

# Comportamento bibliométrico da língua portuguesa, como veículo de representação da informação

Elza Lima e Silva Maia

Instituto Brasileiro de Bibliografia e Documentação  
Escola de Biblioteconomia e Documentação da FEFIEG

## SINOPSE

As pesquisas em Ciência da Informação, orientadas para a Lingüística, apresentam interesse particular para a Bibliometria, que indica o tratamento quantitativo da comunicação escrita; entre as múltiplas aplicações da Bibliometria, encontram-se as pesquisas referentes à freqüência de palavras em determinado texto.

Zipf, baseado em uma relação empírica observada por Estoup, formulou leis estabelecendo relações entre a ordem de série de uma palavra em ordem de freqüência, e a freqüência de seu aparecimento em um texto "suficientemente longo".

Booth, estudou e desenvolveu a segunda lei de Zipf, menos conhecida do que a primeira, e referente a palavras de baixa freqüência.

Estudando as duas leis de Zipf, W. Goffman derivou uma fórmula de transição, que indica, aproximadamente, as ordens de série nas quais se devem encontrar as palavras significativas de um texto em língua inglesa.

As seguintes hipóteses foram levantadas: a primeira lei de Zipf é válida para a língua portuguesa; sua segunda lei e a fórmula de transição de Goffman são, também, válidas para a língua portuguesa, ou podem ser adaptadas a este idioma. Foram escolhidos três Textos, no campo da Bibliografia: Texto A, com 6395 palavras, sendo 1791 palavras diferentes; Texto B, com 2210 palavras, sendo 839 palavras diferentes; e Texto C, com 1290 palavras, sendo 535 palavras diferentes.

Nestes textos, foram aplicadas as leis de Zipf e a fórmula de transição de Goffman, calculando-se suas respectivas fórmulas, e tentando-se obter variações da constante 2, válida para a língua inglesa, para maior precisão dos resultados.

As conclusões foram as seguintes: a primeira lei de Zipf é válida para a língua portuguesa; sua

segunda lei, tal como formulada, parece não ser válida para a língua portuguesa, mas pode ser adaptada a esse idioma; a fórmula de transição de Goffman, também, poderá sofrer uma adaptação para a língua portuguesa.

## 0 – INTRODUÇÃO

A palavra *informação* é usada, na linguagem cotidiana, de diferentes maneiras. Falamos de informação "útil", "valiosa", "real", "fidedigna", "precisa", "verdadeira", "relevante"; mas nenhuma dessas expressões ocorre na teoria matemática da comunicação, que descreve a informação tão somente como uma probabilidade estatística de sinais provindos de uma determinada fonte.

Apesar de todos sabermos o que é informação, não há maneira de defini-la, exceto através de suas características e propriedades, estudadas pela Ciência da Informação.

Dentre as propriedades, a representação da informação é um dos campos mais investigados, em consequência de uma premissa antiga, na qual se identifica o bom funcionamento de um sistema de informação com a representação adequada desta informação — e isto acontece freqüentemente. Assim, os diferentes tipos de representação se refletem nas bibliotecas, através do sistema de

Dissertação apresentada ao Instituto Brasileiro de Bibliografia e Documentação/Universidade Federal do Rio de Janeiro para obtenção do grau de Mestre em Biblioteconomia e Documentação. *Orientador*: Professor Dr. Tefko Saracevic, Case Western Reserve University, Cleveland, Ohio

classificação adotado, dos cabeçalhos de assunto e da catalogação; nos sistemas de recuperação da informação, através das linguagens de indexação; nos sistemas que utilizam o computador, através das linguagens de programação e sua codificação; nos sistemas de transmissão de sinais, através de sua codificação.

A representação da informação tem, na linguagem, o seu principal veículo, estabelecendo uma relação entre Ciência da Informação e Lingüística: enquanto a primeira se interessa por todos os aspectos da comunicação da informação, inclusive tendo a linguagem como veículo básico para tanto, a segunda estuda a linguagem como um sistema para comunicar a informação.

Na prática, esta relação não tem sido devidamente explorada: primeiro, pelo fato da Lingüística ter muito pouco a oferecer na área da Semântica — aspecto da linguagem que mais interessa aos cientistas da informação; segundo pelo aparecimento, no início deste século, da indexação, que, como linguagem processada manualmente, não exigia o conhecimento *explícito* dos sistemas de comunicação da informação.

As pesquisas em Ciência da Informação, orientadas para a Lingüística, apresentam interesse particular para a Bibliometria — conjunto de leis que constituem uma das bases teóricas da Ciência da Informação.

Bibliometria, termo consagrado por Alan Pritchard<sup>1</sup>, em 1969, indica o tratamento quantitativo da

comunicação escrita, em substituição a "estatística bibliográfica", expressão anteriormente usada por E. W. Hulme, em 1922, mas de dúbia interpretação e pouco aceita pelos cientistas da informação. A Bibliometria tem como objetivos:

- 1) esclarecer os processos de comunicação escrita, da natureza e evolução de uma disciplina (na medida em que esta se manifesta através do conhecimento registrado), por intermédio da contagem e análise de suas diversas facetas;
- 2) reunir e interpretar dados estatísticos relativos a livros e periódicos, para demonstrar a evolução histórica, determinar a utilização destes materiais em âmbito nacional e internacional, bem como verificar o uso específico de livros e periódicos em bibliotecas, centros de documentação, etc.

Estas finalidades, particularmente a segunda, descrevem apropriadamente as múltiplas aplicações da Bibliometria, tais como estudos relacionados a citações (Solla Price, Kessler, Garfield, etc.), à distribuição da literatura periódica e respectiva utilização (Bradford, Urquhart, Brookes, etc.), e às pesquisas referentes à frequência de palavras em determinado texto (Zipf, Mandelbrot, Parker-Rhodes e Joyce, Booth, etc.)<sup>2</sup>.

Estas pesquisas levaram ao estabelecimento de

diferentes leis, posteriormente reformuladas e adaptadas por outros cientistas da informação.

Zipf<sup>3,4</sup>, baseado em uma relação empírica observada por S. B. Estoup em 1916, formulou em 1935 e 1949 leis estabelecendo relações entre a ordem de série (rank) de uma palavra em ordem de frequência, e a frequência de seu aparecimento em um texto "suficientemente longo" (long enough text). Mandelbrot<sup>5</sup>, baseado em Zipf e na Teoria da Informação, em 1953 e 1954, introduziu uma outra variável "esforço" ou "custo" de palavras, definida como "espaço de tempo", resultante da transmissão de palavras como uma seqüência de combinações de letras ou fonemas, separadas por espaços ou pausas. Considerando que o objetivo da linguagem é permitir a transmissão da maior variedade de sinais (i.e., informação no sentido atribuído pela Teoria da Informação), no menor espaço de tempo, Mandelbrot empregou técnicas para combinar códigos com o uso da mensagem.

Parker-Rhodes e Joyce<sup>6</sup> (1956), a partir de Zipf, estabeleceram o número de palavras, de um determinado vocabulário, que ocorre com determinada frequência, dentro de um texto suficientemente longo.

Fairthorne<sup>7</sup>, de 1953 a 1958, aplicou o princípio do "menor esforço" a vários problemas de documentação e recuperação da informação, a partir de um determinado código (linguagem).

Booth<sup>8</sup>, em 1967, estudou e desenvolveu a segunda lei de Zipf, menos conhecida do que a primeira, e referente a palavras de baixa frequência. Em 1969, determinou, ainda a partir de Zipf, a disposição de livros nas estantes, baseado na frequência de empréstimo<sup>9</sup>.

Além das leis relativas à alta e baixa frequência de palavras em um texto, Zipf coletou e estudou diversos aspectos de fala e da escrita, e foi em relação ao comportamento humano que ele formulou seu elogiado mas também atacado "princípio do menor esforço", isto é, a tendência de se obter o máximo com o mínimo de esforço possível<sup>10</sup>.

A primeira lei de Zipf, aplicada inicialmente ao campo da Lingüística e referente às palavras de alta frequência em um texto, assim se enuncia: "o número de vezes que cada palavra aparece em um texto é contado. As palavras são, então, ordenadas em uma tabela, na qual a primeira palavra é a mais freqüente, a segunda palavra a segunda mais freqüente e assim sucessivamente.

A ordem de qualquer palavra na tabela é chamada *ordem de série* ( $r$ ), e o número de vezes que ela aparece, *freqüência* ( $f$ ), de tal forma que  $r \times f = C$ , onde  $C$  é uma constante para qualquer texto"<sup>8</sup>.

Esta lei ilustra o "princípio do menor esforço". Zipf observou, ainda, que uma vez estabelecida a frequência de palavras de um texto, as palavras de baixa frequência (i. e., de alta ordem de série) ocorrem de tal forma que várias palavras têm a mesma frequência. Esta segunda lei de Zipf foi

modificada e expressa matematicamente por Booth<sup>8</sup>, para a língua inglesa, da seguinte forma:

$$\frac{I_n}{I_1} = \frac{2}{n(n+1)}$$

onde  $I_n$  é o número de palavras que ocorre  $n$  vezes, e  $I_1$ , qualquer palavra que ocorra uma única vez. O número 2 representa uma constante válida para a língua inglesa, segundo Booth; validade que pode ou não se manter para outras línguas. Estudando as duas leis de Zipf – a de alta e a de baixa frequência – W. Goffman\* derivou uma fórmula de transição, cujo resultado final é uma

Ordem de série T, que indica, aproximadamente, as ordens de série nas quais se devem encontrar as palavras significativas de um texto em língua inglesa:

$$T = \frac{-1 + \sqrt{1 + 8 I_1}}{2}$$

onde:

T – ordem de série qualquer

8 – constante fixa para qualquer idioma

I – frequência 1, correspondente à ordem de

1 série mais alta

2 – constante válida para a língua inglesa

A Fórmula de transição de Goffman, bem como a primeira lei de Zipf, e sua segunda lei modificada por Booth, foram aplicadas, neste estudo, à língua portuguesa — como veículo de representação da informação — excluindo-se toda e qualquer outra implicação.

