66	9	12,0
Tetrassílabo	155	95	60	38,7
Pentassílabo (redondilho menor)	319	317	2	0,6
Hexassílabo	168	168	0	0,0
Heroico quebrado	194	123	71	36,6
Verso de seguidilha	48	52	-4	-8,3
Heptassílabo (redondilho maior)	28.747	28.028	719	2,5
Octossílabo	21		21	100,0
Eneassílabo	107	96	11	10,3
Decassílabo italiano	4.141	3.798	343	8,3
Decassílabo 5-10	30	16	14	46,7
Decassílabo 4-7-10	14	10	4	28,6
Hendecassílabo	100	88	12	12,0
Dodecassílabo	1		1	100,0
Erro	1		1	100,0
	34.395	33.123	1.272	3,7

Começamos com os poemas em castelhano: o sistema certamente os incluiu na contagem, mas não está claro se Chociay o fez; além disso, os próprios números dos versos em castelhano não coincidem. Ele afirma que Gregório escreveu 21 poemas em castelhano, totalizando 584 versos, dos quais 472 redondilhos e 112 decassílabos (CHOCIAY, 1993, p. 51). Com os poemas codificados em formato TEI, obtivemos o mesmo número de poemas, 21, porém um número diferente de versos em castelhano, 555, ou seja, 29 a menos que Chociay. É difícil especular sobre as causas exatas desta divergência numérica, mas podemos demonstrar que os números de Chociay não estão corretos. Ele afirma que são 112 os decassílabos escritos em castelhano; ora, a Tabela 3 mostra as páginas, na edição de James Amado, em que aparecem sonetos decassilábicos em castelhano. São nove sonetos, que totalizam, portanto, $9 \times 14 = 126$ decassílabos, 14 a mais que os encontrados por Chociay. Se ajustarmos a contagem dele para levar em consideração este número de decassílabos, o total de versos em castelhano saltaria de 584 para 598. As contagens de versos feitas pelo Aoidos não levam em consideração os versos dos motes, já que eles se repetem ao longo dos poemas; se, por um instante, deixarmos de lado essa restrição e contarmos todos os versos em

castelhano, incluindo os versos dos motes, o Aoidos acusa 597 versos, o que é muito próximo da quantidade ajustada de 598. Existe, todavia, um poema em castelhano, com *incipit* “Coração: siente tu anhelo”, com estrutura similar a vários outros escritos por Gregório: um mote seguido de quatro décimas. Na terceira décima deste poema, porém, falta um verso, o que se pode constatar às pp. 422-423 da edição de James Amado. Chociay possivelmente consultou a edição da Academia Brasileira de Letras e por este motivo teve acesso ao verso que não consta da edição de James Amado; esse verso, se presente no nosso *corpus*, elevaria a nossa contagem para os 598 versos (ajustados) de Chociay. Se este raciocínio estiver correto, Chociay pulou um soneto em castelhano quando contou versos mas não o pulou quando contou poemas, já que sem nenhum ajuste o Aoidos encontrou os mesmos 21 poemas que Chociay.

Tabela 3. Números das páginas em que os sonetos decassilábicos em castelhano escritos por Gregório de Matos aparecem na edição de James Amado.

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

Não está claro, contudo, como já dito, se os poemas em castelhano foram incorporados na contagem de Chociay junto com os versos em português. Por um lado, os poemas em castelhano são parte indissociável da obra poética de Gregório do Matos e nos parece fazer bastante sentido contá-los; por outro, Chociay afirma:

Para se ter uma ideia dessa pouca importância do castelhano, basta lembrar que são mais de vinte e oito mil os redondilhos escritos por Gregório em português. (CHOCIAY, 1993, p. 51)

Ora, se a tabela produzida por Chociay indica que há 28.028 redondilhos e se ele afirma que há mais de 28.000 redondilhos em português, restaria margem para apenas 27 redondilhos em castelhano; ele mesmo diz, contudo, que há 472 redondilhos naquele idioma. A conclusão lógica, embora implícita, é que a tabela não inclui versos em castelhano. Chociay, contudo, certamente contou alguns versos em castelhano: aqueles que aparecem nos poemas escritos majoritariamente em português. Prova disto é que Chociay conta 88 hendecassílabos, oriundos dos poemas “Sendo pois o alterar da moeda” e “Marinículas todos os dias” (CHOCIAY, 1993, p. 65). Os dois poemas somam $43 + 1 = 44$ quadras que alternam versos de 9 e 11 sílabas, de forma que há $44 \times 2 = 88$ hendecassílabos. Entretanto, o segundo poema mencionado contém estes dois versos (não consecutivos) escritos inteiramente em castelhano:

Los ordeñadores se alquilan aqui
Pudiera cogerlos en un mal Latim

Chociay inclui, portanto, estes dois versos em castelhano na contagem dos hendecassílabos, e provavelmente fez o mesmo com muitos outros versos naquele idioma que aparecem em meio a versos em português.

