

SÔBRE O VOCABULÁRIO POÉTICO BRASILEIRO DA FASE COLONIAL

JEAN ROCHE

INTRODUÇÃO

O presente artigo representa apenas uma parte, reduzida, dos resultados das pesquisas realizadas no nosso Centro de Lexicologia e Estilística de Língua Portuguesa, criado em 1967-68 em Toulouse. Não pretendemos senão apresentar as orientações do levantamento lexical e umas sugestões para a sua exploração.

Responsável pelo Centro, desejamos que o patrimônio inteiro da literatura de língua portuguesa fôsse incluído: estamos conduzindo trabalhos paralelos sôbre a literatura de Portugal e do Brasil. Aqui falaremos unicamente do Brasil e de lexicologia, já que as nossas pesquisas sôbre a estilística justificariam uma outra apresentação (Estilística do romance brasileiro, principalmente contemporâneo).

A nossa primeira preocupação foi associar pesquisa e ensino, como se deve. Professôres e alunos precisam, antes de mais nada, conhecer — ou pelo menos sondar — a língua poética, isto é, a língua literária por excelência, a mais distante da prosa quotidiana. Assim, para começar, escolhemos trinta poetas, desde Bento Teixeira até João Cabral de Melo Neto, extraímos de suas obras poemas-amostras para estabelecer uma lista reveladora do seu vocabulário preferido e para comparar as listas, por século, período ou escola. O que tentamos é balizar uma *terra incógnita* (porque faltam os levantamentos quantitativos) e avaliar melhor, pela relação com as

nossas estacas, a originalidade ou a contribuição de tal poeta, acompanhando em diacronia a evolução da língua poética brasileira.

Começaremos, pois, pelo período chamado de fase colonial, do Século XVI até a entrada do século XIX, de Bento Teixeira até José Bonifácio de Andrada e Silva. Este estudo poderá ser comparado com o artigo que acabamos de publicar sobre o vocabulário poético brasileiro do Romantismo até o Neomodernismo (1), e tanto mais facilmente que o modo de trabalhar foi o mesmo.

Na impossibilidade de reunir, num *corpus* integral, a totalidade do vocabulário usado por todos os poetas desta fase, limitamos a lista provisória a dez:

- | | |
|------------------------------|--------------------------------------|
| 1 Bento Teixeira | 6 Caldas Barbosa |
| 2 Gregório de Matos | 7 J. Basílio da Gama |
| 3 Manuel Botelho de Oliveira | 8 Tomás Antônio Gonzaga |
| 4 Sta. Rita Durão | 9 M. I. Silva Alvarenga |
| 5 Cláudio Manuel da Costa | 10 José Bonifácio de Andrada e Silva |

E extraímos de uma antologia (2) certo número de poesias, indicadas pelo acaso, para evitar uma escolha arbitrária da nossa parte (3). A relação dos poemas arrolados fica anexada mais abaixo (4).

Lembraremos que cada poema, e depois o conjunto dos poemas de um mesmo autor, fornece uma lista alfabética dos vocábulos encontrados, com indicação das formas e das frequências (5). Por poeta, estabelecemos uma lista de vocábulos por ordem de frequência decrescente e várias tabelas numéricas (posição e frequência, distribuição por classe morfológica etc.), destinadas também a serem publicadas (há matéria para um volume) desde que haja possibilidade, porque serão úteis a todos os estudiosos.

Muito mais modesta, a presente explanação constará de considerações sobre a distribuição dos vocábulos arrolados, da reprodução dos vocábulos preferenciais (usados por um poeta com uma frequência superior ou igual a 2), da elaboração da lista dos vocábulos preferenciais comuns (usados por pelo menos dois poetas do período) e de um comentário sobre a distribuição das ocorrências, reveladora de uma maneira pessoal de escrever.

-
1. ROCHE, Jean, *Du vocabulaire poétique brésilien*, CARAVELLE, CMHLB, n.º 15, 1970, Toulouse, p. 73-128.
 2. *Antologia dos poetas brasileiros da fase colonial*, por Sérgio BUARQUE DE HOLLANDA, Dep. de Imp. Nac., Rio de Janeiro, 1952-53, 2 vol. 337 + 297 p.
 3. MULLER, Charles, *Initiation à la statistique linguistique*, Larousse, Paris, 1968, 247 p.
 4. Cf. lista em apêndice, p.
 5. EMORINE, Jacques, *Lexique et Analyse lexicale de l'Auto da Compadecida*. F.U.C. Porto Alegre, 1969, 253 p.

PRIMEIRA PARTE. ESTUDO DOS VOCABULOS

I) DADOS NUMÉRICOS E COMENTARIOS.

Nesta primeira parte, voltada ao estudo dos vocábulos isto é, às "palavras" diferentes encontradas, que constam da lista alfabética com as frequências, utilizaremos os símbolos *V* para a soma total dos vocábulos e *N* para a soma total das ocorrências, *v* para a soma das nocionais e *n* para a soma das ocorrências das palavras nocionais, *v'* para a soma das palavras gramaticais e *n'* para a soma das ocorrências das palavras gramaticais. Eis os resultados do nosso trabalho:

$$\begin{aligned} V &= 3685 & , & & v &= 3106 & , & & v' &= 579 \\ N &= 9206 & , & & n &= 4917 & , & & n' &= 4289 \end{aligned}$$

Isto quer dizer que foram estabelecidas 9206 fichas, materializando tantas ocorrências, ou *unidades-de-texto* (expressão melhor do que "palavra" no sentido comum). E as diversas listas e tabelas foram organizadas a partir dêste material, que vamos resumir, indicando qual foi a contribuição numérica de cada um dos dez poetas (designados a seguir ou pelas suas iniciais ou pela ordem, sempre a mesma, de 1 a 10).

1. VALORES E PERCENTAGENS:

POETAS	V	N	v	n	Vocábulos de freq. 1		
					N.º	= % V	= % N
1 Bento Teixeira	292	581	233	305	222	76	38
2 Greg. de Matos	339	753	277	383	240	70	32
3 M. Bot. de Oliveira	720	2223	624	1153	455	63	20
4 Sta. Rita Durão	491	1085	425	568	373	75	34
5 Cláudio M. da Costa	251	565	203	303	174	69	30
6 Caldas Barbosa	205	642	157	375	115	57	17
7 J. Basílio da Gama	365	858	316	457	262	71	30
8 Tomás Antônio Gonzaga	342	943	279	483	236	69	25
9 M. I. Silva Alvarenga	316	874	269	475	215	68	24
10 José Bonifácio de A. S.	364	682	323	415	282	77	41
Total	3685	9206	3106	4917	2574	—	—
Média	368	920	310	491	257	69	28

A relação V/N tem um valor próprio de cada poeta, revelando a sua originalidade. Por outro lado, o coeficiente de repetição ($9206/3685 = 2,49$) não tem sentido, porque há mais de dois terços de "hapax" (vocabúlos de frequência 1) e porque as palavras nocionais são menos frequentemente repetidas do que as gramaticais (p. ex., Bento Teixeira $v' = 59$, $n' = 276$, Cláudio M. da Costa $v' = 48$, $n' = 262$, José Bonifácio $v' = 31$, $n' = 267$, e no total $v' = 579$, $n' = 4289$ versus $v = 3106$, $n = 4917$). A tabela acima permite uma avaliação da "riqueza de vocabulário" de um autor. Mas como as palavras nocionais têm o papel relevante neste vocabulário, vamos construir outra tabela, com a sua distribuição por classe morfológica, com referência a valores e percentagens de V que representam:

POETAS	Valôres absolutos de V				= Percentagens de V			
	Verbo	Adje- tivo	Subs- tantivo	Ad- vérbio	V	Q	S	AD
1 B. T.	41	74	109	9	14	25	37	3
2 G. M.	71	54	137	15	21	16	40	4
3 M. B. O.	133	163	302	26	18	22	41	3
4 S. R. D.	116	90	202	17	23	18	40	3
5 C. M. C.	58	38	86	21	23	15	34	8
6 C. B.	59	27	54	17	28	13	26	8
7 J. B. G.	95	73	122	26	26	20	33	7
8 T. A. G.	75	41	144	19	22	12	42	5
9 S. A.	80	68	111	10	25	21	35	3
10 J. B. A. S.	72	88	154	9	19	24	42	2
Total	800	716	1421	169	—	—	—	—
Média	80	71	142	16	21,7	19,4	38,5	4,6

2. COMENTÁRIOS.

Ficamos, já, em melhores condições para uma segunda aproximação da variedade do vocabulário de cada poeta, até pela comparação das proporções das classes entre as quais são distribuídos os valores de tantos vocábulos diferentes. Mas, antes de comentar o quadro precedente, apresentaremos uma sinopse:

Percentagens de V	Verbos	Qualificativos	Substantivos	Advérbios	Palavras Nocionais
Médio	21,7	19,4	38,5	4,6	84,1
Máximo	28	25	42	8	88
Mínimo	14	12	15	2	76
Dif. Máx - Mín	14	13	27	6	12
% Dif. / Máx	50%	52%	64%	75%	13%

Se a percentagem do total de palavras nocionais varia pouco, as oscilações das percentagens das classes, nunca inferiores à metade do valor máximo, alcançam a três quartos. A sinopse torna-se como uma moldura enquadrando a posição de um ou outro poeta, e que evidencia a transposição sobre o gráfico n.º 1.

O vocabulário de Bento Teixeira tem uma riqueza média em substantivos e máxima em adjetivos, enquanto a proporção de verbos e advérbios é inferior de um terço à média. O de José Bonifácio tem a mais alta proporção de substantivos e a segunda de adjetivos, mas mais verbos e menos advérbios do que o precedente.

Mas podemos, também, comparar as percentagens dos três períodos em que distribuiremos os poetas da fase colonial:

Períodos	Poetas	V	Q	S	AD	P. N.
XVI-XVII	1, 2, 3,	17,6	21,0	39,3	3,3	81,2
XVIII	4, 5, 6, 7, 8, 9,	24,5	16,5	35,0	5,6	81,6
XIX	10	19,0	24,1	42,4	2,5	88,0

Aparece, por exemplo, a tendência dos poetas do século XVIII em aumentar o papel do verbo e reduzir o do substantivo, com os corolários — redução do adjetivo e acréscimo do advérbio. José Bonifácio pertence ao XIX pela tendência para a ênfase nominal — mas com uma alta proporção adjetival que somente nos Simbolistas reencontraremos. As diferenças entre os períodos tornam-se mais nítidas ainda no gráfico n.º 2, que ilustra uma evolução histórica. Mas revela também a correlação entre a classe do substantivo e a do adjetivo e a correlação entre a do verbo e a

do advérbio, muito mais claramente do que o gráfico correspondente para os poetas românticos-neomodernistas. Entretanto, o que aparece em todos os períodos da poesia brasileira é uma dupla pulsação das proporções de verbos e substantivos em sentido contrário: o século XVIII é de mínimo nominal e de máximo verbal.