Assim, três textos diferentes foram escolhidos, para aplicação das três leis, tendo sido procuradas as fórmulas ótimas para a língua portuguesa.

Um primeiro estudo nesta área foi apresentado pela autora como trabalho final à disciplina de Processamento de Dados na Documentação (Information Retrieval), durante o Curso de Mestrado em Ciência da Informação, IBBD/UFRJ, 1971/72. O presente estudo, após esta introdução, compreende: a hipótese, o material — apresentação dos textos escolhidos — o método empregado para aplicação das leis, os resultados obtidos, conclusões e citações bibliográficas.

\* Inédito. Comunicação pessoal do Prof. Tefko Saracevic.

Agradecimentos muito especiais devem ser apresentados ao Professor Tefko Saracevic, pela sua valiosa e entusiástica orientação; às Professoras Hagar Espanha Gomes, Presidente do IBBD, Alice Príncipe Barbosa, Diretora do Serviço de Intercâmbio de Catalogação, do mesmo Instituto, e Gilda Maria Braga, pelo apoio e estímulo, sem os quais este trabalho não seria realizado; a todos os que direta ou indiretamente contribuíram para esta realização, minha gratidão.

A meus Pais, meu marido e minha filha, dedico este estudo, na falta de palavras que possam expressar meu reconhecimento.

## 1 – HIPÓTESE

Considerando-se que as duas leis de Zipf e a fórmula de transição de Goffman foram observadas e calculadas tomando-se como base a língua inglesa, são levantadas as seguintes hipóteses:

- a) a primeira lei de Zipf é válida para a língua portuguesa;
- b) a segunda lei de Zipf é válida para a língua portuguesa, ou pode ser adaptada a este idioma;
- c) a fórmula de transição de Goffman é válida para a língua portuguesa, ou pode ser adaptada a este idioma.

## 2 – MATERIAL

Foram escolhidos três Textos, no campo da Bibliografia, de autores de reconhecida capacidade profissional neste campo.

O primeiro, de Edson Nery da Fonseca, "Desenvolvimento da Biblioteconomia e da Bibliografia no Brasil", publicado na Revista do Livro, em 1957, p. 95-124 (Texto A).

O segundo, de Hagar Espanha Gomes, "Ensino de Técnicas Bibliográficas nos Cursos de Ciências Agrícolas", apresentado no Seminário sobre Bibliotecas Agrícolas, organizado pelo IBBD, no Rio de Janeiro, 1963, 10 f. (Texto B).

O terceiro, de Laura Maia de Figueiredo, "Bibliografia Corrente Brasileira", apresentado no 3. Congresso Brasileiro de Biblioteconomia e Documentação, em Curitiba, 1961, 6 f. (Texto C).

As características numéricas destes Textos são dadas a seguir:

Texto A: 6395 palavras, sendo 1971 palavras diferentes.

Texto B: 2210 palavras, sendo 839 palavras diferentes.

Texto C: 1290 palavras, sendo 535 palavras diferentes.

## 3 – MÉTODO

A contagem de palavras obedeceu às normas e processos abaixo discriminados.

### 3.1 – Normas

As normas empregadas foram idênticas para os três Textos.

#### 3.1.1 – Palavras compostas consideradas como um todo;

As palavras compostas foram consideradas como um todo (i.e., uma única palavra):

- as ligadas por hífen, constituindo uma unidade semântica. Ex.: catálogo-dicionário;
- as de formas verbais, com pronomes enclíticos. Ex.: dando-lhe, exercê-la, etc.;
- as de formas verbais com o pronome *se* empregado como complemento objetivo ou terminativo. Ex.: realizou-se, deve-se, reorganizaram-se, etc.;
- os vocábulo formados por prefixos, representando formas objetivas. Ex.: norte-americano, tupi-guarani, etc.

#### 3.1.1.2 – Foram ainda considerados como um todo:

- os nomes de repartições, agremiações, estabelecimentos públicos ou particulares, nacionais ou estrangeiros (por extenso, parcialmente, ou em siglas). Ex.: Instituto Brasileiro de Bibliografia e Documentação, IBBD, o Instituto (referindo-se ao IBBD), Library of Congress, LC, etc.;
- os nomes de atos públicos, tais como exposições, congressos, simpósios, etc. Ex.: Exposição de História do Brasil, Congresso Brasileiro de Biblioteconomia e Documentação, etc.;
- os títulos individuais de publicações e impressos de diferentes categorias, tais como: bibliografias, revistas, livros etc. Ex.: Anais da Imprensa Nacional, O ensino da Biblioteconomia no Brasil, etc.;
- os nomes próprios personativos e locativos. Ex.: Edson Nery da Fonseca, Ouro Preto, etc.;
- as expressões latinas. Ex.: "servus servorum scientiae".

#### 3.1.3 – Não foram incluídas na contagem:

- datas. Ex.: 1954;
- indicações numéricas de período de tempo abrangendo mais de um ano. Ex.: décadas, século XX, etc.

#### 3.1.4 – As palavras restantes foram contadas uma a uma.

### 3.2 – Processos

O processo de contagem foi idêntico para os três Textos.

- Foram copiadas em fichas todas as palavras de cada um dos Textos A, B e C, de acordo com as normas do indicativo 3 1;
- os termos incidentes em igual número de vezes

foram agrupados em conjuntos;

c) as palavras foram ordenadas por série, em uma tabela, em ordem decrescente, começando-se pelas palavras que ocorreram maior número de vezes. Estas receberam a ordem de série 1; as que apareceram em segundo maior número de vezes, receberam a ordem de série 2, e assim sucessivamente;

d) foi obtida a *ordem de mm* de duas ou mais palavras diferentes com mesma freqüência, dividindo-se a soma de suas ordens de série, pelo número de palavras a elas correspondentes. Ex.: A ordem de série 10,5, na Tabela 1, foi obtida, somando-se os números 10 e 11, e dividindo-se o resultado por 2 (número de palavras). O número 12, da ordem de série seguinte, foi aquele não envolvido na soma;

e) foi calculado o produto da ordem de série *versus* freqüência de palavras, a fim de se obter a constante C e se verificar a conformidade com a primeira lei de Zipf (Tabelas 1, 2 e 3);

f) as palavras dos Textos com suas respectivas ordens de série, foram listadas nas Tabelas 1A, 2A, e 3A;

g) os Gráficos 1, 2 e 3, correspondentes às Tabelas 1, 2 e 3 (ordem de série *versus* freqüência de palavras), foram calculados e determinados pelos pontos de encontro dos logaritmos das freqüências com as ordens de série das palavras;

h) foi aplicada a segunda lei de Zipf, segundo reformulação de Booth, tentando-se obter variações da constante 2, para maior precisão dos resultados (Fórmulas F1, F1A; F2, F2A e F3, F3A);

i) a fórmula de transição de Goffman foi calculada, tentando-se obter variações da constante 2, para maior precisão dos resultados (Fórmulas T1, T1A, T1B; T2, T2A, T2B e T3, T3A, T3B).

## 4 – RESULTADOS

As Tabelas 1, 2 e 3 mostram a ordenação por série das palavras, em ordem decrescente, onde a *ordem de série* 1 representa a palavra que apareceu em maior número de vezes; a *ordem de série* 2 representa a palavra que apareceu em segundo maior número de vezes, e assim sucessivamente. A seguir, é dado o *número de palavras* correspondentes à ordem de série; isto é, total de palavras que apareceram em uma ordem de série qualquer, seguidas de sua *freqüência*.

O *produto total* representa o número de palavras *versus* a freqüência; e o *constante C*, a ordem de série **versus** a freqüência.

As Tabelas 1A, 2A e 3A mostram as palavras dos Textos analisados com suas respectivas ordens de série.

As características numéricas dos Textos analisados são:

*Texto A*: 6395 palavras, sendo 1791 palavras diferentes.

*Texto B:* 2210 palavras, sendo 839 palavras diferentes.

*Texto C:* 1290 palavras, sendo 535 palavras diferentes.

Estas características numéricas e suas ocorrências até 5 vezes, estão incluídas na Tabela 4.

Os percentuais de ocorrências únicas nos vocabulários dos três Textos, estão incluídos na Tabela 5.

Os Gráficos 1, 2 e 3 representam os resultados obtidos respectivamente nas Tabelas 1, 2 e 3, onde a linha foi obtida unindo-se os pontos dos logaritmos naturais da ordem de série *versus* frequência de palavras.

De acordo com a fórmula de Zipf, modificada por Booth, a constante ideal para a língua inglesa é representada pelo número 2; esta constante foi sucessivamente modificada até chegar ao valor 1,50 considerado como a constante adequada para a língua portuguesa, conforme mostram os resultados das Fórmulas F1, F1A, F2, F2A, F3, F3A.