Poderíamos diminuir a diferença entre os totais de versos de 1.272 para 717 se excluíssemos da contagem do Aoidos os 555 versos em castelhano, seguindo a lógica implícita de que Chociay não os contou. Se considerássemos, contudo, toda a lógica implícita de Chociay, deveríamos também incluir os versos dos motes, como Chociay aparentemente fez nos poemas em castelhano, o que adicionaria 307 à nossa contagem, parcialmente compensando a subtração de 555 e mantendo a diferença acima de 1.000 versos.

Mesmo ignorando a questão dos versos em castelhano, sabe-se que Chociay não contou certos versos. Ele afirma que Gregório de Matos compôs 196 sonetos, mas uma análise automática produz a contagem de 217 sonetos. Uma conferência manual foi feita para averiguar que o número de sonetos é, no mínimo, 217; as páginas em que cada um deles aparece constam da Tabela 4. Mesmo se descontarmos os 9 sonetos em castelhano, a diferença ainda existe: são $217 - 9 - 196 = 12$ sonetos que Chociay não contou, ou seja, 168 versos.

Tabela 4. Números das páginas em que os sonetos decassilábicos, em português e em castelhano, aparecem na edição de James Amado.

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			

O Aoidos encontrou 75 quebrados de *ovillejo*, enquanto Chociay aponta a existência de apenas 66 na obra de Gregório de Matos. Entretanto, Chociay afirma também que quatro foram os poemas escritos em *ovillejos* (CHOCIAY, 1993, pp. 99-103) e cita as páginas da edição de James Amado em que os tais poemas aparecem: 55-56, 56-58, 830 e 1051-1052. Ao consultar a edição impressa e fazer uma contagem manual dos quebrados de *ovillejos* destes quatro poemas, percebemos que eles contêm 18, 27, 9 e 21 destes versos quebrados, respectivamente. A conta correta, portanto, é a do sistema: $18 + 27 + 9 + 21 = 75$. Os 9 versos a mais encontrados pelo sistema correspondem justamente ao número de quebrados do poema da p. 830, o que é evidência de que este poema foi pulado na contagem de Chociay.

A contagem automática do Aoidos revelou a presença de 155 tetrassílabos, mas Chociay contou apenas 95. Dois poemas bastam para que a soma de 95 versos tetrassílabos seja superada: um com *incipit* “Vá de retrato” (pp. 183-186 de James Amado) e outro, “Vá de

aparelho” (pp. 846-849). O primeiro possui 84 tetrassílabos e o segundo, 57, perfazendo 141 versos deste tipo. O próprio Chociay menciona os dois poemas à p. 76 de seu livro, inclusive com referência às páginas da edição de James Amado. A evidência, portanto, de que Chociay pulou no mínimo $141 - 95 = 46$ tetrassílabos é bastante clara. Descobrir exatamente quais versos foram contados por Chociay para chegar à soma de 95 e a origem da diferença entre a contagem dele e a do Aoidos é mais difícil, mas claro está que a quantidade de tetrassílabos é superior àquela apontada por Chociay.

