

OS SONETOS DE CAMÕES (ALGUNS DADOS ESTATÍSTICOS)

Antônio Geraldo da Cunha
Fundação Casa de Rui Barbosa

1. PRELIMINARES. No primeiro trimestre de 1995, por sugestão do Prof. Evanildo Bechara, decidimos preparar o *Índice Analítico do Vocabulário dos 'Sonetos' da 1ª edição (1595) das 'Rhythmas' de Camões*, para comemorar o quarto centenário de sua publicação. Graças ao empenho do Prof. Bechara – a quem deixamos aqui consignados os nossos mais sinceros agradecimentos – e ao alto patrocínio do Real Gabinete Português de Leitura, do Liceu Literário Português, da Real Sociedade Portuguesa Caixa de Socorros D. Pedro V e da Federação das Associações Portuguesas e Luso-Brasileiras, nosso *Índice* veio à luz nos primeiros dias de fevereiro de 1996.

1.1 No curto prazo que fora estabelecido para a elaboração do nosso trabalho, não nos foi possível preparar um estudo sobre o aspecto quantitativo do vocabulário dos *Sonetos*. Vimos hoje, decorridos alguns meses da publicação do *Índice*, apresentar alguns dados estatísticos que nos pareceram pertinentes.

2. DADOS ESTATÍSTICOS. Tendo em vista que o texto dos *Sonetos* não nos permitiria desenvolver considerações muito objetivas sobre o aspecto quantitativo do seu vocabulário, dada a sua pequena extensão, registraremos, a seguir, apenas alguns dados que nos pareceram relevantes.

2.1 Eleva-se a 5.431 o número total (N) de ocorrências vocabulares dos *Sonetos*, correspondente a 1.118 vocábulos (V) distintos. A frequência média ($F = N / V$) é igual, aproximadamente, a 4,857.

2.2 Na Tabela 1 indicamos o número dos vocábulos de frequência 1, os de frequência 2, os de frequência 3, etc., e o número de ocorrências vocabulares que corresponde a cada uma destas frequências. Assim, por exemplo, o consulente pode verificar que existem 585 vocábulos que se

documentam apenas uma vez cada um, totalizando 585 ($585 \times 1 = 585$) ocorrências vocabulares; 189 vocábulos que se documentam 2 vezes cada um, totalizando 378 ($189 \times 2 = 378$) ocorrências vocabulares, e assim por diante.

Tabela 1

v_1	f	n_1	v_2	f	n_2	v_3	f	n_3
585	x	1 = 585	3	x	19 = 57	1	x	46 = 46
189	x	2 = 378	2	x	20 = 40	1	x	47 = 47
84	x	3 = 252	1	x	21 = 21	2	x	48 = 96
55	x	4 = 220	2	x	22 = 44	1	x	57 = 57
22	x	5 = 110	3	x	23 = 69	1	x	60 = 60
31	x	6 = 186	1	x	24 = 24	1	x	68 = 68
30	x	7 = 210	2	x	25 = 50	1	x	70 = 70
13	x	8 = 104	3	x	27 = 81	1	x	72 = 72
16	x	9 = 144	3	x	28 = 84	1	x	82 = 82
9	x	10 = 90	1	x	30 = 30	1	x	85 = 85
7	x	11 = 77	1	x	33 = 33	1	x	89 = 89
8	x	12 = 96	1	x	34 = 34	1	x	93 = 93
3	x	13 = 39	1	x	35 = 35	1	x	103 = 103
4	x	14 = 56	1	x	36 = 36	1	x	119 = 119
4	x	15 = 60	1	x	39 = 39	1	x	147 = 147
4	x	16 = 64	1	x	41 = 41	1	x	150 = 150
2	x	17 = 34	2	x	42 = 84	1	x	172 = 172
3	x	18 = 54	1	x	44 = 44	1	x	270 = 270
1.069		2.759	30		846	19		1.826