O exame dos dados puramente numéricos já nos proporciona várias possibilidades: a de determinar a maneira que tem um poeta (ou um grupo) de diversificar a distribuição ou a composição do seu vocabulário (sincronia); a de acompanhar uma evolução histórica (diacronia) nos mesmos domínios; a de verificar as relações entre os termos correlativos ou antinômicos dos binômios S+Q, V+AD, SxV, quer dizer, tradicionais classes morfológicas ou oposição paradigma-sintagma. Mas seriam resultados desproporcionados ao trabalho de coleta dos vocábulos. E convém, agora, voltar às próprias listas.

II) OS VOCÁBULOS PREFERENCIAIS.

Dada a impossibilidade de reproduzir aqui os 3685 vocábulos diferentes encontrados nos poemas arrolados, limitar-nos-emos a indicar os que cada poeta empregou pelo menos 2 vezes, distribuídos por classe morfológica e na ordem de frequência decrescente. Eles perfazem um total de 507 palavras nocionais (5,2% do total V) com 2564 ocorrências (27,8% do total N), ou respectivamente 16,0% de *v* e 52,1% de *n*, quer dizer, mais da quarta parte das ocorrências globais e mais da metade das ocorrências de palavras nocionais. As referidas percentagens confirmam o interesse do estudo a seguir, pela alta representatividade daquelas séries: elas vão permitir a determinação dos vocabulários preferenciais individuais. Os "hapax", ainda que numerosos (2574/3685), não podem ser estudados nos limites do presente artigo; menos do que ninguém negaremos o papel que eles têm na expressão temática ou pessoal; mas eles merecem outro tipo de exame e análise, que devemos por enquanto adiar (6). E passamos a considerar os vocábulos mais frequentes dos nossos dez poetas.

1) AS LISTAS DE VOCÁBULOS PREFERENCIAIS INDIVIDUAIS:

1. BENTO TEIXEIRA

1 ser	8	mar	6	alto	3	tão	3
2 fazer	5	valor	4	bom	2		
3 ver	—	Albuquerque	3	duro	—		
4 cantar	4	céu	—	escuro	—		

6. Cf. tabela em apêndice.

5	chamar	—	rei	—	inculto	—
6	estar	—	ar	2	quieto	—
7	esperar	3	carro	—	vivo	—
8	ter	—	entrada	—		
9	dar	2	fama	—		
10	derivar	—	fúria	—		
11	dever	—	irmão	—		
12	romper	—	lajem	—		
13			língua	—		
14			licor	—		
15			muro	—		
16			obra	—		
17			pedra	—		
18			pôrto	—		
19			voz	—		

2. G. DE MATOS GUERRA

1	ser	8	dia	4	arrependido	3	mais	7
2	ver	—	senhor	—	cruel	2	tão	4
3	querer	6	vida	—			bem	2
4	ofender	4	ar	3			mui	—
5	saber	—	flor	—				
6	dar	3	luz	—				
7	delinquir	—	ôlho	—				
8	dizer	—	ovelha	—				
9	governar	—	pena	—				
10	bastar	2	abraço	2				
11	crescer	—	ai	—				
12	gozar	—	aurora	—				
13	matar	—	beleza	—				
14	ouvir	—	cinza	—				
15	passar	—	coração	—				
16	pecar	—	dor	—				
17	perder	—	gemido	—				
18	poder	—	glória	—				
19	sentir	—	hora	—				
20	trocar	—	Jesus	—				
21			mal	—				
22			maldade	—				
23			medida	—				
24			mulato	—				
25			mundo	—				
26			negro	—				
27			piedade	—				
28			pranto	—				

29	salvação	—
30	sol	—
31	vaidade	—
32	verdade	—

3. MANUEL BOTELHO DE OLIVEIRA

1	ser	69	fruto	16	belo	9	mais	16
2	ter	20	gosto	8	vário	—	muito	11
3	ver	14	terra	7	doce	6	tão	10
4	fazer	11	mundo	6	gostoso	5	sempre	9
5	dar	10	agrado	5	próprio	—	também	—
6	ctiar	9	céu	—	apetecido	4	aqui	6
7	querer	—	côr	—	fermoso	—	porém	4
8	estar	6	Europa	—	deleitoso	3	acaso	3
9	parecer	—	jardim	—	galhardo	—	bem	—
10	faltar	4	ôlho	—	grande	—	enfim	—
11	gerar	—	açúcar	4	liberal	—	menos	—
12	ir	—	esperança	—	maior	—	docemente	2
13	produzir	—	peixe	—	melhor	—	fêrtilmente	—
14	amar	3	regalo	—	mimoso	—	quase	—
15	desejar	—	vantagem	—	parecido	—	só	—
16	desprezar	—	amor	3	pequeno	—	tanto	—
17	dizer	—	Anarda	—	picante	—		
18	levar	—	capela	—	puro	—		
19	morrer	—	farinha	—	amarelo	2		
20	mostrar	—	legume	—	amoroso	—		
21	poder	—	maio	—	assado	—		
22	pôr	—	mar	—	azêdo	—		
23	preferir	—	maré	—	breve	—		
24	prezar	—	mês	—	cândido	—		
25	tirar	—	pão	—	copioso	—		
26	adornar	2	rei	—	cozido	—		
27	alcançar	—	rosa	—	estranho	—		
28	alentar	—	sabor	—	excelente	—		
29	apurar	—	sustento	—	fácil	—		
30	beber	—	água	2	famoso	—		
31	cair	—	alegria	—	fecundo	—		
32	calar	—	ano	—	feliz	—		
33	celebrar	—	apetite	—	frio	—		
34	conhecer	—	ar	2	lêdo	—		
35	cozer	—	arvoredo	—	obliquo	—		
36	dedicar	—	Bahia	—	pobre	—		
37	desfazer	—	bôca	—	raro	—		
38	encerrar	—	bondade	—	regalado	—		

39	fabricar	—	cana	—	rubicundo	—
40	falar	—	casta	—	sadio	—
41	ficar	—	castanha	—	salgado	—
42	gostar	—	cheiro	—	singular	—
43	gozar	—	conduto	—	soberano	—
44	levantar	—	cristal	—	suave	—
45	lograr	—	cuidado	—	verde	—
46	matar	—	cultura	—	vermelho	—
47	perder	—	desfavor	—	vivo	—
48	plantar	—	engenho	—		
49	servir	—	Espanha	—		
50	solicitar	—	espinho	—		
51			favor	—		
52			figo	—		
53			flor	—		
54			flora	—		
55			gente	—		
56			gentio	—		
57			glória	—		
58			graça	—		
59			Holanda	—		
60			ilha	—		
61			império	—		
62			lagostim	—		
63			limão	—		
64			majestade	—		
65			mandioca	—		
66			mangava	—		
67			manjar	—		
68			mantimento	—		
69			mão	—		
70			mato	—		
71			melancia	—		
72			milho	—		
73			natureza	—		
74			nome	—		
75			pé	—		
76			pimenta	—		
77			pomar	—		
78			pomo	—		
79			primor	—		
80			rôsto	—		
81			sêde	—		
82			sol	—		
83			substância	—		
84			tempêro	—		

85	tempo	—
86	trabalho	—
87	vale	—
88	ventura	—
89	vestido	—
90	vida	—

4. J. DE SANTA RITA DURÃO

1	ser	11	donzela	4	vasto	4	mais	8
2	dar	7	Caeté	3	belo	3	tão	6
3	ver	5	copia	—	doce	—	ali	2
4	fazer	4	gente	—	formoso	—	então	—
5	ter	—	guerra	—	ardente	2	pouco	—
6	buscar	3	Jararaca	—	bravo	—		
7	correr	—	mar	—	claro	—		
8	ficar	—	nome	—	possante	—		
9	morrer	—	planta	—	prezado	—		
10	tomar	—	Paraguaçu	—	prodigioso	—		
11	vir	—	sertão	—	puro	—		
12	acender	2	sombra	—				
13	arder	—	água	2				
14	chamar	—	amante	—				
15	cobrir	—	bôca	—				
16	convir	—	campo	—				
17	espremer	—	cana	—				
18	estar	—	chama	—				
19	fugir	—	erva	—				
20	ir	—	espôso	—				
21	olhar	—	face	—				
22	parecer	—	fama	—				
23	pôr	—	foz	—				
24	querer	—	frescura	—				
25	saber	—	gôsto	—				
26	transplantar	—	imagem	—				
27	voar	—	lado	—				
28			légua	—				
29			milho	—				
30			mimo	—				
31			modo	—				
32			nação	—				
33			ôlho	—				
34			pai	—				
35			pais	—				
36			peito	—				
37			príncipe	—				

38	raiz	—
39	razão	—
40	rio	—
41	rosto	—
42	suspiro	—
43	temor	—
44	terra	—
45	terror	—

5. C. MANUEL DA COSTA

1 ser	16	Nize	8	certo	2	aonde	4
2 estar	12	tronco	4	diferente	—	onde	—
3 ver	8	monte	3			quanto	—
4 fazer	5	saudade	—			tão	—
5 deixar	4	alma	2			agora	2
6 andar	3	dia	—			aqui	—
7 ouvir	—	formosura	—			como	—
8 poder	—	gado	—			menos	—
9 seguir	—	memória	—			tanto	—
10 vir	—	montanha	—				
11 chorar	—	pastor	—				
12 cuidar	—	pranto	—				
13 mostrar	—	rio	—				
14 querer	—	rosto	—				
15 respirar	—	vale	—				
16		vez	—				

6. DOMINGOS CALDAS BARBOSA

1 morrer	27	ternura	11	brasileiro	12	devagar	10
2 ser	22	amor	7	último	4	sempre	6
3 ir	15	gente	—	cruel	2	etc.	5
4 ter	9	morte	6	doce	—	mais	—
5 ver	8	chama	4	lisonjeiro	—	bem	3
6 acabar	7	doçura	—	triste	—	já	—
7 poder	—	gosto	—	verdadeiro	—	cá	2
8 saber	—	Brasil	2			lá	—
9 gostar	5	ciúme	—				
10 dar	4	coração	—				
11 querer	—	esperança	—				
12 vir	—	exemplo	—				
13 amar	3	mel	—				
14 expirar	—	mundo	—				
15 perder	—	Nhanhá	—				
16 viver	—	parte	—				