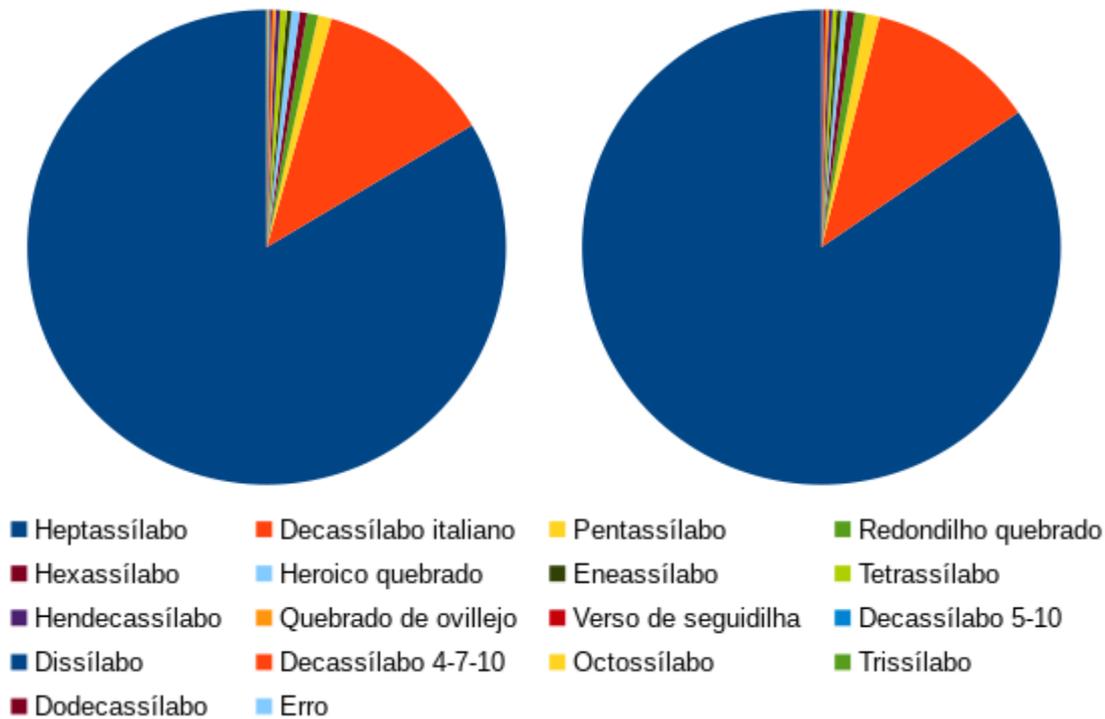
O único tipo de verso em que a balança se volta para o lado de Chociay é o verso de seguidilha; o Aoidos contou 48, e Chociay, 52. À p. 58 de seu livro, Chociay afirma que Gregório de Matos utilizou este verso em dois poemas; o primeiro deles, “Retratar ao bizarro”, foi corretamente contabilizado pelo sistema, resultando nos 48 versos apontados. O sistema, entretanto, não conseguiu identificar a seguinte estrofe do segundo poema indicado por Chociay como uma quadra de versos de seguidilha:

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

O Aoidos não foi capaz de fazê-lo porque não conseguiu escandir o segundo verso da quadra em quatro sílabas. Não está claro como Chociay escandiu o referido verso. O sistema, ao não classificar estes versos como de seguidilha, classificou-os em outras categorias.

Entre todas as discrepâncias mencionadas, queremos ressaltar a concordância que emerge quando os números são considerados na sua globalidade. Considerem-se, por exemplo, os dois gráficos exibidos na Figura 1; o gráfico à esquerda representa os números do Aoidos e aquele à direita, os de Chociay. As proporções dos tipos de versos são extremamente similares nos dois gráficos. Em ambos os gráficos, salta aos olhos a predominância maciça dos heptassílabos; em ambos, os decassílabos italianos participam como coadjuvantes; em ambos, os demais tipos de versos são minoria absoluta. Quando nos distanciamos um pouco dos detalhes, o que vemos é que tanto os versos pulados de Chociay (algo entre 700 e 1300) quanto os erros de escansão feitos pelo sistema (aproximadamente 288 versos) não nos impedem de caracterizar o poeta.

Figura 1. Proporção dos vários tipos de versos, de acordo com o Aoidos (à esquerda) e Chociay (à direita).



5 CONCLUSÃO

Os resultados quantitativos produzidos pelo Aoidos são similares, mas não idênticos, àqueles relatados por Chociay. Por um lado, Chociay certamente possui uma sofisticação muito superior à do sistema. Ao encontrar um verso em que parece faltar uma sílaba, Chociay pode consultar edições alternativas para encontrar uma versão mais coerente e a sílaba que faltava; o sistema, na melhor das hipóteses, produzirá uma escansão incorreta e, na pior, produzirá um erro. Chociay pode empregar sua larga experiência para definir um esquema de classificação de versos que seja adequado à obra de Gregório de Matos, por exemplo separando os decassílabos em mais de um tipo; o Aoidos, para classificar os versos, precisa de um humano que diga quais são os tipos e como identificá-los. O livro de Chociay, *Os Metros do Boca*, vai, com efeito, muito além da contagem de versos: o autor explora profundamente a obra de Gregório de Matos e fornece uma rica experiência versificatória ao leitor; o sistema produz apenas números.