A Fórmula de transição de Goffman — aquela cujo resultado final é uma ordem de série T, que indica, aproximadamente, as ordens de série nas quais deveria se encontrar as palavras significativas de um texto em língua inglesa — foi experimentada nos três Textos, empregando-se as constantes 2, 1,50 e 7, e os resultados obtidos encontram-se na Fórmula de transição de Goffman, Textos A, B e C.

Observa-se, pela comparação das Tabelas 1, 2 e 3, que a constante C se manteve nos três Textos, comprovando, assim, o "princípio do menor esforço" de Zipf.

Estranhamente, na Tabela 1A, as palavras "bibliografia" e "bibliotecas", que correspondem às ordens de série 13 e 14, respectivamente, estão incluídas entre as palavras consideradas como de alta frequência, isto é, artigos, preposições, conjunções, etc.

De acordo com as Fórmulas F1 a F3A, observa-se que a constante 2, aplicada a língua portuguesa, produziu resultados discrepantes. Por exemplo: no Texto A, a frequência 2, que deveria corresponder, de acordo com a fórmula, a 372 palavras, corresponde na realidade a 281 palavras (Tabela 1); no Texto B, a frequência 2, que deveria corresponder, de acordo com a fórmula, a 180 palavras, corresponde na realidade a 122 palavras (Tabela 2); no Texto C, a frequência 2, que deveria corresponder, de acordo com a fórmula, a 120 palavras, corresponde na realidade a 78 palavras (Tabela 3). A constante 1,50, de acordo com os mesmos exemplos, foi aquela que mais se aproximou dos resultados reais:

*Texto A:* 279,5 (281 palavras na Tabela 1)

*Texto B:* 135 (122 palavras na Tabela 2)

*Texto C:* 90 (78 palavras na Tabela 3)

A Fórmula de Goffman, com a constante 2, resultou na ordem de série 47, como sendo a mais próxima das palavras significativas do Texto A, o

que não corresponde à realidade, de acordo com a Tabela 1A.

A constante 1,50, para o mesmo Texto, resultou na ordem de série 62,6, bem mais próxima da realidade, do que o obtido pela constante 2, conforme se observa nas ordens de série 63,5 e 79,5 onde aparecem, respectivamente, as palavras "brasileiras", e "bibliografias", "corrente", "Cursos de Biblioteca Nacional da Biblioteca Nacional".

A constante 7, para o mesmo Texto, resultou na ordem de série 13, relativa à palavra "bibliografia", uma das palavras significativas do Texto em questão, e que aparece, também, no plural, na ordem de série 79,5 bem próxima da 62,6, resultado obtido pela constante 1,50.

A Fórmula de transição de Goffman, com constante 2, aplicada ao Texto B, ajustou-se à ordem de série 32, que se aproxima bastante da realidade.

A constante 1,50, para o mesmo Texto, ajustou-se à ordem de série 43,5 que, curiosamente, está entre as ordens de série 37,5 e 51, e que incluem respectivamente, as palavras "informação" e "especialista", consideradas significativas para o Texto em apreço.

A constante 7, para o mesmo Texto, não correspondeu à realidade,

A Fórmula de transição de Goffman, com constantes 2 e 1,50, aplicadas ao Texto C, não correspondem à realidade.

A constante 7, foi a mais apropriada, pois, nas ordens de série 8 e 9, respectivamente, aparecem as palavras "bibliografia" e "Instituto Brasileiro de Bibliografia e Documentação", consideradas significativas para o Texto em questão.

## 5 – CONCLUSÕES

A primeira lei de Zipf é válida para a língua portuguesa, uma vez que a *constante* deve ser tomada no seu sentido mais alto. Há, na realidade, uma grande aproximação numérica, que pode ser notada nas Tabelas 1, 2 e 3, para a constante C. A segunda lei de Zipf, tal como formulada, parece não ser válida para a língua portuguesa, mas pode ser adaptada a esse idioma. A fórmula original

$$\frac{I_n}{I_1} = \frac{2}{n(n+1)}$$

poderia ser modificada para:

$$\frac{I_n}{I} = \frac{1,50}{n(n+1)}$$

A Fórmula de transição de Goffman, assim como a segunda lei de Zipf, talvez tenham que sofrer

uma adaptação para a língua portuguesa, em face dos resultados mostrados no indicativo 4.

Estudos posteriores mostrarão se a constante 2 válida para a língua inglesa, segundo a Fórmula de transição de Goffman, deverá ser substituída pela constante 1,50 ou 7.

Os resultados obtidos nesse trabalho representam uma investigação em um campo de linguagem técnica, restrito ao assunto "Bibliografia". Certamente aplicados a outras áreas específicas do conhecimento, poderão ser modificados, não só em relação às palavras, mas, principalmente, em relação aos índices de suas respectivas frequências.

Outros estudos são necessários, principalmente para determinar a constante ideal para a segunda lei de Zipf e a Fórmula de transição de Goffman.

A metodologia usada talvez seja a responsável pela inclusão das palavras "bibliografia" e "bibliotecas" entre aquelas de alta frequência na Tabela 1A, pois suas formas singular e plural foram tomadas separadamente. Este critério, válido para a língua inglesa, talvez não o seja para outras línguas, inclusive a portuguesa.

O presente estudo foi abordado sob um aspecto teórico. No entanto, convém assinalar que a descoberta da frequência de distribuição de palavras em um texto, pode ter importantes aplicações práticas com diversas finalidades, tais como:

- a) indexação automática de textos completos de documentos (onde as palavras significativas podem ser racionalmente selecionadas, segundo seus vários graus de especificidade);
- b) associação de documentos baseada nos percentuais de superposição em altas frequências;
- c) comparação de termos de documentos para construção de "thesaurus", ou controle de outros tipos de vocabulários;
- d) estabelecimento de novas palavras a serem incluídas em um "thesaurus", ou em qualquer outro instrumento de vocabulário;
- e) estudo da riqueza de vocabulários de diferentes autores, ou diferentes assuntos.

## 6 – CITAÇÕES BIBLIOGRÁFICAS

- 1 – PRITCHARD, A. – Statistical bibliography or Bibliometrics? *Journal of Documentation*, 25(4):348-9, Dec. 1969.
- 2 – SARACEVIC, T. – *Introduction to Information Science*. New York, Bowker, 1970. 751 p.
- 3 – ZIPF, G. K. – *Human behaviour and the principle of least effort; an introduction to human ecology*. Facsim. of 1949, ed. New York, Hafner, 1965. 573 p.
- 4 – ZIPF, G. K. – *The Psycho-biology of Language; an introduction to dynamic philology*. Cambridge, Mass., M. I. T. press, 1965. 336 p.

5 – MANDELBROT, B. – Structure formelle des textes et communication. *Word*, 10(1):1-27, Apr. 1954.

6 – PARKER-RHODES, A. F. & JOYCE, T. – A theory of word-frequency distribution. *Nature*, 178 (4545): 1308, Dec. 1956.

7 – FAIRTHORNE, R. A. – Progress in Documentation empirical hyperbolic distributions for Bibliometric description and prediction. *Journal of Documentation*, 25 (4): 319-43, Dec. 1969.

8 – BOOTH, A. D. – A "law" of occurrences for words of low frequency. *Information and Control*, 10(4): 386-93, Apr. 1967.

9 – BOOTH, A. D. – On the geometry of libraries. In: SARACEVIC, T. - *Introduction to Information Science*. New York, Bowker, 1970. p. 456-73.

## SYNOPSIS

The researches in Information Science oriented to Linguistics have a particular interest to Bibliometrics; the latter indicates the quantitative treatment of written communication; researches related to word frequency in a given text are found among several applications of Bibliometrics.

Zipf propounded laws that establish relationships between the rank of a word in a frequency order, and the frequency of its appearance in a long enough text; his studies were based on an empirical relationship observed by Estoup.

Booth studied and developed Zipf's second law, less known than the first one, which refers to low frequency words.

W. Goffman studied Zipf's two laws and derived a transition formula which indicates approximatively the ranks in which significant words in a text in English must be found.

The following hypothesis were expressed: Zipf's first law is valid for the Portuguese language; his second law and Goffman's transition formula are also valid for the Portuguese language, or they can be adapted to this language.

Three texts in Portuguese dealing with Bibliography were chosen: Text A, with 6395 words, 1791 of which are different; Text B, with 2210 words, 839 of which are different; and Text C, with 1290 words, 535 of which are different.

Zipf's laws and Goffman's transition formula were applied to these texts, and calculations were made to obtain variations of the constant 2 (valid for the English language), for finding more accurate results.

The following conclusions were drawn: Zipf's first law is valid for Portuguese language; Zipf's second law, as well as Goffman's transition formula, do not seem to be valid for Portuguese language, but can be adapted to this language.