17 chegar	2	reino	—
18 conhecer	—	saudade	—
19 haver	—	veneno	—
20 matar	—	vida	—
21 prometer	—		
22 sentir	—		
23 sofrer	—		
24 sustentar	—		

7. J. BASÍLIO DA GAMA

1 ser	5	chama	6	alto	3	mais	5
2 poder	4	peito	—	triste	—	já	4
3 ter	—	rio	—	verde	—	ainda	3
4 correr	3	incêndio	5	abrasador	2	entanto	—
5 envolver	—	mão	—	escuro	—	agora	2
6 fugir	—	vento	—	largo	—	ali	—
7 ver	—	campina	4	negro	—	aqui	—
8 arrebatat	2	margem	—	pátrio	—	debalde	—
9 cobrir	—	valeroso	—	valeroso	—	pouco	—
10 dizer	—	tempo	—			quando	—
11 durar	—	água	3			sempre	—
12 entregar	—	braço	—				
13 espalhar	—	Cacambo	—				
14 fazer	—	chão	—				
15 ir	—	índio	—				
16 saber	—	morte	—				
17 sacudir	—	norte	—				
18 sentir	—	pé	—				
19 tornar	—	aljava	2				
20		ar	—				
21		arco	—				
22		areia	—				
23		Cepé	—				
24		céu	—				
25		cinza	—				
26		fogo	—				
27		fumo	—				
28		gado	—				
29		gente	—				
30		gruta	—				
31		imagem	—				
32		luz	—				
33		natureza	—				
34		ôlho	—				
35		onda	—				

36	parte	—
37	pena	—
38	rosto	—
39	seta	—
40	sorte	—
41	voz	—

8. T. ANTÔNIO GONZAGA

1 ser	14	Marília	18	belo	9	já	5
2 mudar	10	sorte	17	só	8	mais	—
3 ter	9	graça	14	feliz	3	sempre	2
4 dar	8	estrêla	7	gentil	—		
5 ir	—	ôlho	5	bom	2		
6 ver	—	tronco	4	fino	—		
7 fazer	4	amor	3	frio	—		
8 querer	3	cabelo	—	próprio	—		
9 acabar	2	céu	—	quente	—		
10 arrastar	—	côr	—	vivo	—		
11 cair	—	corpo	—				
12 cobrir	—	dia	—				
13 fugir	—	gôsto	—				
14 gastar	—	luz	—				
15 haver	—	monte	—				
16 levantar	—	pastor	—				
17 padecer	—	pastôra	—				
18 procurar	—	planta	—				
19 supor	—	primavera	—				
20 tornar	—	tempo	—				
21 viver	—	ano	2				
22		bruto	—				
23		deus	—				
24		doença	—				
25		estação	—				
26		face	—				
27		flor	—				
28		gado	—				
29		homem	—				
30		mão	—				
31		mundo	—				
32		passo	—				
33		rebanho	—				
34		rosto	—				
35		terra	—				
36							

9. M. I. DA SILVA ALVARENGA

1	entregar	12	Glaura	13	triste	11	já	5
2	ser	10	amor	6	doce	8	pois	—
3	deixar	8	terra	—	duro	7		
4	brotar	5	alma	5	desgraçado	5		
5	negar	—	bem	—	lindo	—		
6	submergir	—	cajueiro	—	belo	4		
7	ver	—	calma	—	brando	3		
8	dizer	4	fado	—	ditoso	—		
9	chorar	3	lida	—	cruel	2		
10	vir	—	rigor	—	escuro	—		
11	voar	—	senhor	—	infeliz	—		
12	alegrar	2	suspiro	—	ingrato	—		
13	chegar	—	desventura	4	mimoso	—		
14	existir	—	flor	—	mudo	—		
15	levar	—	saudade	—	precioso	—		
16	oferecer	—	ardor	3	terno	—		
17	padecer	—	lirio	—	cerde	—		
18	procurar	—	pastor	—				
19			piedade	—				
20			rosa	—				
21			sombra	—				
22			céu	2				
23			côr	—				
24			dia	—				
25			favor	—				
26			fruto	—				
27			giro	—				
28			luz	—				
29			melancolia	—				
30			sepultura	—				
31			sol	—				
32			ternura	—				
33			tronco	—				
34			vento	—				

10. JOSÉ BONIFÁCIO DE ANDRADA E SILVA

1	cantar	7	filho	4	divino	5	então	4
2	fazer	4	ouro	—	doce	4	assim	2
3	levantar	3	terra	—	nôvo	3	manso	—
4	sentir	—	dia	3	aceso	2		
5	ter	—	frente	—	altivo	—		
6	ver	—	pátria	—	alvo	—		
7	açoitar	2	poesia	—	horrendo	—		

8 aprender	—	vate	—	negro	—
9 arrancar	—	virtude	—	próprio	—
10 celebrar	—	voz	—	soberbo	—
11 ditar	—	altar	2	sonoro	—
12 reinar	—	camena	—	velho	—
13 restar	—	canto	—	vermelho	—
14		corda	—	vil	—
15		crime	—		
16		deusa	—		
17		filha	—		
18		fome	—		
19		gênio	—		
20		gente	—		
21		glória	—		
22		Grecia	—		
23		luz	—		
24		mãe	—		
25		monstro	—		
26		musa	—		
27		sangue	—		
28		som	—		
29		verso	—		

2. A LISTA DOS VOCABULOS PREFERENCIAIS COMUNS:

Antes de comparar os vocábulos preferenciais individuais, será mais conveniente estabelecer uma classificação preliminar dos vocábulos que apareceram, com uma frequência superior ou igual a 2 em pelo menos duas listas, e a que chamaremos de Vocábulos Preferenciais Comuns (isto é, comuns aos poetas da fase colonial), Símbolo V. P. C.

	Verbos			Substantivos			Substantivos			Adjetivos					
	P	F	NP	P	F	NP	P	F	NP	P	F	NP			
Acabar	17	9	2	Água	28	7	3	Rosa	33	6	2	Alto	8	6	2
Amar	24	6	2	Alma	28	7	2	Rosto	16	10	5	Belo	1	25	4
Cair	37	4	2	Amor	2	19	4	Sangue	33	6	2	Bom	12	5	2
Cantar	16	11	2	Ano	47	4	2	Saudade	20	9	3	Cruel	8	6	3
Celebrar	37	4	2	Ar	20	9	4	Senhor	20	9	2	Doce	2	23	5
Chamar	24	6	2	Bôca	47	4	2	Sol	33	6	3	Duro	4	9	2
Chegar	37	4	2	Cana	47	4	2	Sorte	2	19	2	Escuro	8	6	3
Chorar	29	5	2	Céu	9	15	5	Suspiro	28	7	2	Feliz	12	5	2
Cobrir	24	6	3	Chama	33	6	2	Tempo	20	9	3	Formoso	6	7	2
Conhecer	37	4	2	Cinza	47	4	2	Ternura	11	13	2	Frio	15	4	2
Correr	24	6	2	Côr	16	10	3	Terra	1	21	5	Negro	15	4	2
Dar	5	34	6	Coração	47	4	2	Tronco	16	10	3	Proprio	4	9	3

Deixar	13	12	2	Dia	10	14	5	Vale	47	4	2	Puro	12	5	2
Dizer	13	12	4	Esperança	33	6	2	Vento	28	7	2	Triste	3	16	3
Entregar	11	14	2	Face	47	4	2	Vida	25	8	3	Verde	6	7	3
Estar	8	26	5	Fama	47	4	2	Voz	28	7	3	Vermelho	15	4	2
Fazer	4	35	7	Favor	47	4	2					Vivo	8	6	3
Ficar	29	5	2	Flor	15	11	4								
Fugir	20	7	3	Fruto	4	18	2								
Gostar	20	7	2	Gado	33	6	3								
Gozar	37	4	2	Gente	6	16	5								
Haver	37	4	2	Glória	33	6	3								
Ir	7	31	5	Gôsto	5	17	4								
Levantar	20	7	3	Graça	6	16	2								
Levar	29	5	2	Imagem	47	4	2								
Matar	24	6	3	Luz	12	12	5								
Morrer	6	33	3	Mão	20	9	3								
Mostrar	29	5	2	Mar	12	12	3								
Ouvir	29	5	2	Milho	47	4	2								
Parecer	19	8	2	Monte	33	6	2								
Perder	20	7	3	Morte	20	9	2								
Poder	10	19	5	Mundo	12	12	4								
Por	29	5	2	Natureza	47	4	2								
Querer	8	26	6	Nome	44	5	2								
Romper	37	4	2	Olho	6	16	5								
Saber	13	12	4	Parte	47	4	2								
Sentir	17	9	4	Pastor	25	8	3								
Ser	1	163	9	Pé	44	5	2								
Ter	3	52	7	Peito	25	8	2								
Tornar	37	4	2	Piedade	44	5	2								
Ver	2	67	10	Planta	33	6	2								
Vir	12	13	4	Pranto	47	4	2								
Viver	29	5	2	Rei	33	6	2								
Voar	29	5	2	Rio	16	10	3								

Advérbios

Agora	9	4	2
Ali	9	4	2
Aqui	5	10	3
Bem	6	8	3
Então	7	6	2
Já	4	17	4
Mais	1	46	6
Menos	8	5	2
Pouco	9	4	2
Sempre	3	19	4
Tanto	9	4	2
Tão	2	27	5

A lista dos vocábulos preferenciais comuns conta 44 verbos, com 716 ocorrências; 60 substantivos, com 472 ocorrências; 17 adjetivos, com 148 ocorrências; 12 advérbios, com 154 ocorrências; ou seja no total 133 vocábulos, com 1490 ocorrências (= 16% de *N*, 30,3% de *n*).

A lista dos V.P.C. representa menos da vigésima parte dos vocábulos nocionais, mas quase a terça parte das suas ocorrências, o que lhe assegura uma representatividade inegável. Ela vai constituir o fio de Ariana para orientar nosso estudo do vocabulário dos dez poetas.

A primeira observação é que os substantivos são mais numerosos do que os verbos. Dêstes devemos subtrair os *ser*, *estar*, *ter*, *haver*, que têm um papel quase gramatical, e que totalizam 278 ocorrências. Os 40 verbos

restantes tem apenas 438 ocorrências, isto é menos do que os nomes. Na análise dos vocabulários individuais devemos, portanto, examinar em primeiro lugar os substantivos.