Por outro lado, os resultados produzidos pelo Aoidos são plenamente compatíveis com aqueles apresentados por Chociay. As interpretações e conclusões que se fazem com base nos números de Chociay são as mesmas que se fazem quando são tomados os números do sistema. Se os resultados são equivalentes, uma vantagem da utilização do sistema é a economia de tempo e de paciência, já que o Aoidos gera seus resultados em questão de

poucos minutos. Outra vantagem, bastante evidente na comparação feita neste artigo, é que o sistema não comete certos tipos de erros; em particular, o Aoidos não pula versos. Se considerarmos que o sistema produz escansões erradas para menos de 1% dos versos, concluímos que o número de versos pulados por Chociay é maior do que os escandidos incorretamente pelo Aoidos.

Apesar da comparação constante entre o Aoidos e Chociay neste artigo, o por humanos e computadores só faz sentido quando se quer demonstrar que os computadores fazem um trabalho comparável àquele dos humanos. Na prática, o que se deseja é que pesquisadores e estudiosos possam utilizar o Aoidos como uma ferramenta auxiliar; o intuito certamente não é substituir o humano. Ao empregar um sistema automático de escansão, o pesquisador evita a tarefa árdua de contar as sílabas de milhares de versos e pode ocupar seu tempo com atividades mais nobres.

REFERÊNCIAS

- BEAUDOUIN, V. *Rythme et rime de l'alexandrin classique: Étude empirique des 80 000 vers du théâtre de Corneille et Racine*. 2000. Tese (Doutorado) – École des Hautes Études en Sciences Sociales.
- CHOCIAIY, R. A Identidade Formal do Decassílabo em “O Uruguai”. *Revista de Letras*, v. 34, p. 229–243, 1994.
- CHOCIAIY, R. *Os Metros do Boca: Teoria do Verso em Gregório de Matos*. São Paulo: UNESP, 1993.
- CHOCIAIY, R. *Teoria do Verso*. São Paulo: McGraw-Hill do Brasil, 1974.
- DELENTE, É.; RENAULT, R. Traitement automatique des formes métriques des textes versifiés. In: 22ème conférence sur le traitement automatique des langues naturelles, 2015. p. 432–438.
- GERVÁS, P. A Logic Programming Application for the Analysis of Spanish Verse. In: 1st international conference on computational logic, 2000. p. 1330–1344.
- MAMEDE, N.; TRANCOSO, I.; ARAÚJO, P.; VIANA, C. An Electronic Assistant for Poetry Writing. In: 9th ibero-american conference on artificial intelligence, 2004. p. 286–294.
- MATOS, G. de. *Gregório de Matos: Poemas Atribuídos, Códice Asensio-Cunha*. Edição: João Adolfo Hansen e Marcello Moreira. Belo Horizonte: Autêntica Editora, 2014. v. 3.

MATOS, G. de. *Obra Poética*. Edição: James Amado. 3 ed. Rio de Janeiro: Record, 1992.

MATOS, G. de. *Obras Poéticas de Gregório de Matos Guerra*. Florianópolis: Literatura Digital, 2016. Disponível em: <<https://goo.gl/zwJ01s>>. Acesso em: 15 dez. 2016.

MITTMANN, A. *Escansão Automático de Versos em Português*. 2016. Tese (Doutorado) – Universidade Federal de Santa Catarina.

MITTMANN, A.; WANGENHEIM, A. von; LUIZ DOS SANTOS, A. Aoidos: A System for the Automatic Scansion of Poetry Written in Portuguese. In: 17th international conference on intelligent text processing and computational linguistics, 2016.

NAVARRO-COLORADO, B.; LAFOZ, M. R.; SÁNCHEZ, N. Metrical Annotation of a Large Corpus of Spanish Sonnets: Representation, Scansion and Evaluation. In: 9th international conference on language resources and evaluation, 2016.

ROBEY, D. Scanning Dante's the *Divine Comedy*. A Computer-based Approach. *Literary and Linguistic Computing*, v. 8, n. 2, p. 81–84, 1993.

Recebido: 16 de junho de 2017

Aceito: 06 de julho de 2017