**TABELA 1**  
**ORDEM DE SÉRIE *VERSUS* FREQUÊNCIA DE PALAVRAS**  
**TEXTO A**

Ordem de Série	Nº de palavras	Frequência	Produto total	R x F = C
1	1	320	320	320
2	1	303	303	606
3	1	246	246	738
4	1	177	177	708
5	1	156	156	780
6	1	126	126	756
7	1	112	112	784
8	1	88	88	704
9	1	69	69	621
10,5	2	58	116	609
12	1	56	56	672
13	1	55	55	715
14	1	54	54	756
15	1	53	53	795
16,5	2	50	100	825
18	1	49	49	882
19	1	46	46	874
20,5	2	40	80	820
22	1	35	35	770
23	1	34	34	782
24,5	2	33	66	808,5
26	1	32	32	832
27	1	30	30	810
28	1	29	29	812
29	1	28	28	812
31	3	26	78	806
33,5	2	25	50	837,5
35,5	2	23	46	816,5
37,5	2	22	44	825
40	3	21	63	840

**TABELA 1 (cont.)**

Ordem de Série	Nº de palavras	Frequência	Produto total	$R \times F = C$
42,5	2	20	40	850
44	1	19	19	836
46	3	18	54	828
49	3	17	51	833
51,5	2	15	30	772,5
55	5	14	70	770
59	3	13	39	767
63,5	6	12	72	762
70,5	8	11	88	775,5
79,5	10	10	100	795
91	13	9	117	819
109	23	8	184	872
130	19	7	133	910
152	25	6	150	912
185,5	42	5	210	927,15
236	59	4	236	944
329	127	3	381	987
533	281	2	562	1066
1232,5	1118	1	1118	1232,5
	1791		6395	

**TABELA 2**  
**ORDEM DE SÉRIE VERSUS FREQUÊNCIA DE PALAVRAS**  
**TEXTO B**

Ordem. de Série	Nº de palavras	Frequência	Produto total	R x F = C
1	1	128	128	128
2	1	72	72	144
3	1	68	68	204
4	1	53	53	212
5	1	50	50	250
6,5	2	36	72	234
8,5	2	27	54	229,5
10,5	2	26	52	273
12,5	2	21	42	262,5
14	1	20	20	280
15	1	18	18	270
16	1	16	16	256
17	1	15	15	255
19	3	14	42	266
21,5	2	13	26	279,5
23	1	12	12	276
26	5	11	55	286
29	1	10	10	290
32	5	9	45	288
37,5	6	8	48	300
43,5	6	7	42	304,5
51	9	6	54	306
65	19	5	95	325
90,5	32	4	128	362
141	69	3	207	423
236,5	122	2	244	473
568,5	542	1	542	568,5
	839		2210	

**TABELA 3**  
**ORDEM DE SÉRIE VERSUS FREQUÊNCIA DE PALAVRAS**  
**TEXTO C**

Ordem de Série	Nº de palavras	Frequência	Produto total	$R \times F = C$
1	1	64	64	64
2,5	2	56	112	140
4	1	31	31	124
5	1	30	30	150
6	1	27	27	162
7	1	24	24	168
8	1	20	20	160
9	1	17	17	153
10,5	2	15	30	157,5
12,5	2	12	24	150
15	3	11	33	165
17,5	2	10	20	175
19,5	2	9	18	175,5
22	3	8	24	176
24,5	2	7	14	171,5
29,5	8	6	48	177
40	13	5	65	200
56,5	20	4	80	226
82	31	3	93	246
136,5	78	2	156	273
355,5	360	1	360	355,5
	535		1290	

TABELA 1 A

PALAVRAS DO *TEXTO A* COM RESPECTIVAS ORDENS DE SÉRIE

Ordem de Série			
1	de	40	à, não, primeira
2	a	42,5	ao, história
3	e	44	cursos
4	em	46	Instituto Brasileiro de Bibliografia e
5	o		Documentação (IBBD), parte, sendo
6	do	49	senhora (sra.), seu, trabalhos
7	da	51,5	além, trabalho
8	que	55	hoje, Instituto Nacional do Livro, nos,
9	no		organização, São Paulo
10,5	os, por	59	publicado, segundo, sua
12	foi	63,5	ainda, até, brasileiras, já, sob, volume
13	bibliografia	70,5	cargo, diretor, está, mas, pelos, pu-
14	bibliotecas		blicação, publicados, só
15	as	79,5	ano, bibliografias, brasileiros, corrente,
16,5	um, uma		Cursos de Biblioteconomia da Biblio-
18	com		teca Nacional, grande, literatura, ou,
19	dos		reforma, seus
20,5	Brasil, como	91	artigos, cada, Catálogo da Exposição
22	na		de História do Brasil, colaboração,
23	é		deste, especializadas, esse, Europa,
24,5	biblioteconomia, pelo		Manoel Cícero Peregrino da Silva,
26	foram		Manual Bibliográfico de Estudos Bra-
27	para	109	sileiros, muito, obra, publicou
28	bibliotecários		apenas, assuntos, Benjamin Franklin
29	sobre		Ramiz Galvão (Ramiz Galvão), bi-
31	Biblioteca Nacional (Real Bibliotheca), brasileira, das		bliográfica, ciências, desde, direção,
33,5	pela, primeiro		Distrito Federal, documentação, duas,
35,5	se, senhor		ele, essa, esta, livros, nacional, outros,
37,5	mais, ser		partir, primeiras, Primeiro Congresso
			Brasileiro de Biblioteconomia, pri-
			meiros, sido, também, tem

## TABELA 1 A (cont.)

Ordem de  
Série

130	Alfredo do Valle Cabral (Valle Cabral), aos, bibliográficas, bibliográficos, bibliotecária, Boletim Bibliográfico [da Biblioteca Nacional], dentre, documentos, entre, fundada, iniciativa, maior, pode, Rio de Janeiro, serviço, técnicos, tendo, vem, volumes		Conferência sobre o Desenvolvimento dos Serviços de Bibliotecas Públicas na América Latina, congressos, correspondente, criado, cuja, Departamento de Cultura [de São Paulo], depois, desse, desses, escolas, escritor, especial, especializados, estado, estrangeiros, estudar, estudos, existem, existentes, feita, Fidelino de Figueiredo, filologia, gerais, informes, Instituto Internacional de Bibliografia, julho, legal, maio, meio, mesmo, ministrado, normas, norte-americano, novo, número, onde, organizar, país, promovendo, público, qual, quase, realizados, relatório, reorganizada, Rubens Borba de Moraes, têm, todos
152	anos, Antônio Simões dos Reis, brasileiro, catálogo, classificação, curso, edifício, então, fazer, fez, fichas, fim, geografia, muitos, nas, obras, professor (prof.), promovido, publicações, publicada, Recife, retrospectiva, seguintes, teve, Unesco		
185,5	administração, arranjo, às, autor, Biblioteca Pública Municipal de São Paulo, bibliotecário, Bruxelas, catalogação, concursos, criação, Curitiba, desenvolvimento, devem, dirigida, era, Estados Unidos da América do Norte (Estados Unidos), estão, fase, filosofia, grupo, há, influência, intensivos, modo, outras, pelas, periódicos, promoveu, quando, realizou-se, recentemente, referência, são, século, segunda, serviços, suas, Sully Brodbeck, técnica, tempo, vários	329	aliás, alunos, Anais da Biblioteca Nacional, antropologia, apresentação, apresentado, àquele, arte, artes, aspectos, Associação dos Servidores Civis do Brasil, Associação Paulista de Bibliotecários, associações, atualização, atualizada, auspícios, Bahia, bairros, basta, Belo Horizonte, bem, bibliográfico, biblioteconômica, Boletim Informativo [do IBBD], capítulos, catálogos, central, cidade, Classificação Decimal Universal (CDU), constituída, contendo, contribuições, cultura, Curso de Pesquisas Bibliográficas em Ciências Naturais, depósito, descrição, dezembro,
236	acordo, alguns, Antônio Caetano Dias, anual, assunto, biblioteca, Biblioteca Central da Universidade de São Paulo, capítulo, coletivo, Comissão Nacional de Bibliografia, concluintes,		

**TABELA 1 A (cont.)**

Ordem  
de Série

329	<p>(cont.)</p> <p>diplomática, ela, embora, escola, especializada, esses, excelente, folclore, formação, fundação, gênero, geral, Herbert Coblans, historiador, iconografia, igualmente, importante, índice, infantil, infantis, infelizmente, informe, iniciada, iniciado, instaladas, instituições; intensa, internacional, introdução, junho, lei, Lenyra Fraccaroli, logo, Lydia de Queiroz Sambaquy, medicina, melhor, mencionadas, mencionados, mencionar, ministérios, modernas, monumental, norte-americano, nossa, nosso, nossos, numismática, organizações, organizada, orientação, outubro, participação, passou, periódicas, pesquisas, pois, precursor, prêmio, preocupação, promovidos, própria, publica, públicas, química, realizada, realizado, rede, registra, registrando, representando, respeito, responsável, revistas, Rio Grande do Sul, Salvador, senhores (srs.), serão, Série Bibliográfica de Estudos Brasileiros, sociais, tanto, tarde, território, título, todas, tomaram, unidades, Universidade do Recife, vez, virtude, vista</p>	<p>ra de Normas Técnicas (ABNT), atividades, atuação, atualmente, autarquias, autora, autores, avulsos, base, Bibliografia Brasileira, Bibliografia Brasileira de Documentação, Bibliografia das Bibliografias Brasileiras, bibliógrafo, Biblioteca da Escola Politécnica da Universidade da Bahia, Biblioteca do Departamento Administrativo do Serviço Público, Biblioteca Infantil Carlos Alberto, Biblioteca Infantil Monteiro Lobato, Biblioteca Municipal [do Distrito Federal], Biblioteca Pública da Bahia, Biblioteca Pública do Espírito Santo, Biblioteca Pública do Estado do Maranhão. Biblioteca Pública do Estado do Paraná, Biblioteca Pública Infantil [de São Paulo], biblioteconômicos, biologia, Boletim Bibliográfico Brasileiro, bolsistas, botânica, brevemente, Camillo de Monserrate (Camillo Cléau), catálogo-dicionário, católica, Centro de Pesquisas de Geografia do Brasil da Faculdade Nacional de Filosofia, centros, chamado, chamados, científicos, cientistas, Classificação Decimal de Dewey, classificada, começou, comentada, Comissão Técnica de Documentação [da ABNT], Comissão Técnica de Documentação da International Or-</p>
533	<p>acompanhada, acrescentando, adiante, administrativamente, adotada, agora, alfabético, algumas, âmbito, anais, apresentava, aqui, Associação Brasilei-</p>	