Apesar da apresentação por ordem alfabética, que facilitará os confrontos ulteriores, devemos prestar atenção à frequência, e correlativamente à posição dos vocábulos, comuns ou individuais. Observamos que as frequências são geralmente mais baixas do que se podia esperar e que não existe nome preponderante. Assim para os adjetivos.

Outra averiguação é que o próprio número de poetas em cujas listas individuais aparecem os V. P. C. é também mais baixo do que se pensava: há 1 verbo usado pelo 10 poetas, 1 por 9, 2 por 7, 2 por 6, e todos os outros de 5 para baixo; há 6 substantivos usados por 5 poetas, e todos os outros de 4 para baixo; idem 1 adjetivo; 1 advérbio por 6, 1 por 5, e os outros de 4 para baixo. É portanto indispensável examinar com cuidado a repartição dos V. P. C., e a proporção que eles representam em cada lista individual.

III) PONDERAÇÕES.

Passando agora à exploração das listas individuais (à luz da geral), contentar-nos-emos em sugerir orientações, já que seria impossível desenvolver inteiramente tôdas as investigações, e que indicamos algumas em outra publicação.

1. Uma orientação será agrupar os vocábulos por idéia ou tema, estudando as sinonimias, ou as antinomias. O primeiro tema podia ser o do *Amor*, apesar da raridade relativa do verbo *amar*, por causa da alta frequência do nome *amor*, e de todos os outros que aparecem no mesmo campo semântico, quer como circunstante, quer como símbolo. Outro tema, o antagonismo entre *viver* e *morrer*, ou *querer* e *poder* (que aparece como uma constante em toda a poesia brasileira). Outro tema, a *terra* (1.^a posição pela frequência e pelo número de poetas), e por *terra* devemos entender a brasileira, afirmação muito interessante do telurismo. Há também temas contingentes (*pastor*, *gado*, etc.), e outros infinitesimais (*céu* no 9.^o lugar, *Deus* ausente; *saudade*, 20.^o lugar). É verdade que os adjetivos *doce* (2.^o lugar) x *triste* (3.^o lugar) sugerem um estado de alma semelhante, mas a beleza, a vida, as côres levam a vantagem nesta classe.

E podemos comparar com este vocabulário temático "comum" os vocábulos preferenciais primeiro, e depois os vocábulos próprios de frequência 1 quer de um período, quer de uma escola, quer de um poeta. Eis um meio eficiente para identificar um modo de expressão.

2. Uma segunda orientação é examinar uma ou várias listas individuais, avaliando a sua extensão até à frequência 2, observando o número, a percentagem ou até a posição dos Vocábulos Preferenciais Comuns que cada lista contém. Vamos exemplificar.

Números de V.P.C. encontrados nas listas individuais (Freq. sup. ou = 2).

	1	2	3	4	5	6	7	8	9	10	Tot.	Méd.
V	9	13	25	20	11	20	13	16	10	7	144	14,4
Q	5	1	11	4	0	3	5	6	8	4	47	4,7
S	6	14	32	19	11	12	22	22	20	7	165	16,5
AD	1	3	7	5	5	4	7	3	1	1	37	3,7
VQSAD	21	31	75	48	27	39	47	47	39	19	393	39,3

Percentagens destes V.P.C./n.º de Vocábulos preferenciais individuais:

V	75	62	50	74	73	83	68	76	55	53		65,4
Q	71	50	23	36	0	42	55	60	47	28		37,3
S	31	44	35	42	68	60	53	62	57	24		45,5
AD	100	75	43	100	55	50	65	100	50	33		59,6
VQSAD	53	52	35	54	64	66	58	68	54	32		51,0

Percentagens destes V. P. C./n.º V. P. I. por períodos e escolas:

	XVI-XVII	XVIII	XIX	Épicos	Líricos	Média
V	56	72	53	71	68	65,4
Q	30	46	28	45	35	37,3
S	36	75	27	47	62	45,5
AD	52	52	33	75	68	59,6
VQSAD	42,5	42,3	32,2	56	62	51,0

As tabelas indicam como se distingue um poeta, um século, ou uma época. Quanto a sua percentagem mais próxima da média, tanto mais integrado no grupo, ou vice-versa. Imediatamente, p. ex., José Bonifácio (n.º 10) aparece como diferente dos outros poetas "clássicos", porque ele tem percentagens sempre inferiores à média. Os do século XVIII são os que usam a maior proporção de V. P. C., mas entre eles há diferenças sensíveis entre épicos e líricos, principalmente no que toque aos substantivos. E para têmos a medida direta da originalidade de cada poeta, escola ou período, basta calcular o complemento da percentagem de V. P. C. a 100: o complemento indica a percentagem de V. P. I., quer dizer uma avaliação da individualização: José Bonifácio usa 76% de V. P. I. x 24% de V. P. C., quer dizer 3 de seus vocábulos preferenciais entre 4 lhe são próprios...

Voltando às listas individuais de vocábulos preferenciais, podemos progredir na análise, e extrair não os seus números ou valores, mas os vocábulos. Trataremos aqui apenas de uma classe (substantivos), em quatro listas (poetas n.º 1, 5, 8, 10). — Bento Teixeira emprega 19 substantivos com uma frequência igual ou superior a 2, entre os quais 6 (= 31%) VPC:

mar (1), *céu* (3), *rei* (3), *ar* (6), *fama* (6), *voz* (6); os substantivos próprios de B. T. são: valor, Albuquerque, ano, entrada, fúria, irmão, lajem, língua, licor, muro, obra, pedra, pôrto.

— Cláudio Manuel da Costa recorreu a somente 16 substantivos "frequentes" entre os quais 10 (= 62%) VPC: *tronco* (2), *monte* (3), *saudade* (3), *alma* (5), e ex-aequo *dia*, *gado*, *pastor*, *pranto*, *rio*, *rosto*; os pessoais são Nize (1), formosura (5), e ex-aequo memória, montanha, vale, vez.

— Tomás Antônio Gonzaga fornece 35 substantivos frequentes, entre os quais 19 (= 54%) VPC: *sorte* (2), *graça* (3), *ôlho* (5), *amor* (6), *céu* (6), *côr* (6), *monte* (6), *planta* (6), *gôsto* (6), *pastor* (6), *tempo* (6), *tronco* (6), *flor* (21), *mundo* (21), *terra* (21), *ano* (21), *mão* (21), *rosto* (21), *face* (21), *gado* (21); os próprios: Marília (1), estrêla, cabelo, corpo, pastôra, primavera, bruto, Deus, doença, estação, homem, passo, rebanho.

— José Bonifácio tem na sua lista 29 substantivos de frequência superior ou igual a 2, entre os quais apenas 7 (24%) VPC: *terra* (3), *dia* (4), *voz* (4), *luz* (11), e ex-aequo *glória*, *gente*, *sangue*; os pessoais são: filho (1), ouro, frente, pátria, poesia, vate, virtude, altar, Camenas, canto, corda, crime, deusa, filha, fome, gênio, Grecia, mãe, monstro, musa, som, verso.

A presença e frequência dos VPC indicam o grau de integração do poeta no grupo "Poetas brasileiros da fase colonial", e as dos vocábulos preferenciais próprios os individualizam. Claro que tal caracterização só seria completada pela consideração dos "hapax", trabalho que cada um poderá realizar quando publicado todo o *corpus*. Mas, mesmo assim, podemos ver que os dois poetas tomados como exemplos do século XVIII são os mais integrados no grupo "fase colonial" e os mais reveladores do genro poético do tempo, pastoral e idílico; que fica Bento Teixeira o mais "português" e José Bonifácio o idealizador da Independência. Será interessante continuar o estudo do substantivo em tôdas as listas, e das outras classes...

3. Não devemos esquecer a utilidade de outro método, procedendo pelo exame dos V. P. C. *ausentes* de uma ou várias listas individuais, porque a *exclusão* também é significativa. O confronto da lista de V. P. C. e das listas individuais alimentará duas observações em correlação: levantamento dos VPC excluídos e dos VPC escolhidos em cada caso. Para começar vamos indicar os V. P. C. ausentes das listas 1, 2 e 3:

V: correr, vir, cobrir, fugir, voar, deixar, chorar, acabar, viver, chegar, haver, entregar, tornar.

Q: triste, negro.

S: planta, chama, face, peito, rio, suspiro, tronco, monte, saudade, alma, gado, pastor, vento, sangue, morte, parte, sorte, ternura.

AD: ali, então, pouco, agora, já.

Isto quer dizer primeiro que os 3 poetas anteriores ao XVIII não usavam tais palavras com frequência e segundo que elas são, em alta frequência, privativas dos outros 7 poetas.

Mas, já que temos por mera subtração as palavras próprias das listas 1, 2, 3, vamos poder comparar esta nova lista à de José Bonifácio, que não tem vocábulo preferencial exclusivo, e de que ficam também ausentes:

V: *acabar*, *amar*, *cair*, *chamar*, *chegar*, *chorar*, *cobrir*, *conhecer*, *correr*, *dar*, *deixar*, *dizer*, *entregar*, *estar*, *ficar*, *fugir*, *gostar*, *gozar*, *haver*, *ir*, *levar*, *matar*, *morrer*, *mostrar*, *ouvir*, *parecer*, *perder*, *poder*, *pôr*, *querer*, *romper*, *saber*, *ser*, *tornar*, *vir*, *viver* *voar*.

Q: *alto*, *belo*, *bom*, *cruel*, *duro*, *escuro*, *feliz*, *formoso*, *frio*, *puro*, *triste*, *verde*, *vivo*.

S: *água*, *alma*, *amor*, *ano*, *ar*, *bôca*, *cana*, *céu*, *chama*, *cinza*, *côr*, *coração*, *esperança*, *face*, *fama*, *favor*, *flor*, *fruto*, *gado*, *gôsto*, *graça*, *imagem*, *mão*, *mar*, *milho*, *monte*, *morte*, *mundo*, *natureza*, *nome*, *ôlho*, *parte*, *pastor*, *pé* *peito*, *piedade*, *planta*, *pranto*, *rei*, *rio*, *rosa*, *rostro*, *saudade*, *senhor*, *sol*, *sorte*, *suspiro*, *tempo*, *ternura*, *tronco*, *vale*, *vento*, *vida*.