$$V = v_1 + v_2 + v_3 = 1.069 + 30 + 19 = \mathbf{1.118}$$

$$N = n_1 + n_2 + n_3 = 2.759 + 846 + 1.826 = \mathbf{5.431}$$

$$F = N / V = 5.431 / 1.118 \cong \mathbf{4,857}$$

2.3 Na tabela 2 indicamos, por ordem de frequência decrescente, os cinquenta vocábulos que se documentam um maior número de vezes nos *Sonetos*. Nesta tabela vão referidos os vocábulos (seguidos de sua classe gramatical, mas apenas quando houve necessidade de distingui-los de seus homógrafos) e o número de vezes que cada um deles ocorre nos *Sonetos* (isto é, a sua frequência absoluta: f_a). Pareceu-nos conveniente indicar também, de dez em dez vocábulos, os cinco subtotais relativos à frequência absoluta de cada grupo, cumulativamente: do 1º ao 10º, do 11º ao 20º, do 21º ao 30º, do 31º ao 40º e do 41º ao 50º. Adiante de cada um dos subtotais foram indicados os percentuais de sua frequência absoluta, isto é, a relação entre o número de ocorrências desse grupo de vocábulos e o total de ocorrências vocabulares dos *Sonetos* ($N= 5431$). Estes percentuais definem a frequência relativa: f_r . Assim, por exemplo, os 10 primeiros vocábulos mais freqüentes correspondem a 24, 12% do total das ocorrências vocabulares do poema; os 20 mais freqüentes correspondem a 34,43% ,e assim por diante.

2.3.1 Observe-se que os 50 vocábulos mais freqüentes respondem, aproximadamente, pela metade de todas as ocorrências vocabulares ($N= 5.431$) dos *Sonetos*. Este valor não constitui uma particularidade dos *Sonetos* de Camões, nem deve causar estranheza, pois é praticamente uma constante, a qual não se altera, substancialmente, em textos de diferentes épocas e de estilos totalmente diversos. Não deve causar estranheza também o fato de que entre os 10 primeiros vocábulos mais freqüentes dos *Sonetos* só ocorra um *lexical* – o verbo *ser* –; todos os outros nove são *gramaticais*. Dos 50 vocábulos mais freqüentes, 17 são *lexicais* e 33 *gramaticais*.

Tabela 2: Relação dos 50 vocábulos mais freqüentes dos *Sonetos*, por ordem de frequência decrescente.

v ₁	f _a	f _r	v ₂	f _a	f _r
o ¹ (a, os, as) <i>art.</i>	270		fazer	28	
que ² <i>coni.</i>	172		olho	28	
de	150		como	27	
e	147		estar	27	
que ¹ <i>pron.</i>	119		porque	27	
me	103		já	25	
não	93		querer	25	

v1	f _a	f _r	v2	f _a	f _r
em	89		haver	24	
ser <i>vb.</i>	85		alma	23	
do (da, dos, das)	82		quem	23	
1º subtotal	1.310	24,12%	4º subtotal	2.487	45,79%
se ¹ <i>pron.</i>	72		vós	23	
no (na, nos, nas)	70		este (esta, ...)	22	
meu (minha, ...)	68		mal ² <i>sm.</i>	22	
ver	60		saber <i>vb.</i>	21	
se ² <i>coni.</i>	57		senhora	20	
a ³ <i>prep.</i>	48		seu (sua, seus,..)	20	
tão	48		assim	19	
com	47		quando	19	
ter	46		sentir	19	
amor	44		ao	18	
2º subtotal	1.870	34,43%	5º subtotal	2.690	49,53%
por <i>prep.</i>	42		v ₁ , v ₂ ⇒ vocábulos f _a ⇒ freqüência absoluta f _r ⇒ freqüência relativa		
um (uma, uns,...)	42				
mas	41				
poder <i>vb.</i>	39				
mais	36				
o ² (a os as) <i>pron.</i>	35				
vos	34				
vosso (vossa...)	33				
vida	30				
dar	28				
3º subtotal	2.230	41,06%			

2.4 Na Tabela 3 vão relacionados os 20 verbos, os 20 substantivos e os 20 adjetivos mais freqüentemente usados nos *Sonetos*.

Tabela 3: Relação dos 20 verbos, 20 substantivos e 20 adjetivos mais freqüentes dos *Sonetos*.

VERBOS		SUBSTANTIVOS		ADJETIVOS	
ser	85	amor	44	duro	12
ver	60	vida	30	grande	11
ter	46	olho	28	triste	11
poder	39	alma	23	claro	10
dar	28	mal	22	lindo	10
fazer	28	senhora	20	longo	9
estar	27	esperança	17	puro	9
querer	25	mundo	16	doce	8
haver	24	pena	13	alto	7
saber	21	bem	12	áspero	7
sentir	19	lembrança	12	contente	7
tomar	16	pensamento	12	formoso	7
viver	15	tempo	12	novo	7
perder	14	dia	11	vão	7
pôr	13	vista	11	belo	6
ficar	12	água	10	gentil	6
ir	12	mágoa	10	vivo	6
passar	10	parte	10	brando	5
tornar	10	dor	9	ditoso	5
andar	9	fogo	9	grão	5

3. COMENTÁRIOS. Cumpre assinalar que, dos valores estatísticos aqui apresentados, não se pode tirar conclusões definitivas, pois além da pequena extensão dos *Sonetos*, considerada do ponto de vista estatístico, não

dispomos ainda de um número suficiente de índices de freqüência para um estudo comparativo. De qualquer maneira, apenas a título ilustrativo, podemos confrontar os substantivos mais freqüentes dos *Sonetos* com os de *Os Lusíadas* (cf. IAVL) e com os da *Prosopopéia*, de Bento Teixeira, cujo *Índice Analítico* já foi por nós preparado, mas ainda se conserva inédito.