**TABELA 1 A (cont.)**

Ordem de  
Série

533

(cont.)

ganization for Standardization (ISO/TC 46), compilação, completa, compreendendo, comprovar, comunicações, conferências., Congresso Internacional de Bibliotecas e Centros de Documentação, considerada, constituem, construção, construídos, contava, contribuição, contribuíram, correlatos, criada, crítico, cujo, data, decimal, decisivo, Denise Fernandes Tavares, dentro, Departamento Administrativo do Serviço Público (DASP), Departamento de Documentação e Cultura [do Recife], desenvolvido, dessa, desta, destacam-se, deve, devendo, deveria, dicionários, diferentes, difícil, diplomada, diplomados, direito, dirigido, discutidos, disso, diversos, divulga, divulgando, doações, doutor (dr.), duração, edifícios, educação, eles, elogiável, encontrada, ensino, entretanto, época, esboço, escritores, especialmente, estando, estava, este, etc., etnologia, europeu, excelentes, Faculdade de Filosofia, Ciências e Letras da Universidade de São Paulo, falar, Federação Internacional das Associações de Bibliotecários, feito, filiada, física, folhetos, fosse, frei, Fundação Getúlio Vargas, governo, guia, havendo, Henri La Fontaine (La Fontaine), igreja, im-

portância, importantes, impressão, inaugurou-se, independentes, indicaremos, iniciados, iniciou, instituição, instituído, Instituto Brasileiro de Educação, Ciência e Cultura (IBECC), inteiramente, interesse, interrompido, introduções, Irene de Menezes Dória, Jordão Emerenciano, José Honório Rodrigues, lançou, levantamento, Library of Congress, Lilá do Rego Maciel, língua, lista, literária, maneira, manuais, Maria Luisa Monteiro da Cunha, Mário de Andrade, matemática, material, Meier, melhores, Melvil Dewey, Mercedes de Moura Reis Pequeno, mesma, ministrados, modelar, modelo, muitas, municipal, música, nacionais, nasceu, nem, nesta, neste, nova, novas, oficiais, ora, organizado, organizadores, órgão, outra, outro, papel, Paul Otlet, Pedro Gomes Ferrão Castello Branco, pequena, período, períodos, Pernambuco, pesquisadores, pioneiro, planejada, podem, podendo, poderiam, populares, Porto Alegre, portuguesa, posteriormente, preparo, presentemente, principais, principal, principalmente, problemas, produção, professora, projeto, promover, proposta, próprios, pública, publicadas, publicar, públicos, quase-europeu, recebi-

**TABELA 1 A (cont.)**

Ordem  
de Série

533	<p>(cont.)</p> <p>das, referências, reformadas, registrar, registro, regulamento, regulares, relativamente, relatórios, renovador, reorganização, reorganizadas, repertórios, representa, reprodução, respectivamente, resultado, resultou, resumos, revista, rigorosamente, Robert Smith, Rodolfo Garcia, salões, seção, seções, Segundo Congresso Brasileiro de Biblioteconomia, sem, senhoras (sras.), senhorita (sta.), sentido, setores, sistema, sociologia, sugestão, surto, tenham, tentativa, toda, tomos, trata, últimas, Universidade da Bahia, Universidade de [São Paulo], Universidade do Brasil, universidades, universitárias, várias, vêm, verdadeiramente, verdadeiro, verificar, vulto, Washington, D.C., zoologia.</p>	<p>Imprensa Nacional, analisava, analítica, anexo, antigas, Anuário Estatístico do Brasil, anunciado, anunciava, apareceu, aparecido, aplicadas, apresentadas, apresentados, apresentam, apresentando-se, aprovadas, aquisição, Ariceya Moreira Lima, Armindo Guaraná, arqueologia, arquiteto, arquivo, artigo, artistas, artística, Assembléia de Bibliotecários das Américas, assinala, assinalando, assinalou, assistência, assistiram, associação, Associação Baiana de Bibliotecários, Associação Brasileira de Bibliotecários, Associação Paranaense de Bibliotecários, Associação Pernambucana de Bibliotecários, Associação Riograndense de Bibliotecários, Astrogildo Pereira, atenção, atender, atingiram, atividade, atravessa, atribuições, atual, Augusto Meyer, Augusto Victorino Alves Sacramento Blake, aulas, áureo, Áureo Ottoni, autêntica, autêntico, autoria, bairro, barão, bases, básicos, belas-artes, beneditinos, beneficiaram, benefício, benemérito, Bernadette Sinay Neves, Bibliografia Bibliotecológica Brasileira, Bibliografia Brasileira Corrente, Bibliografia Brasileira de Ciências Sociais, Bibliografia Brasileira de Educação, Bibliografia Cartográfica Bra-</p>
1232,5	<p>abafa, abono, abordadas, abrangendo, abreviatura, abrigada, absoluta, absorvido, acabou, acadêmica, aceitação, acervo, acontecimento, adaptados, Adelpha Silva Rodrigues de Figueiredo, administrativa, adotar, afirmativa, Afrânio Proença, agremiação, agremiações, aí, Aida Furtado, alcançado, Alcides Bezerra, alfabética, Alfredo de Carvalho, almejado, Alpheu Diniz Gonçalves, alta, Amazonas, ambos, América latina, amplitude, Anais da</p>	

**TABELA 1 A (cont.)**

Ordem  
de Série

1232,5 (cont.)

sileira, Bibliografia da História da Literatura Brasileira, Bibliografia de Literatura Infantil em Língua Portuguesa, Bibliografia Econômico-social, Bibliografia Geográfico-estatística Brasileira, Bibliografia Nacional, Bibliografias, bibliograficamente, Bibliographie Cartographique Internationale, A Biblioteca, Biblioteca Central |da Universidade do Brasil|, Biblioteca Demonstrativa Castro Alves, Biblioteca do Departamento de Profilaxia da Leprosia, Biblioteca do Instituto Mackenzie, Biblioteca do Instituto Nacional de Tecnologia, Biblioteca do Instituto Tecnológico do Estado do Rio Grande do Sul, |Biblioteca do Ministério| da Educação e Cultura, |Biblioteca do Ministério| da Fazenda, Biblioteca do Ministério da Viação e Obras Públicas, |Biblioteca do Ministério| das Relações Exteriores, |Biblioteca do Ministério| do Trabalho, Indústria e Comércio, biblioteca-piloto, Biblioteca Popular da Encruzilhada, Biblioteca Popular de Afogados, Biblioteca Popular de Casa Amarela, Biblioteca Popular de Santo Amaro, |Biblioteca Pública do Estado de| Minas Gerais, bibliotecárias, Bibliotecas Infantis, Bibliotecas Populares, Bibliotecas Populares da

Prefeitura do Distrito Federal, Bibliotecas Públicas e Intercâmbio, Bibliotecas Universitárias e alguns de seus Problemas, bibliotecografia, biblioteconômico, bimestrais, bimestral, biografias, biógrafo, bispo, boas, Boletim Bibliográfico |da Biblioteca Pública Municipal de São Paulo|, bolsas, bolsista, bom. Botafogo, Brunet, buscamos, bustos, caberia, Câmara Brasileira do Livro, Campo Grande, capital, casos, Cecil Knight Jones, célebre, centralização, centro, Centro Regional da Unesco, certamente, certos, cessou, chamar, chefe, chefes, chegada, cifra, Cim, circulação, citada, citadas, citado, citados, citar, civil, classes, classificações, cobrem, códigos, coincide, colaborador, Coleção B-1: Bibliografia, Coleção B-2: Biblioteconomia, coleções, colégio, colégios, Colloquium Internacional de Estudos Luso-brasileiros, começa, começam, começaram, começo, comentadas, comentários, comerciais, Comissão Consultiva de Bibliografia, Comissão de Estudo dos Textos da História do Brasil, Comissão do Quarto Centenário, Comissão Nacional do Catálogo Coletivo, comissões, compareceu, competentes, compila, compila-

**TABELA 1 A (cont.)**

Ordem  
de Série

1232,5

(cont.)