AD: *agora*, *ali*, *aqui*, *bem*, *já*, *mais*, *menos*, *pouco*, *sempre*, *tanto*, *tão*.

Podemos observar por um lado as ausências notáveis na lista 10, tais como *ser* (José Bonifácio é o *único* entre todos os poetas estudados a não usar *ser*, que é o verbo mais frequente) e *saudade* que floresce nas listas 4-9, etc.; e por outro lado fazer o levantamento das ausências encontradas tanto na lista 10 como nas 1, 2, 3 (itálicas).

Enfim não custa, senão um instante de atenção, obter a lista dos VPC próprios das listas 4-9, dos poetas do século XVIII: será o último arrolamento, diminuído de *cantar*, *celebrar*, *glória* e *vermelho*, ausentes a partir de Sta. Rita Durão até Silva Alvarenga, pelo menos como palavras de alta frequência.

Pelo mesmo processo podemos estabelecer comparação entre qualquer lista e outra, para verificar semelhança ou avaliar diferença no material lexical mais usado, como também entre a lista individual tomada como referência e a lista do vocabulário empregado pelo mesmo autor em outra obra, não arrolada aqui. Não há limite, senão o levantamento completo de toda a obra, e nesta altura, será dispensável a nossa amostragem.

4. A última comparação a oferecer aqui será a da lista alfabética de V. P. C. dos nossos 10 poetas XVI-XIX (grupo I, ou lista I) com a dos 15 poetas arrolados nos séculos XIX-XX, do Romantismo ao Neo-Modernismo (grupo II, ou lista II). Em apêndice o leitor encontrará a referida lista de vocábulos, com indicação da posição pela frequência global e do número de poetas, isto é, de listas individuais em que o vocábulo preferen-

cial aparece (O n.º total é 25 poetas). Eis o levantamento dos *Vocábulos Preferenciais Comuns encontrados nas listas I e II*:

Verbos			Substantivos			Adjetivos			Advérbios		
	P	NP		P	NP		P	NP		P	NP
Acabar	22	4	Água	18	7	Alto	7	4	Agora	5	9
Amar	15	5	Alma	1	9	Belo	2	8	Aqui	7	5
Cantar	13	8	Amor	2	12	Doce	1	7	Bem	6	7
Chamar	25	4	Céu	4	12	Escuro	7	5	Já	2	10
Chorar	18	4	Chama	22	4	Frio	5	6	Mais	1	16
Correr	21	4	Coração	20	5	Puro	9	4	Sempre	4	8
Dar	7	13	Dia	8	11	Triste	3	7	Tão	2	9
Deixar	18	6	Esperança	20	5	Verde	4	5			
Dizer	12	9	Flor	3	10	Vivo	5	6			
Estar	8	12	Luz	11	10						
Fazer	10	13	Mão	16	6						
Ficar	17	7	Mar	6	9						
Fugir	22	6	Morte	13	7						
Haver	9	13	Mundo	11	8						
Ir	4	15	Olho	9	8						
Morrer	11	9	Rosa	13	5						
Ouvir	24	5	Saudade	17	5						
Poder	15	7	Sol	13	8						
Querer	5	14	Terra	5	10						
Saber	14	9	Vento	18	4						
Sentir	18	10	Vida	10	8						
Ser	1	24	Voz	7	7						
Ter	2	18									
Ver	3	21									
Vir	6	13									
Viver	25	4									

Estes 26 verbos, 22 substantivos, 9 adjetivos e 7 advérbios comuns de 25 poetas, entre Bento Teixeira e João Cabral de Melo Neto, não vamos apresentá-los como a quintessência da poesia brasileira, o que nada significaria. A lista é apenas a indicação dos vocábulos mais frequentes nas amostras analisadas, e ela não tem finalidade senão fornecer o primeiro elemento de uma série de comparações, frutuosas, e um exemplo de que se pode fazer para conhecer melhor a língua poética do Brasil. As nossas observações serão breves. Apesar de mais numerosos, os verbos tem um papel talvez inferior ao dos nomes: devemos primeiro subtrair *ser* (1), *estar* (8), *ter* (2), *haver* (9), *ir* (4) e *vir* (6), cuja carga semântica é fraca; e no estudo dos temas já sugerido supra, amar vem muito longe atrás de amor, chorar no 18.º lugar, longe atrás de cantar... No campo dos substantivos, podemos notar a ausência de *Deus*, enquanto *alma* e *céu* expri-

Antes de comentar os dados, extraímos da tabela acima uma outra indicando os limites de oscilação das percentagens das classes:

<i>Percentagem de N</i>	V	Q	S	AD	P.N.	P.G.
Médio	16,2	10,4	22,9	3,8	53	47
Máximo	28	15	28	7	60	49
Mínimo	12	7	15	2	51	40
Dif. Máx.-Miní.	16	8	13	5	9	9
% Dif./Máx.	57	53	46	71	15	18

É a percentagem global das palavras nocionais que, mais uma vez, apresenta a maior estabilidade, ainda que com uma amplitude um pouco superior a $1/6$ do Máximo ela permite situar um ou outro poeta com uma nitidez suficiente. Mas a identificação do autor será inegavelmente mais clara pelo valor relativo da percentagem de cada uma das classes: as amplitudes representam entre $1/2$ e $3/4$ do Máximo. O gráfico n.º 3 ilustra perfeitamente as diferenças de posição: nunca dois poetas apresentam as mesmas percentagens nas quatro classes de nocionais. Eis realmente um excelente critério de definição da maneira de escrever própria de cada um.

2. COMENTÁRIOS

Bento Teixeira, por exemplo, usa poucos verbos e advérbios, muitos substantivos e proporcionalmente muito mais adjetivos: faltam-lhe o dinamismo e a densidade. Podemos notar que êle é um dos dois poetas a empregar mais adjetivos que verbos (o outro é José Bonifácio), e um dos três cujas percentagens de adjetivos são superiores à média da classe.

José Bonifácio, por sua conta, que tem a mais alta percentagem de adjetivos é também o campeão do substantivo, e usando a mesma fraca proporção de advérbios que B. T., emprega mais verbos: a sua língua é francamente descritiva, mais descritiva do que a de B. T., mas, mais densa (proporção de adjetivos/substantivos mais fraca) e mais dinâmica (percentagem de verbos superior à de B. T.). No seu caso devemos prestar atenção a três classes (S, Q, V), e no caso de Bento Teixeira a duas (S, Q).

Mas a nossa atenção é chamada pelas características de Claudio Manuel da Costa e Caldas Barbosa: êles têm a mesma percentagem mínima de adjetivos, provavelmente porque eles têm também as duas percentagens mais baixas de substantivos; mais ainda, êles têm as duas mais altas dos verbos; mais do que isso, êles constituem um caso, nunca encontrado em todo o resto das nossas pesquisas sobre a poesia brasileira — são os únicos a possuir uma percentagem de verbos superior à de substantivos, sendo a diferença de 0,2 e 13 pontos a favor dos verbos, o que parece incrível (seria uma proporção de diálogo de romance). No gráfico n.º 3 podemos ver, com efeito, uma saliência considerável da linha dos verbos e uma queda ainda mais profunda da linha dos substantivos, sendo ambas entrecruzadas.

A percentagem de substantivos de Caldas Barbosa cai abaixo da média dos verbos, e a sua percentagem de verbos, ultrapassando a média dos substantivos, atinge o mesmo nível do máximo de substantivos (28,2% x 28,7%). O caso C. B. provém de apenas 2 poemas, *Vou morrendo devagar* e *Doçura de amor* (duas cantigas), nas quais encontramos 110 verbos versus 34 substantivos, isto é 3,2 vezes mais, (os totais sendo 181 V x 102 S para os 4 poemas arrolados). E podíamos continuar pelo mesmo modo na análise da maneira de escrever de cada um dos nossos dez poetas.

3. DIFERENÇAS E QUOCIENTES

Entretanto vamos encontrar mais um critério no cálculo da diferença (em pontos) entre os seguintes binômios de percentagens: Substantivos-Verbos, Substantivos-Adjetivos, Verbos-Advérbios.

Dif. de Percent.	1	2	3	4	5	6	7	8	9	10	média
S — V	11	8	6	8	-0,2	-12	9	8	7	14	6,1
S — Q	9	16	9	14	12	8	13	19	10	3	10,9
S — AD	10	12	11	12	13	21	9	14	12	11	11,9

Nunca dois poetas têm as mesmas diferenças nas três séries, e o conjunto dos "gradientes" permite uma delineação da singularidade de cada autor. Mas será interessante tomar em consideração não somente a diferença como os valores de que se extrai.

Vamos pois calcular e comparar quocientes, isto é, números médios de substantivos por um verbo, de substantivos por um adjetivo, de verbos por um advérbio. Tais quocientes esclarecem a relação verbo-substantivo, base da frase, e as relações de determinação adjetivo/substantivo e advérbio/verbo. E recorremos mais uma vez a uma tabela:

	1	2	3	4	5	6	7	8	9	10	média
N.º S/1 V	1,8	1,6	1,4	1,6	0,9	0,5	1,6	1,5	1,4	2,0	1,4
N.º S/1 Q	1,7	2,0	1,8	2,4	2,8	2,1	2,4	3,7	1,7	1,8	2,2
N.º V/1 AD	6,0	4,4	3,4	5,3	3,0	4,0	3,0	5,1	7,7	6,4	4,9

Quanto mais baixo o primeiro número, tanto mais rico o texto em verbos, assim podemos avaliar a coesão e o dinamismo da frase: há uma nítida oposição entre José Bonifácio (10) e Caldas Barbosa (6), o texto de C. B. sendo 4 vezes mais dinâmico; por outro lado Bento Teixeira e José Bonifácio têm as maiores densidades nominais, enquanto as suas percentagens de substantivos em relação às ocorrências são diferentes (23,7 e 28,7%). Portanto não era inútil proceder a mais um cálculo, o que fica inscrito na primeira linha supra.

A segunda linha da tabela será talvez a mais expressiva, revelando a firmeza do nominal: quanto mais elevado o número, tanto mais sólido o texto, porque em vez de encontrar um adjetivo por menos de dois substantivos (1, 3, 9, 10), notamos um por três (8). Mas observamos que somente 4 dos 10 poetas ultrapassam a média, ela mesma baixa (principalmente em relação com os poetas dos séculos XIX e XX, cuja média é 3,2). Podemos afirmar que a poesia da fase colonial cultivava com grande favor a determinação pelo adjetivo qualificativo, o que criava ênfase senão inflação no setor nominal, tornado mais mole senão mais brando.