Tabela 4: Relação dos 10 substantivos mais freqüentes dos *Sonetos* de Camões, comparados com os 10 mais freqüentes de *Os Lusíadas* e os 10 mais freqüentes da *Prosopopéia*, de Bento Teixeira.

SONETOS	OS LUSÍADAS	PROSOPOPÉIA
amor	gente	mar
vida	terra	rei
olho	rei	reino
alma	mar	tempo
mal	água	céu
senhora	mundo	peito
esperança	céu	gente
mundo	reino	sorte
pena	peito	fama
bem	deus	parte

3.1 Não nos cabe emitir um critério de valor sobre o paralelismo aqui estabelecido entre *Os Lusíadas* e a *Prosopopéia*, pois, como referimos acima, este confronto visa apenas a apontar as características diferenciais dos vocabulários dos dois poemas. Note-se, a propósito, que dos 10 substantivos mais freqüentes dos *Sonetos*, apenas 1 – *mundo* – ocorre em *Os Lusíadas*, enquanto que, na *Prosopopéia*, existem seis substantivos – *céu*, *gente*, *mar*, *peito*, *rei* e *reino* – que são também comuns a *Os Lusíadas*. A explicação desta aparente anomalia decorre do fato de que os *Sonetos* pertencem ao gênero lírico, enquanto que *Os Lusíadas* e a *Prosopopéia* são poemas épicos.

3.2 Jorge de Sena, profundo conhecedor da obra de Camões, deixou numerosos estudos, quase todos redigidos em Santa Bárbara, na Califórnia, em março e abril de 1971, os quais foram reunidos num belo livro, intitulado

« Estudos sobre o Vocabulário de ‘Os Lusíadas’ », publicado postumamente, em 1982, em Lisboa, com uma Nota Prévia de Mécia de Sena e um Prefácio de Luís de Sousa Rebelo. Referindo-se à diferenciação entre os vocabulários de obras de estilos diferentes, observa o ilustre camonista, à pág. 21: “Porque, se o vocabulário de uma obra é característico de um autor *nela* (será óbvio que o vocabulário depende, também em grande parte, da natureza da obra e das intenções estilísticas do seu autor ao compô-la), e nos permite analisar o nível linguístico, mais ou menos refinado e peculiar, que o autor pretendeu realizar com ele, só verdadeiramente será característico e definidor um *índice analítico*, em que todas as ocorrências vocabulares sejam registadas e classificadas.”

3.3 Refere-se Jorge de Sena ao nosso *Índice Analítico do Vocabulário de ‘Os Lusíadas’* (o IAVL), cuja 1ª edição foi publicada em 1966. Com base nele, o grande camonista expendeu eruditíssimos comentários sobre algumas centenas de vocábulos, nos quatorze capítulos dos seus *Estudos*. São estudos que revelam, de par com uma fina sensibilidade no exame estilístico do Poema, um extenso conhecimento da obra camoniana, cujo vocabulário ele trata com perícia e rigor científico. Lamentando a inexistência de índices analíticos de outros autores do século XVI, Jorge de Sena volta a comentar, à pág. 23: « Claro que, para o comparativismo estilístico, continuamos limitados pela falta de um índice analítico do vocabulário do século XVI ».

4. CONCLUSÃO. Outros aspectos do vocabulário dos *Sonetos* poderiam ser aqui focalizados. Estávamos mesmo tentados a tratar de algumas questões relativas a este ou àquele vocábulo, a esta ou àquela variante, mas julgamos prudente adotar uma atitude cautelosa, não nos arriscando a emitir juízos precipitados. Nosso propósito neste artigo – repetimos – foi apresentar alguns dados estatísticos do vocabulário dos *Sonetos*, para ulterior confronto com os *índices analíticos* de outras obras que, como esperamos, venham a ser publicados. No Curso de Lexicografia que ministramos na Pós-Graduação em Letras da Universidade do Estado de São Paulo (USP), no primeiro semestre de 1996, e na Universidade do Estado do Rio de Janeiro (UERJ), no segundo semestre, mencionamos a importância da elaboração destes *índices analíticos*, entre outras tarefas urgentes para o desenvolvimento da nossa Lexicografia. Dado o interesse manifestado pelos mestrandos, acreditamos que estes trabalhos venham a ser em breve publicados.