das, compilador, compilar, compilou, complemento, completada, completamente, compõe-se, compra, compreendeu, compreendia, compreendiam, comum, comunicação, concebida, concedendo, conceituada, conclave, concluídos, concurso, condição, condigna, confessadamente, confortáveis, Congresso de Bibliotecas do Distrito Federal, conhecedor, conhecendo, conhecido, conhecimento, conjunta, conseguir, Conselho Nacional de Pesquisas, considera, Considerações sobre a Biblioteca Universitária, considerado, considerar, considerou-se, consolidou-se, constante, Constantin, constitucional, constitui, constituindo, consulta, consultados, consultar, consultor, contém, continental, continua, continuada, continuado, conventos, cooperação, cooperativa, coordenação, Copacabana, correção, correspondência, Cousin, criadas, criança, criavam, criou, crise, críticas, críticos, cultural, cumprida, currículo, Curso de Biblioteconomia de Manaus, Curso de Biblioteconomia do Departamento Administrativo do Serviço Público, custeou, dando-lhe, daquele, dar, datam, datilográfica, década, décadas, Décio Pereira de Vasconcelos, decla-

ração, decorrentes, decretos, dedicado, deficiência, deixando, delas, deliberada, Delisle, demais, demográfica, denominada, densidade, Department of State, Departamento de Fisiologia Geral e Animal, Departamento de Imprensa Nacional, depositárias, descobrimento, descritivo, desempenhado, dessas, destacada, destacadas, destacados, destacamos, destes, destinados, detalhadas, detalhes, deu-lhe, deve-se, devemos, devida, devoção, dia, Dictionario Bibliographico Brasileiro, dicionário, Dicionário Bio-bibliográfico, diferente, dificuldades, diretamente, diretores, Diretoria de Estatística, Propaganda e Turismo [Recife], diria, dirigia, dirigidas, dirigidos, dirigiu, dispersos, dispõe, dissemos, distinguido, distribuindo, dita, divide, Divisão de Bibliotecas Infantis [São Paulo], divulgação, divulgadas, divulgados, divulgar, divulgou, dizer, Do Serviço de Importação de Material Bibliográfico para as Bibliotecas Universitárias, doação, documentalista, documento, Dolores Iglésias, Dom Pedro II, domínio, Dorothy Muriel Geddes (Mrs. Arthur E. Groppe), duplicação, Duque de Berry, durante, eclesíastica, École des Chartes, econômica, econômicos, Edi-

**TABELA 1 A (cont.)**

Ordem  
de Série

1232,5 (cont)

son Carneiro, edita, editada, editar, editora, editores, editoriais, Edson Nery da Fonseca, efeito, efetuada, elaborar, elas, emanados, empolgado, emprega, encarregado, encarregou, encerramento, encontram-se, encontramos, encontraremos, enfadonha, engenharia, engenheira, ensaio, ensaistas, O Ensino da Biblioteconomia no Brasil, entidade, entradas, entregou, entregues, entretimentos, entusiasta, enumeração, enviou, eram, Ernesto Manuel Zink, escapar, escapou, esclarece, esclarecer, esclarecimentos, Escola de Biblioteconomia de São Paulo, Escola, Modelo "Caetano de Campos", escolar, escreve, escreveu, escritos, especialista, especialistas, especializado, espécie, esperar, esquecidos, essas, estabeleceu, estabelecido, estaduais, estadual, estampados, A Estante, estatística, estatísticas, estátuas, estimula, estimular, estivera, estrangeiras, estrangeiro, estreita, estudantes, estudioso, estudiosos, Eugênio Gomes, europeia, europeias, europeus, evita, exame, exclusividade, executado, exerce-la, exerciam, êxito, experiência, explica-se, exposição, Exposição de História do Brasil, exposições, expressão, extensão, extinta, extraordinariamente, extraor-

dinário, extraordinários, face, facsímile, Faculdade de Direito [de Recife], Faculdade de Medicina da Universidade de São Paulo, faculdades, falaremos, família, fantasia, fará, faremos, fascículo, fascículos, fases, fato, faziam, Federação Internacional de Documentação, federal, fevereiro, ficaram, ficção, ficou, figura, figurar, filho, finalidade, financeira, fins, fiscalização, físicas, fontes, fora, formados, fornecendo, fotográfica, francês, francesa, franciscanos, franqueada, frente, freqüentadas, freqüentemente, função, funciona, funcionar, funcionários, fundadora, fundados, fundaram, gabinetes, Gávea, genealogia, geográfica, geológicas, gerações, Gilberto Freyre, governos, Graesel, grandes, gravuras, grupos, guiando-se, Guilherme Studart, Guiomar de Carvalho Franco, Havana, hemerotecas, heráldica, Herbert Baldus, histórias, histórica, holandês, homem, honra, hora, iconográfica, iconográficas, ida, ideal, identificação, imperador, império, impossível, impressa, impressas, impresso, impressos, imprimindo, imprimiu, inadequado, inaugurada, inauguradas, inaugurando, inaugurando-se, inaugurou, incluídos, inclusive, independente, Index Transla-

**TABELA 1 A (cont.)**

Ordem  
de Série

1232,5

(cont)

tionum, indexação, indicados, indicando, Índice Tecnológico, índices, indispensável, individuais, individual, industrial, inédita, inestimáveis, inexplorados, influenciados, influxo, informa, informações, inglês, iniciadores, inicial, inicialmente, iniciar, iniciativas, insertos, inspirada, instalações, instalada, instalar-se, instalaria, instalou, instituiu, instituto, Instituto Mackenzie, Instituto Nacional de Estudos Pedagógicos, Instituto Nacional de Pesquisas da Amazônia, instrutivas, insuficiente, integrar, inteira, intenção, intensificados, intensificou-se, intensiva, interessadas, interessantes, interior, internacionais, International Organization for Standardization (ISO), interrompida, intimamente, introduzidas, introduziu, inúmeras, inventariados, inventário, irregular, isoladamente, isto, Itaim, J. Matoso Câmara Júnior, Jango Fisher, Jannice Monte-Mor Alves de Moraes, jesuitas, João Capistrano de Abreu, João Francisco Velho Sobrinho, João Pessoa, John Casper Branner, Jorge Abrantes dos Santos, Jorge de Andrade Maia, José Carlos Rodrigues, José Césio Regueira Costa, José Valladares, José Veríssimo, Josué Camargo Mendes, Josué de Sousa

Montello, jovens, Juliano Moreira, justas, justo, L. de Castro Faria, Laborde, lacunas, lado, lamentável, lançada, larga, latim, O Leitor e o Bibliotecário, leitores, leitura, lepra, levada, levantada, lido, lidos, ligação, ligado, limitado, limites, listas, literárias, livrarias, livro, localização, Louise-Noëlle Malclés, Luis Heitor Correia de Azevedo, Luis Washington Vita, Luisa Fonseca, Luisa Keffer, luso-brasileira, lutando, M. da Nóbrega, Maceió, Madrid, maiores, maioria, maioridade, Maire, Manaus, manifestar, mantém, manter, Manuel des Échanges Internationaux de Publications, manuscritos, marcante, marco, Margarit Bates, Maria Amélia Araújo, Maria de Lourdes Meneghezzi, marinhas, Martinho da Fonseca, matemáticas, matéria, matérias, mediante, medidas, membros, memoráveis, memória, mencionada, menores, merece, mérito, mesmos, mestres, Miguel Arrojado Lisboa, milhão, militar, Minas Gerais, Ministério das Relações Exteriores, ministradas, ministro, moderna, modernamente, modernos, Morei, Mosteiro de São Bento, mostra, mostrando, motivos, movimento, multilith, Museu do Estado da Bahia, musicólogo, Myriam Gusmão de

**TABELA 1 A. (cont.)**Ordem  
de Série

1232,5 (cont.)

Martins, Nações Unidas, Namur, não-ibérica, nascido, Natal, naturais, natural, natureza, necessária, necessárias, necessário, necessidades, negociações, negócios, nela, neles, nenhum, nessa, nesse, nesses, nível, nome, nomeado, nomes, normalização, nós, nota-se, notando, noticiadas, novos, num, numa, numeração, obedecendo, objetivo, objeto, observações, observadores, ocasião, ocorre, ocupou, oficial, VIII Reunião Anual da Sociedade Brasileira para o Progresso da Ciência, Oliveira Lima, oportuna, oportunidade, organismo, Organização de Bibliotecas de Música, Organização de Bibliotecas Médicas, Organização dos Estados Americanos, Organização e Administração de Bibliotecas Agrícolas, organizador, organizadora, organizaram, organizaria, organizou, órgãos, orientado, originais, Oscar Niemeyer, Otto Maria Carpeaux, Ouro Preto, padres, paisagens, países, paleografia, palestras, Pandiá Calógeras, Paraná, parece, pareceres, Paris, partes, participaram, particulares, partiu, passado, passar, passaram, patrocínio, Paulo Sawaya, Paulo Sérgio Nery, pavilhão, pedagógica, pedido, Pedro Taulois, Peignot, Penha, pensador, pensadores, pequeno,

porcentagem, perfeita, periodicidade, Periódicos Brasileiros de Cultura, permanente, permuta, permutadas, per-tence, pessoal, pessoas, Petzholdt, Pierre Mombeig, pioneira, planejado, planejados, plano, Plano para o Estabelecimento de numa Bibliotheca Pública na Cidade de S. Salvador, plásticas, plenárias, Plínio Ayrosa, podemos, poder, poderia, poesia, Poetas do Brasil, política, políticos, pontos, porém, pormenorizada, porque, portanto, possível, possuem, possui, possuindo, poucas, pouco, precursores, prédios, prefácio, Prefeitura de São Paulo, Prefeitura do Distrito Federal, Prefeitura Municipal do Recife, preliminares, premiar, prenomes, prepara, preparação, preparada, preparados, preparando, presente, presidente, prestando, pretendia, Primeiro Congresso Ibero-americano de Arquivos, Bibliotecas e Propriedade Intelectual, privadas, Processos Técnicos, procuramos, procurando, professores, professou, profética, profeticamente profissional, programa, progressiva, projetos, prolongou-se, promovida, proporcionada, proporcionar, proporções, propriamente, próprias, próprio, propunha, proveito, providências, pseudônimos, publica-se,