A terceira linha contém uma informação semelhante, na relação de determinação do verbo. Quanto maior o número, tanto menor o papel do advérbio. Da mesma maneira podemos identificar a maneira de um ou outro poeta pela comparação do seu quociente com a média ou o quociente de outro ou outros poetas. Os dois últimos quocientes têm a vantagem de nos informar sobre o binômio paradigma-sintagma.

4. OBSERVAÇÕES

Se voltarmos agora ao exame do gráfico n.º 3, será para fazê-lo horizontalmente (em vez de verticalmente como fizemos para ver as combinações individuais das 4 classes) e acompanhar as flutuações das percentagens de N próprios de cada classe.

As linhas que representam as médias globais ficam superpostas de baixo para cima na ordem advérbio, adjetivo, verbo e substantivo, isto é na ordem crescente. Começaremos pois pelo substantivo, que apresenta (5) e (6) a queda excepcional de que já falamos. Nos outros casos as oscilações são reduzidas, mas podemos observar a partir de (7) uma ligeira ascensão, que culmina com José Bonifácio (10). O adjetivo parece acompanhar de um poeta para outro, ainda que amortecidas, as tendências do substantivo, grosso modo paralelamente; entretanto existem movimentos em sentido contrário (nos casos 2, 3, 4, 8): a percentagem do adjetivo não é diretamente proporcional à do substantivo, é somente nos poemas de Basílio da Gama e de José Bonifácio que podemos encontrar a correlação (e J. B. bate o recorde em ambas as classes), e em todos os outros devemos examinar não tão somente as duas percentagens como a frequente oposição: muitos nomes versus muitos adjetivos (1-3-9). A relação advérbio-verbo é mais simples, a percentagem do advérbio sendo quase sempre alta ou baixa com a do verbo, acompanhando de longe o seu movimento. A percentagem de verbo fica mais frequentemente acima da de adjetivo e abaixo da de substantivo. Entretanto devemos examinar os casos (1) e (10), em que o verbo é menos frequente do que o adjetivo, e lembrar mais uma vez os casos (5) e (6), em que o verbo é mais frequente do que o substantivo (fato que nunca observamos em poesia).

Numa palavra o gráfico n.º 3 torna manifesta singularidade das taxas de constituição lexical em quatro poetas: Cláudio Manuel da Costa (5) e Caldas Barbosa (6) por um lado, e por outro Bento Teixeira (1) e José Bonifácio (10). Mas passamos assim o primeiro ao último dos poetas da fase colonial, numa orientação já diacrônica.

II) PONDERAÇÕES

1. PERÍODOS

Vamos portanto agrupar os poetas por períodos: XVI + XVII, XVIII, XIX, para verificar se houve diferenças de taxas, indicadoras de evolução, ou não. Procederemos como de costume, organizando uma tabela e traduzindo-a pelo gráfico n.º 4.

Porcentagem/N	XVI-XVII	XVIII	XIX	Média
Verbos	14	18,1	14	16,2
Adjetivos	11	8,6	15	10,4
Substantivos	22,6	21,6	28	22,9
Advérbios	3	4,3	2	3,8
Palavras nocionais	51,3	53,5	60	53,4

Tanto o gráfico como a tabela podem ser lidos vertical (para estudar um período) ou horizontalmente (para acompanhar a evolução de uma classe).

Os poetas do primeiro grupo realizam um vocabulário marcado pela sobriedade em verbo-advérbio, pela relativa abundância de adjetivo e a superioridade do substantivo sobre o verbo (+ 8,6 pontos). Os do segundo período (XVIII) caracterizam-se por taxas relativamente baixas de substantivo e adjetivo (cuja redução é mais intensa do que a de substantivo, quer dizer com uma força nominal), e taxas altas de verbo e advérbio. A superioridade da taxa de substantivo é de apenas 3,5 pontos sobre o verbo. O poeta que representa a transição da fase colonial para a Independência apresenta um vocabulário altamente descritivo, com taxa baixa de verbo e muito baixa de advérbio, taxas muito altas de substantivo e adjetivo: a superioridade da percentagem de substantivo sobre a de verbo é de 14 pontos, ou seja um valor igual ao da percentagem verbal, ela mesma ultrapassada pela do adjetivo.

Isto quer dizer que durante a fase colonial houve uma evolução cujas irregularidades provêm do Século XVIII, que marca um aumento das proporções de verbos e advérbios, uma redução das de substantivos e adjetivos em relação ao primeiro período. No fim da fase, se José Bonifácio é o exato

intérprete do período, os movimentos são os contrários: redução dos verbos e advérbios, aumento espetacular dos substantivos e adjetivos, a frase retomando em densidade o que ela perdeu em dinamismo em relação com os períodos anteriores. E no decurso do século XIX verificaremos a tendência do elemento nominal para a supremacia, ainda mais nítida a partir do Simbolismo.

É também o que confirma a tabela seguinte, que toma em consideração o quociente dos valores dos binômios fundamentais:

	S/V	V/AD	S/Q
XVI-XVII	1,6	4,6	1,8
XVIII	1,2	4,8	2,5
XIX	2,0	6,4	1,8
Média	1,4	4,9	2,2

Ela traduz a "febre" verbal do XVIII, a ascensão do nominal e a menor adjetivação no primeiro quartel do XIX.

2. ESCOLAS OU GÊNEROS

Sendo os mais numerosos, os poetas do século XVIII merecem uma última análise. Acompanhando opinião da maioria dos historiadores da literatura brasileira, podemos distinguir duas escolas, a dos épicos (I) p. ex. Sta Rita Durão (4) e J. Basílio da Gama (7) e a dos líricos (II) p. ex. Cláudio Manuel da Costa (5), Tomás Antônio Gonzaga (8) e Silva Alvarenga (9)), acrescentando eventualmente (III) Caldas Barbosa (6), geralmente considerado inferior aos "Mineiros". E procederemos como acima, com duas tabelas:

			Percentagens		sobre	N
			V	Q	S	AD
I)	4 + 7	épicos	14,5	9	23,5	4
II)	5 + 8 + 9	líricos	17,6	9	22,9	3,6
III)	5 + 8 + 9 + 6 -		20,2	8,5	20,9	4,5
	média		16,8	10,0	22,9	3,7

			Quocientes		
			S/V	S/Q	V/AD
I)	épicos		1,6	2,4	4,1
II)	líricos		1,26	2,7	5,2
III)	- (6 + 6)		1,07	3,4	4,9
	média		2,0	1,8	6,4

Será o lirismo o agente do aumento verbal? As diferenças entre as escolas I e II parecem indicar, e o papel de Caldas Barbosa sublinhar a idéia.

Pelo contrário, o gênero épico recorre mais ao nominal, isto é ao substantivo, reduzindo o emprêgo do verbo, o que pode surpreender. E a segunda tabela confirma o papel de Caldas Barbosa como símbolo da embriaguez verbal, com redução da proporção de substantivos e de adjetivos; mas ela revela na linha II que os outros líricos andam com moderação sôbre o mesmo caminho, e outro paradoxo é em corolário a sua discreção no uso do adjetivo. É cedo demais para tirar conclusões, porque devemos antes multiplicar os levantamentos, mas no estado atual da pesquisa podemos chamar a atenção dos leitores sôbre a correlação entre a escola ou o gênero e a frequência do elemento verbal ou nominal, tomada como hipótese de trabalho, a ser verificada ou confirmada à luz das observações quantitativas que eles aceitariam consignar.

CONCLUSÃO

Para concluir êste artigo, extenso demais e, entretanto, fragmentário, agradecendo a benevolência ao leitor, desejamos reafirmar o quanto cremos na eficiência das pesquisas quantitativas e seriais, com que os "literatos" ainda são meros acostumados do que os historiadores, que já renovaram vastos setores da sua disciplina.

Os índices, só os índices, apontam os vocábulos usados pelos autores, revelam as frequências de emprêgo e fundamentam qualquer trabalho estatístico. O ideal seria possuir o índice da totalidade das obras de um autor, de todos os autores... Mas é inacessível. Contentemo-nos em analisar "amostras" extraídas de um certo número de autores e de obras, que tentaremos sempre aumentar.

Num terreno em que sentimos pessoalmente a cruel ausência de dados concretos e exatos, tentamos abrir uma picada e esboçar um mapa sumário, mas que pode ter alguma utilidade, assim como fotografias batidas de diversos ângulos. Demos a maior atenção à listas de vocábulos mais frequentes, focalizando as proporções que representam em relação com o total ou as classes morfológicas. E tentamos mostrar como, pela combinação das listas e dos dados puramente numéricos, se podiam caracterizar cada poeta pelas suas escolhas lexicais, ou uma escola ou um período pelas preferências manifestadas. A seguir, retomamos o exame numérico das ocorrências, isto é das unidades-de-texto, isto é das frequências de emprêgo, o que é a base de um estudo estilístico quantitativo em que pretendemos prosseguir para definir um autor, uma escola, um período e uma tipologia dos gêneros.

E lembraremos que temos à disposição de eventuais interessados em lexicologia, semântica ou estilística, dezenas de milhares de fichas (uma para cada ocorrência) com a localização na obra e a sequência de que o

vocábulo faz parte: — "As suas ordens!", ou melhor o nosso Centro de Lexicologia de Toulouse lhes diz: — "A casa é sua!".

JEAN ROCHE

Professor da Universidade de
Toulouse-Le Mirail
Sócio correspondente da
Academia Brasileira de Letras e da
Academia Internacional de Cultura
Portuguesa

APÊNDICE

1. Relação dos poemas analisados (pág. 29).
2. Lista dos vocábulos nacionais de frequência 1:
 - Valores.
 - Percentagens de N.
 - Percentagens de V. (pág. 30).
3. Lista dos Vocábulos Preferenciais Comuns (V.P.C.) dos 15 poetas do Romantismo até o Neomodernismo. (pág. 31).
4. Gráficos
 - N.º 1 — Vocábulos nacionais por classe morfológicas, em % de V — (Poetas) (pág. 32).
 - N.º 2 — " " " — (Períodos)
 - N.º 3 — Vocábulos nacionais por classes morfológicas, em % de N (Poetas) (pág. 34)
 - N.º 4 — " " " — (Períodos) (pág. 35).