**TABELA 1 A (cont)**

Ordem  
de Série

1232,5

(cont)

publicam, publicou-se, puras, quadro, qualquer, quanto, quem, questões, químicas, raros, real, realização, realizações, realizar-se, recebidos, recente, reclamar, reclamava, recomendações, reconhecida, reconhecimento, reconquista, recordaremos, recreação, re-europeização, refere, referente, referimo-nos, reformada, reformas, reformava, regiões, registrados, registrava, regras, regulamentado, regularidade, reinado, reiniciado, reiniciou-se, relação, relativa, religião, remoto, reorganizar-se, reorganizaram, reorganizaram-se, repertório, Repertório Bibliográfico Universal, representantes, reproduzido, residências, responsáveis, ressaltar, restauração, retratos, reunidos, reunindo, Revista do Serviço Público, revolução, rigor, riquezas, Roquette Pinto, sábio, sacerdote, saído, saiu, salientada, salientar, Santana, Santo Amaro, São Luis, Seção de Manuscritos [da Biblioteca Nacional], Seção de Referência [da Biblioteca Nacional], secretária, secretária-executiva, secretarias, séculos, segregação, seguida, seguinte, seguir, seguiram, seguiram-se, Segundo Colloquium Internacional de Estudos Luso-Brasileiros, selecionar, seletiva, semelhantes, semes-

tre, seminário, Seminário sobre Permuta Nacional e Internacional de Publicações, sempre, senão, sentir, separadas, serem, seria, seriam, série, Les Services Bibliographiques dans le Monde, Serviço Central de Informações Bibliográficas [da Universidade da Bahia], Serviço de Informações do Centro Europeu para a Pesquisa Nuclear (CERN), Serviço Gráfico do Instituto Brasileiro de Geografia e Estatística, servir, sessões, setor, Setor de Bibliotecas Populares [da Prefeitura do Distrito Federal], Silvio Romero, Simpósio sobre Bibliografia e Documentação Científica, sinais, Sindicato Nacional das Empresas Editoras de Livros e Publicações Culturais, sistemático, Situação atual da Biblioteconomia Brasileira, Situação das Bibliotecas Infantis Brasileiras, situada, sobrenomes, sobretudo, sociedade, Sociedade Brasileira para o Progresso da Ciência, sociedades, sorte, soube, sozinho, subordinada, subordinadas, subsidiários, substituição, substituído, subúrbio, subúrbios, subvencionadas, sucessor, suficiente, Sugestões para Organização duma Pequena Biblioteca Infantil, Sugestões para uma Cooperação Intensa entre as Bibliotecas Espe-

**TABELA 1 A (cont.)**

Ordem  
de Série

1232,5

(cont)

cializadas do Brasil, superada, superar, superior, superiores, supor, surgir, surgiram, surgiu, tal, tão, tarefas, Tatuapé, teatro, técnico, técnico-profissional, tecnologia, Ternas de Documentação, tendo-se, tenha, ter, terceiro, teremos, terminada, territorial, tinham, tipográfica, tipos, tivermos, títulos, todo, tomadas, tomam, tome, tornam, tomou, total, totalidade, trabalha, trabalhando, traduções, trajes, transformada, transformaram, tratados, trimestralmente, tudo, tupi-guarani, turma, última, ultimamente, último, União

Geográfica Internacional, União Pan-americana, uniformes, universal, universidade, |Universidade de| Minas Gerais, Universidade de Pennsylvania, |Universidade do| Rio Grande do Sul. universitário, usados, uso, usos, utilidade, vai, válido, vamos, vantagem, vão, vastidão, vê, veja-se, velhas, vencedor, verdade, veremos, vezes, viagens, vias, Vila Mariana, vimos, vindo, visão, visita, visitas, vistas, vivamente, viveu, vocabulário, voltaremos, William Berrien, Zeferino Ferreira Paulo.

**TABELA 2 A**

**PALAVRAS DO *TEXTO B* COM RESPECTIVAS ORDENS DE SÉRIE**

Ordem de Série			
1	de		tas, especializado, especializados, esse,
2	que		esses, está, já, material, mesmo, neces-
3	a		sidades, pelo, qual, suas, técnicas
4	o	90,5	aos, aqueles, aspecto, assim, assunto,
5	e		devem, fazer, guia, informações, lado,
6,5	do, em		maior, maneira, muito, nos, nosso, no-
8,5	com, um		vas, num, ora, pessoal, pois, porque,
10,5	os, para		profissão, realizar, são, sem, sendo,
12,5	cursos, no		setor, só, tão, técnica, tipo, vez
14	se	141	acordo, agronomia, além, alto, am-
15	é		bos, aqui, atitude, bem, cada, ca-
16	como		so, centro, centros, científico, cien-
17	da		tista, colaboração, condições, conhe-
19	das, mais, pesquisa		cimentos, contados, dando, dentro,
21,5	não, por		despertar, deve, devemos, dizer, edu-
23	bibliotecário		cação, ênfase, escolas, especializa-
26	ao, bibliográfica, bibliotecários, curso, trabalho		das, essa, faz, grande, guias, há,
29	bibliotecas		importância, iniciação, instrumentos,
32	à, especialização, Instituto Brasileiro de Bibliografia e Documentação, ser, técnicos		literatura, lugar, mostrar, nacional,
37,5	as, dos, formação, informação, na, problemas		necessidade, nós, novos, onde, orga-
43,5	agrícolas, ainda, às, científica, entre, uma		nização, ou, pela, pode, precisa, pré-
51	ciências, eles, especialista, etc., outro, país, serviços, sua, também	236,5	ciso, problema, quase, queremos, rea-
65	alunos, conhecimento, desenvolvimento, documentação, ensino, especialis-		lizado, referência, reformular, sempre,
			sentimos, sentir, seus, tanto, tarefa, ta-
			refas, tecnológica, tempo, todos, uni-
			versidade, Universidade Nacional de
			Brasília, universitário
			acesso, afirma, algum, apenas, apro-
			ximação, aquele, até, atividades, atra-
			vessa, básicos, bibliografia, bibliográ-

**TABELA 2 A (cont.)**

Ordem  
de Série

236,5	<p>(cont)</p> <p>ficas, bibliográfico, bibliográficos, biblioteconômico, Brasil, campo, casos, certeza, certo, conclusões, consequência, dar, demais, depois, desde, desses, determinado, deveriam, diálogo, dinheiro, dirigida, diversos, econômico, eficiente, elaboração, ele, entanto, espírito, esta, estabelecer, estágios, estão, este, estrangeiros, exceção, êxito, falar, fatores, favorável, fazendo, feita, feito, fim, finalidade, foi, futuramente, idéias, importante, imprescindível, incentivadas, interesse, lhe, maioria, melhor, mentalidade, mercado, mesma, mudar, nada, necessária, nível, nossas, nossos, novo, numa, número, objetivos, orientação, outros, papel, perceber, periódicos, pesquisador, pesquisadores, pesquisas, poderá, ponto, pós-graduação, posição, possua, prestar, primeiro, profissões, progresso, próprio, publicações, reais, realiza, realização, reforma, relação, responsáveis, sejam, Seminário [sobre Bibliotecas Agrícolas], será, serão, sistema, tais, tal, tecnológico, têm, temos, tenham, teorias, terem, torna, torna-se, trabalham, traz, universitária, vista</p>	<p>alguns, aliás, alisar, altamente, aluno, ampliados, ampliam, ampliará, ampliarão, analíticos, ângulo, Anísio Teixeira, anos, anteceder, antigos, anula, aparelho, aprender, apresentado, apresentam, aproveitamento, aristocrática, arroladas, artigo, assumida, atenção, atendem, atenta, atinentes, atividade, através, atravessamos, atuação, atual, aumentará, autoridades, auxiliados, auxiliar, avanço, bancos, base, baseadas, bases, biblioteca, Biblioteca do Instituto Agronômico de Campinas, Bibliotecas Especializadas Brasileiras, biblioteconomia, biblioteconômicas, bomba, bons, brasileiras, brasileiros, campos, capacidade, caráter, catalogação, Catálogo Coletivo Nacional, categoricamente, cerca, chamado, chega, ciência, cinco, citar, claramente, clientelas, coisa, coisas, colaborar, coleções, colegas, competente, competentes, complemento, completas, complexa, complexidade, complexos, concluir, conhecem, conjunto, consciência, conseguimos, conseguirmos, consentânea, constitui, contacto, contento, contrário, contratação, controle, convênio, convidado, convívio, convocados, coordenação, coordenados, creio, cremos, crescentes, criação, criar, critério, criticar, cujos,</p>
568,5	<p>abordado, ação, acompanhar, adequada, afastadas, afirmar, age, agora, agrônomos, alcançados, alcançar, alguma,</p>	

**TABELA 2 A (cont.)**

Ordem  
de Série

568,5

(cont.)