LISTA DAS POESIAS ARROLADAS

1. BENTO TEIXEIRA.
Fragmentos da "Prosopopéia" — Proposição
Descrição do Recife de Pernambuco
2. GREGÓRIO DE MATOS.
Implorando de Cristo, um pecador contrito, perdão...
Pecador contrito aos pés de Cristo crucificado
A procissão de cinza, feita em Pernambuco
Aos Srs. Governadores do mundo em seco da Cidade da Bahia...
Vendo o autor a primeira vez a D. Angela...
Expressões amorosas a ãa dama a quem queria
Saudosamente sentido na ausência da dama...
Lamenta ver-se no tal degrêdo em terra tão remota...
3. MANUEL BOTELHO DE OLIVEIRA.
Vendo a Anarda depõe o sentimento
Ponderação do rosto e olhos de Anarda
Esperanças sem lôgro
A Ilha de Maré

4. JOSÉ DE SANTA RITA DURÃO.
Fragmentos do "Caramuru" (Canto IV)
Fragmentos do "Caramuru" (Canto VII)
5. CLAUDIO MANUEL DA COSTA.
Pastôres, que levais ao monte o gado
Brandas ribeiras, quanto estou contente
Onde estou! Este sítio desconheço
Este é o rio, a montanha é esta
Nize? Nize? onde estás? Aonde espera
Quando, formosa Nize, dividido
6. DOMINGOS CALDAS BARBOSA.
Vou morrendo devagar
Sem acabar de morrer
Doçura de amor
A ternura brasileira
7. JOSÉ BASILIO DA GAMA.
Fragmentos de "O Uruguai" (Do canto terceiro)
8. TOMAS ANTONIO GONZAGA.
Da primeira parte — Marília de Dirceu — Lira I
Da segunda parte — Marília de Dirceu — Lira III
Da segunda parte — Marília de Dirceu — Lira IV
9. MANOEL INACIO DA SILVA ALVARENGA.
O cajueiro — Rondó III
O beija-flor — Rondó VII
Madrigal III, VII, XIII, XV, XXXIV, LVI
10. JOSÉ BONIFACIO DE ANDRADA E SILVA.
Ode à poesia (em 1785)

Nota sobre os vocábulos nocionais de frequência 1.

Números	1	2	3	4	5	6	7	8	9	10	Total	=%N
V	29	50	83	89	43	34	76	54	62	59	579	6,3
Q	67	52	116	79	36	19	64	31	51	74	589	6,4
S	89	105	209	156	70	35	80	109	77	125	1055	11,4
AD	7	11	10	12	12	9	15	16	8	6	106	1,1
Voc. Nocion.	192	218	418	336	161	97	235	210	198	264	2329	25,3
Tot. Voc. F. 1	222	240	455	373	174	115	262	236	215	282	2574	28

Percentagens / N (Total das Ocorrências)

	1	2	3	4	5	6	7	8	9	10	
V	4,9	6,6	3,7	8,2	7,6	5,3	8,8	5,7	7,1	8,6	
Q	11,5	6,9	5,2	7,3	6,4	3,0	7,4	3,3	5,8	10,8	
S	15,3	13,9	9,4	14,3	12,4	5,4	9,3	11,5	8,8	18,3	
AD	1,2	1,4	0,4	1,1	2,1	1,4	1,7	1,7	0,9	0,9	
V. Noc.	33	29	18,8	31	28,5	15,1	27,4	22,3	22,6	38,7	

Percentagens / V (Total dos Vocábulos)

V	9,9	14,7	11,5	17,9	17,1	16,6	20,8	15,8	19,6	16,2
Q	22,9	15,3	16,1	15,9	14,2	9,2	17,5	9,1	16,1	20,3
S	30,5	30,9	29,0	31,3	27,8	17,1	21,9	31,9	24,4	34,3
AD	2,4	3,2	1,3	2,4	4,8	4,4	4,1	4,7	2,5	1,6
V. Noc.	65,7	64,3	58,0	67,6	64,1	47,3	64,4	61,4	62,6	72,5
Tot. Voc. F. 1	76,2	70,8	63,2	75,0	69,3	57,0	71,8	69,0	68,0	77,5

Percentagens / N por períodos:

	XVI-XVII	XVIII	XIX
V	5,0	7,1	8,6
Q	7,8	5,5	10,8
S	12,8	8,6	18,3
AD	1,0	1,4	0,9
Voc. Noc.	26,6	22,6	38,7
Tot. Voc. F. 1.	30,2	27,2	41,3

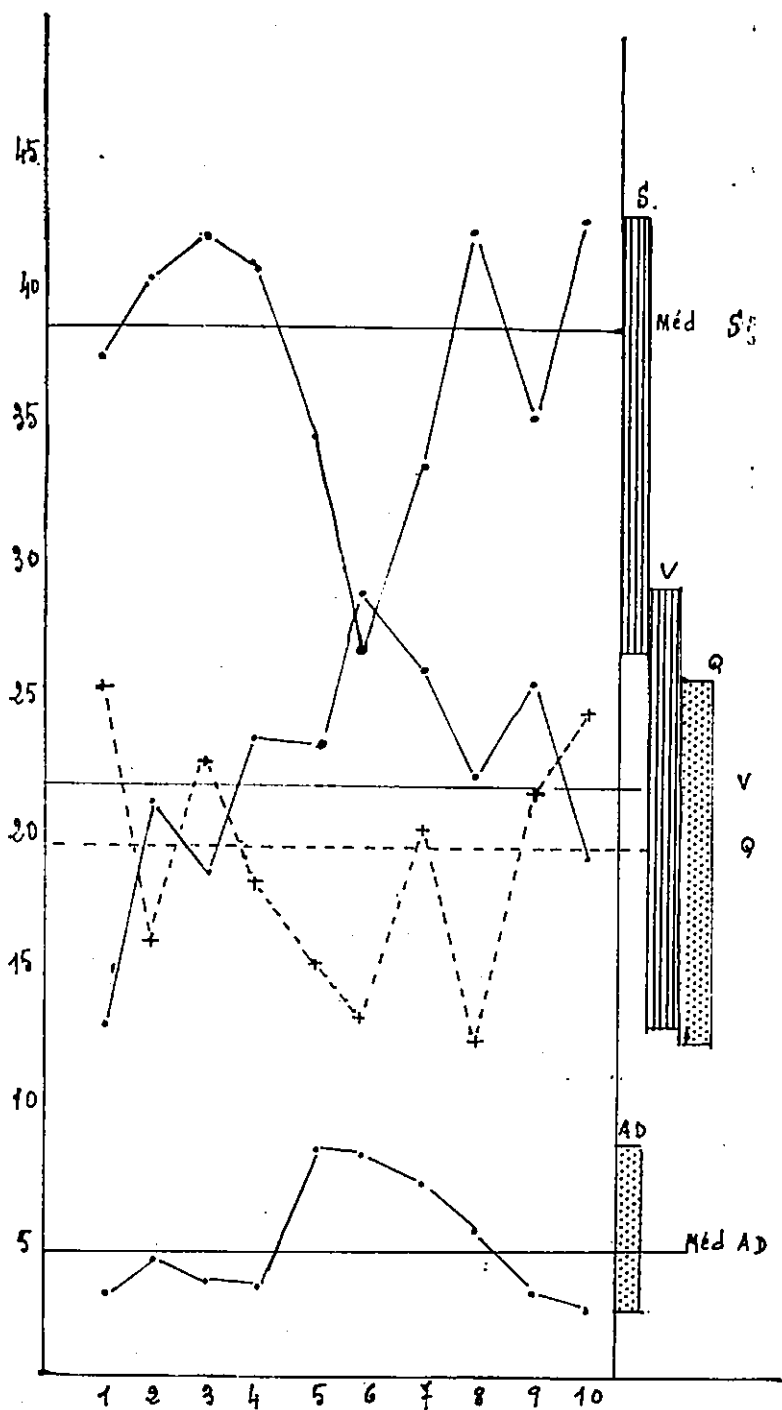
Percentagens / V por períodos:

V	12,0	17,9	16,2
Q	18,1	13,7	20,3
S	30,1	25,7	34,3
AD	2,3	3,8	1,6
Voc. Noc.	61,3	62,6	72,5
Tot. Voc. F. 1.	70,0	68,3	77,5

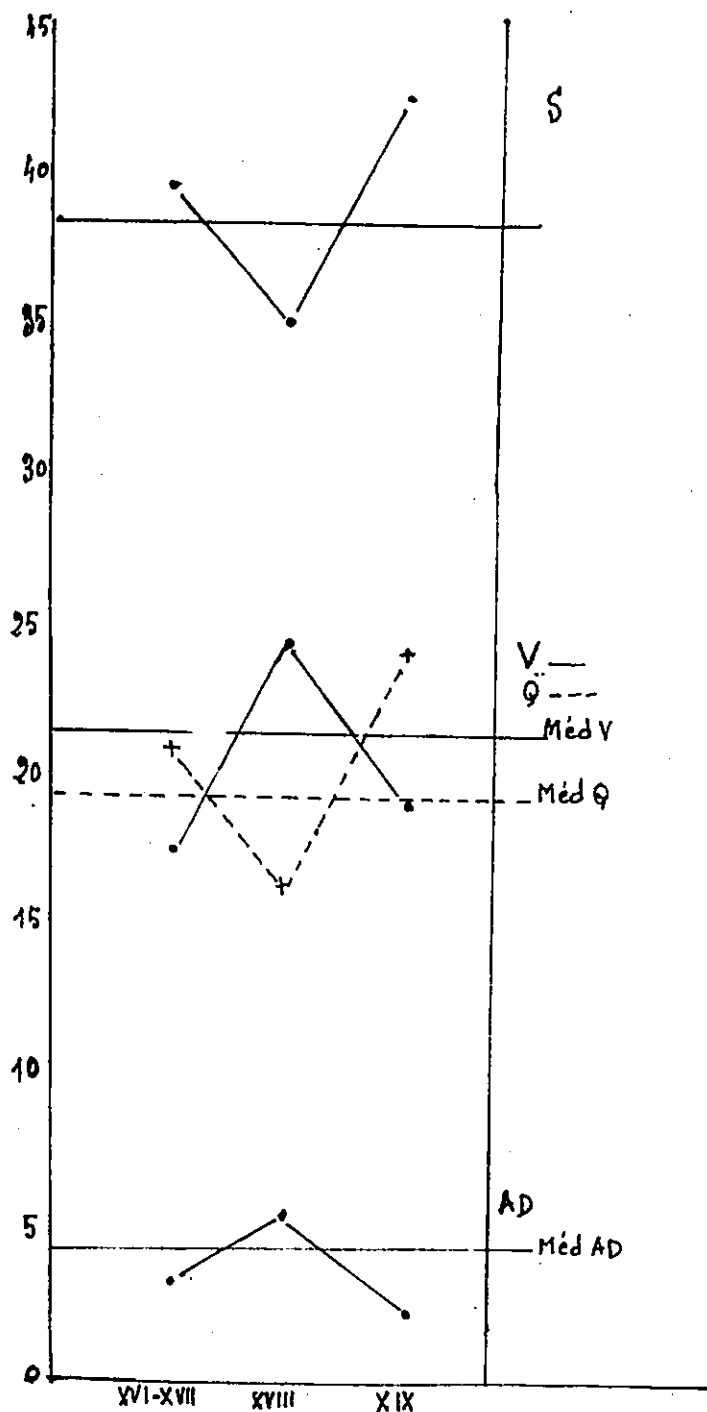
VOCABULOS PREFERENCIAIS COMUNS DOS 15 POETAS XIX-XX

Verbos			Substantivos			Adjetivos			Adverbios		
	P	NP		P	NP		P	NP		P	NP
abrir	25	3	água	27	4	aflito	26	2	agora	3	7
acabar	38	2	alma	1	7	alegre	18	2	ainda	6	6
amar	12	5	amigo	42	2	alto	18	2	aqui	12	2
aprender	41	2	amor	3	8	azul	8	2	assim	13	2
bater	28	2	anjo	29	2	belo	4	4	bem	8	4
beijar	23	2	aroma	40	3	branco	3	3	hoje	11	4
brilhar	46	2	árvore	20	2	cheio	5	3	já	2	6
buscar	41	2	asa	29	3	doce	1	2	longe	6	4
cantar	11	6	astro	48	2	escuro	18	2	mais	1	10
chamar	38	2	BELJO	20	3	estranho	15	2	muito	10	3
chorar	18	4	brisa	36	2	feito	8	2	sempre	8	4
contar	32	3	caminho	45	2	frio	7	4	só	5	4
correr	27	2	canto	13	3	infinito	13	2	tão	4	4

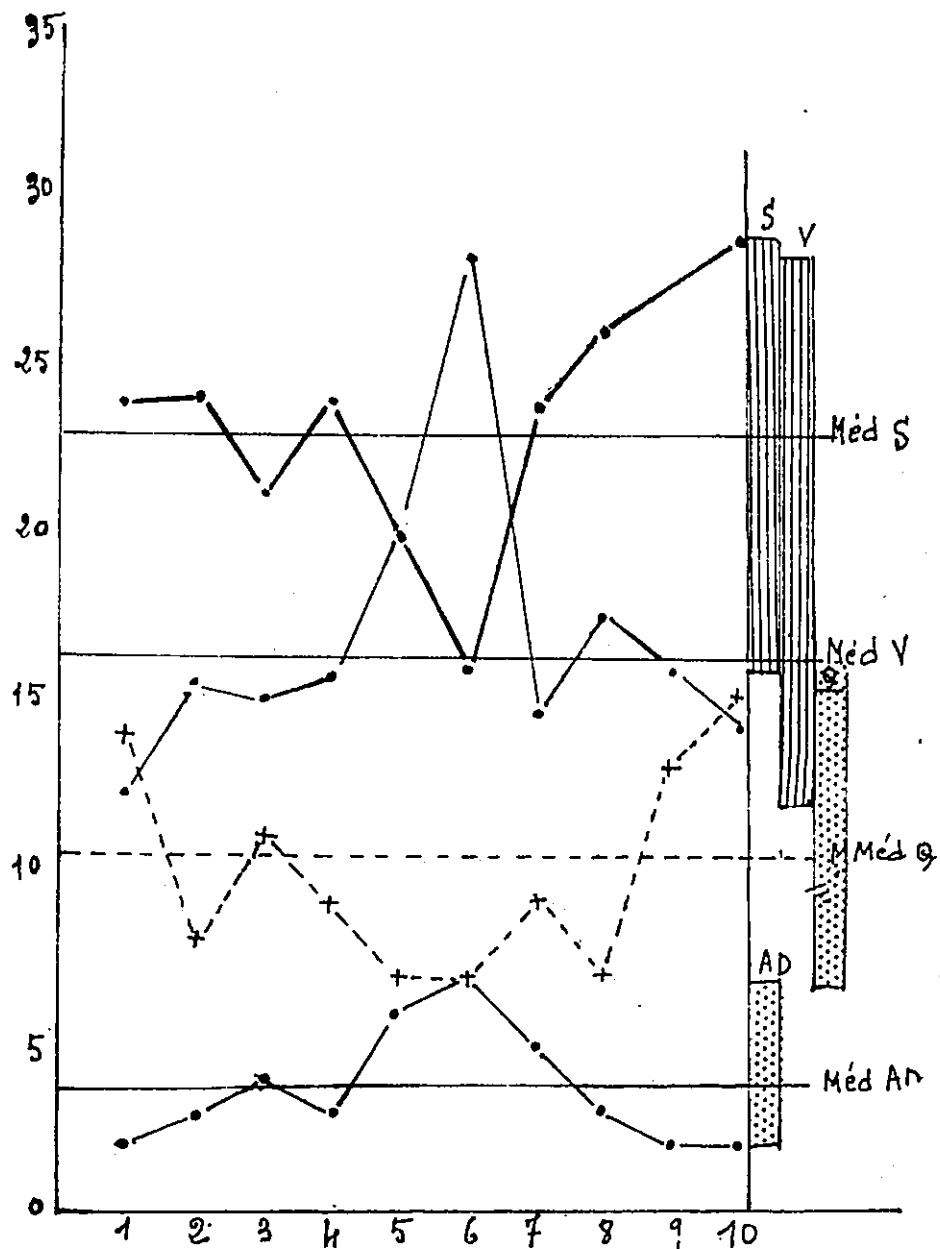
dar	9	7 carne	48	2 largo	18	2
deixar	25	4 catedral	45	2 Morto	18	2
desaparecer	41	2 céu	4	7 novo	6	2
descer	38	2 CHAMA	48	2 nu	15	2
dizer	10	5 cidade	42	2 pálido	18	2
dormir	20	3 coisa	36	3 profundo	8	2
escutar	32	3 coração	36	3 puro	18	2
estar	8	7 corpo	29	3 santo	8	2
falar	41	2 criança	48	2 sonoro	18	2
fazer	12	6 Deus	20	3 triste	1	4
ficar	17	5 dia	9	6 velho	15	2
fugir	32	3 esperança	42	3 verde	8	2
gemer	32	2 estrela	29	2 vivo	13	3
guardar	21	2 flor	2	6		
haver	6	11 gênio	48	2		
ir	4	10 homem	29	3		
lembrar	41	2 lua	45	2		
morrer	12	6 luz	16	5		
olhar	28	3 manhã	25	2		
ouvir	28	3 mão	20	3		
passar	19	3 mar	7	6		
poder	28	2 morte	15	5		
PROCURAR	36	2 mulher	20	3		
querer	5	8 mundo	16	4		
saber	12	5 noite	8	7		
sentir	21	6 olho	18	3		
ser	1	15 ouro	19	2		
sonhar	16	5 poesia	36	2		
subir	23	3 rosa	11	3		
ter	2	11 rua	40	2		
ver	7	11 saudade	29	2		
vir	3	9 sol	11	5		
viver	36	2 sonho	14	4		
		tarde	25	3		
		terra	5	5		
		vento	27	2		
		verso	29	2		
		vida	9	5		
		voz	6	4		



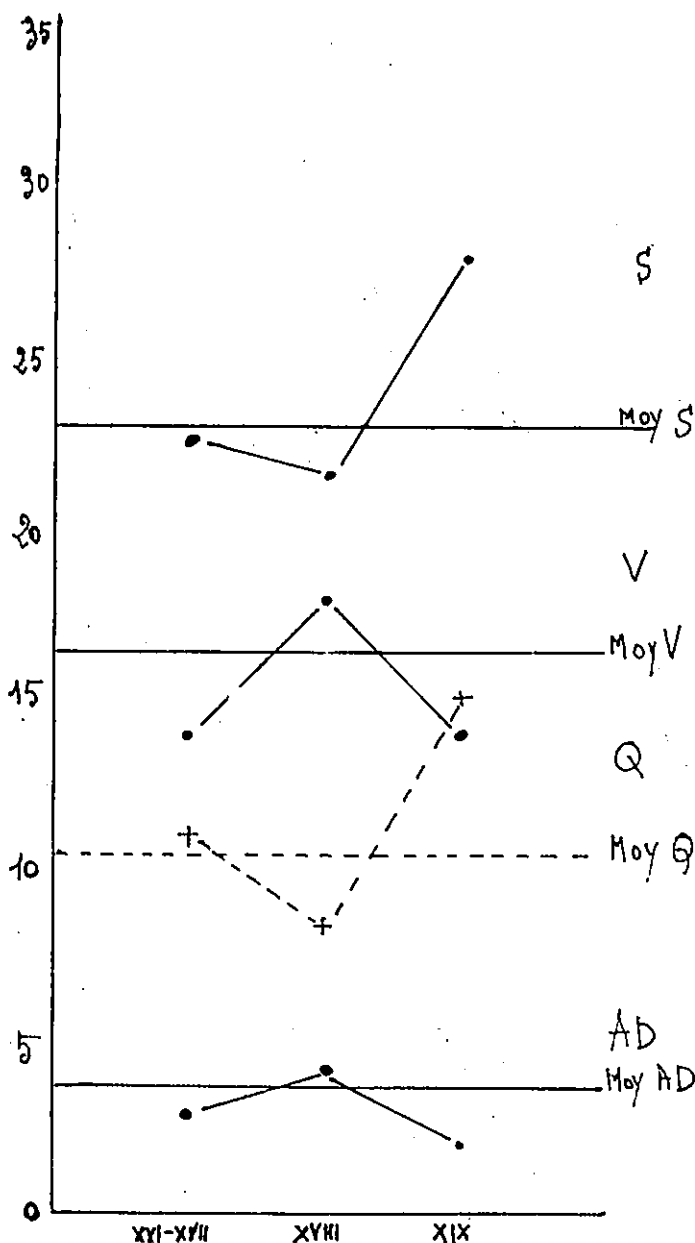
Vocabúlos nacionais por classes morfológicas, em % de V (Poetas).



Vocabulos nacionais por classes morfológicas, em % de V (Períodos).



Vocábulos notacionais por classes morfológicas, em % de N (Poetas).



Vocábulos nacionais por classes morfológicas, em % de N (Períodos).