cultura, currículo, currículos, Curso de Documentação Científica, Curso de Pesquisas Bibliográficas em Ciências Naturais, dados, daí, decorre, decorrência, dedicarem, deem, definidas, definir, delas, demanda, dentre, derrubadas, descobrir, descuradas, desejada, desenvolve, desenvolvidas, desordenado, desperdício, dessa, desse, determinados, deveria, devido, diferente, difícil, dificuldades, dinâmica, diploma, direito, dispensar, dispor, dispuser, disso, distantes, distribuição, divisas, diz, dizemos, documentalista, dona, duplicidade, econômica, economicamente, economista, economizar, educacional, Education for Documentation, efeito, efetiva, eficiência, ela, elaborado, elementos, empenham, empíricas, empresa, empresas, encaminhar, encontro, energia, engenheiro, engrenagem, enquanto, ensinar, então, entidades, enviados, enviou, envolvem, Ernesto Manoel Zink, escolhido, escolhidos, esforço, espantosa, especial, especializada, específica, esperam, essas, estabelece, Estados Unidos, estariam, estarrecedora, estudante, estudioso, estudiosos, estudos, evidente, evitando, evitando-se, exata, execução, exemplo, existentes, explorado, expõe, exposto, fácil,

facilita, faltará, faltou, famosa, fato, fatos, favorecendo, feição, feitos, fez, fichários, ficou, fiéis, fiquem, firmar, folhetos, for, foram, força, formados, formar, formará, frente, freqüentam, freqüente, fronteiras, funcionária, funcionários, funcione, funcionem, fundamental, futuros, geral, geralmente, grandemente, grandes, grau, graus, guardado, haja, havido, hipótese, homem, honrosa, horizontes, humanístico, igual, imediato, importantes, impresso, imprimindo-lhes, incentivado, incentivar, incluem, índices, indispensável, indústria, industrial, industrialização, indústrias, inerentes, infelizmente, inferiorizados, iniciar, instituições, Instituto Agrônomo de Campinas, institutos, integrado, integrar, intercâmbio, interesses, intrínseca, introdução, ir, isolado, isso, isto, laboratório, latino-americano, latino-americanos, Lei de Diretrizes e Bases da Educação Nacional, leitura, lema, levantamento, levantar, lidam, livros, lhes, locais, longo, Lydia de Queiroz Sambaquy, Mme. Malclés, maiores, mantendo, manter, manuseio, manutenção, máquina, mas, massa, máximo, medicina, medida, meio, menor, mergulhado, métodos, microbiologia, microfil-

**TABELA 2 A (cont.)**

Ordem  
de Série

568,5 (cont.)

mes, Ministério da Ciência e da Tecnologia, ministrado, Mr. Shera, modalidades, modificação, momento, mundo, naquele, nelas, nem, nessa, nesse, ninguém, noção, nomes, nossa, números, nutrição, obras, observando, obter, ocorrerá, odontologia, ordem, organismos, organizar, orienta, orientadores, outra, outras, pari-passu, parte, participação, peças, pedirmos, pelos, pensamento, pequenos, percebe, percebe-se, perda, periódicas, período, permite, permitirá, planejamentos, plano, Plano Orientador da Universidade Nacional de Brasília, plenamente, podemos, poderão, poderemos, poderíamos, polimento, portadores, portanto, possam, possibilidade, possível, possuam, povos, precisam, preferência, preocupação, preparo, prescindirá, presença, presente, presidente, pretende, pretendem, primeira, primeiras, principalmente, proceder, processo, procurada, procurar, procurarem, procuraremos, professores, profissional, profunda, programas, projetada, projetos, promove, propomos, própria, próprios, proveito, providos, provoca, quais, quaisquer, qualificados, quanto,

ramos, realidade, realizados, realizaria, realmente, reconhecer, reconhecimento, recrutado, refazendo, referimos, refletem, regionais, registradas, regulamentação, relevante, rendimento, representa, representam, resolução, respectivamente, respeito, resta, resultado, resultados, retardada, rica, riqueza, sã, saem, São Paulo, satisfatória, seguro, seja, seletiva, senhor, sentem, sentindo, seriam, Serviço de Intercâmbio de Catalogação, serviriam, "servus servorum scientiae", setores, seu, si, sido, Simpósio de Biblioteconomia e Documentação, sob, soberania, sobre, sobremodo, sociológica, solicitada, sólidos, solução, somente, somos, subsistir, superiores, suprir, surgem, surto, talvez, técnico, tem, tendo, tenha, terá, tirando, títulos, tivessem, tomado, tomaram, tópico, tornado, totalidade, totalmente, trabalha, trabalhos, tradução, tratados, trazer, treinamento, troca, tudo, último, últimos, único, uniforme, uniformidade, Universidade do Brasil, universidades, utilizados, vai, valiosa, vantagem, variadas, vários, velhas, vem, ver, vinda, virão, visa, visto, vital, volta, voltarem.

TABELA 3 A

PALAVRAS DO *TEXTO C* COM RESPECTIVAS ORDENS DE SÉRIE

Ordem  
de Série

1	de		ção, conjunto, continuam, corrente,
2,5	a, o		correspondente, criação, dezembro, di-
4	que		divulgadas, dois, editar, estas, gerais,
5	e		Instituto Nacional de Estudos Pedagó-
6	da		gicos (INEP), número (nº, n.), ou-
7	em		tros, pois, possível, primeiro, quatro,
8	bibliografia		setor, sinalética, sistema, sua, todos,
9	Instituto Brasileiro de Bibliografia e		vantagens
	Documentação (IBBD)	136,5	alfabético, América latina, anos, anuais,
10,5	bibliografias, no		apesar, apresentação, às, Associação
12,5	para, se		Brasileira de Normas Técnicas, bem,
15	ciências, com, volume		Bibliografia Brasileira de Ciências So-
17,5	ao, foi		ciais, Bibliografia Brasileira de Medi-
19,5	do, um		cina, bibliográficas, bibliográfico, bo-
22	como, mais, os		tânica, brasileiros, compiladas, conser-
24,5	período, trabalho		vando, correntes, dá, data, dez, dia,
29,5	arranjo, brasileira, especializadas, pu-		divulgar, documentação, embora, en-
	blicado, sendo, também, trabalhos,		tre, escolher, está, etc., física, geral, in-
	uma		cluir, independentes, informação, iní-
40	as, bibliográfica, científica, compilar,		cio, Instituto Nacional do Livro, inte-
	das, é, não, organização, ou, referente,		resse, isto, já, maneira, março, mate-
	ser, sociais, teve		mática, médicas, melhor, nacional,
56,5	aplicadas, assuntos, básicas, bibliográ-		normas, obra, órgão, parece, parte,
	ficos, Brasil, cinco, dos, fase, material,		partir, pela, pelas, periodicidade, pes-
	na, países, pelo, por, quanto, são, Ser-		quisa, prelo, primeira, problema, pro-
	viço de Bibliografia do Instituto Bra-		blemas, produção, promover, qualquer,
	sileiro de Bibliografia e Documenta-		quando, realização, resolver, retrospec-
	ção, seu, sobre, solução, suas		tiva, segundo, Seminário Regional de
82	à, ainda, além, analítica, assunto, Clas-		Bibliografia, Documentação e Permu-
	sificação Decimal Universal, colabora-		ta de Publicações, senhora (sra.), se-

**TABELA 3 A (cont.)**

Ordem  
de Série

136,5 (cont.)

ria, sistemático, só, tão, técnico-científica, temos, último, volumes, zoologia

355,5 abranger, acordo, acumulações, acumulada, acumulando, adesão, adiante, admitem, adotá-las, adotada, adotar, agora, agricultura, alguns, América Central, análise, anteriormente, aos, apareceu, aplausos, apreciação, apresentada, apresentam, apresentamos, aquela, aquele, Argentina, arrojada, artigos, assinalou, até, atingiu, atividades, através, atualizada, atualizado, auspícios, autoras, autores, autoria, base, Bernadette Sinay Neves, Bibliografia Brasileira [do Instituto Nacional do Livro], Bibliografia Brasileira de Agricultura, Bibliografia Brasileira de Bibliografias, Bibliografia Brasileira de Botânica, Bibliografia Brasileira de Educação, Bibliografia Brasileira de Matemática e Física, Bibliografia Brasileira de Química, Bibliografia Brasileira de Zoologia, Bibliografia da América Latina - BAL, Bibliografia de Centro America y del Caribe, Bibliografia e Índice de Geologia do Brasil, Biblioteca Nacional, Boletim Bibliográfico [da Biblioteca Nacional], Boletim Bibliográfico Brasileiro, breve, cabe, cabeçalhos, caberá, cada, Câmara Brasileira do Livro, campo, car-

go, Centro Brasileiro de Pesquisas Educacionais, certa, classificação, classificado, compilador, completa, completado, completas, concurso, condições, conheçam, conhecimentos, Conselho Nacional de Pesquisas, considerando, constituem, contentar, continuação, contrapõem, contribuição, contribuisse, conveniente, coordenar, correlacionados, correlatos, cuidadas, cuidado, cuja, cujos, cumulativas, daí, decidir, decidiram, deixa, delegados, demais, Departamento Administrativo do Serviço Público, Desenvolvimento da Biblioteconomia e da Bibliografia no Brasil, desta, deste, destinado, desvantagens, determinado, deva, deve, deveria, Dewey, diante, difícil, dificuldade, digam, digna, dispendioso, dispondo-se, disso, divulgá-la, divulgação, divulgada, divulgará, dizer, documentos, Dolores Iglésias, duas, dúvida, edições, Edson Nery da Fonseca, educação, efeito, efetivo, elaboração, elaborada, eles, empregados, encarregado, encontra, encontra-se, então, esclarecemos, Escola Politécnica da Universidade da Bahia, escolhido, esforços, especialistas, especializada, especialmente, espera-se, esses, esta, estando, este, estudado, estudar, estu-

## TABELA 3 A (cont.)

Ordem  
de Série

355,5

(cont.)

diosos, estudo, estudou, evidente, evidentemente, executar, exigências, experimentação, extremamente, fácil, facilitar, falaremos, faria, fato, fazer, final, fizeram, focalizando, folhetos, for, fora, freqüente, freqüentes, funções, Fundação Getúlio Vargas, fundamentais, geológicas, graças, grandes, grupo, hoje, humanos, IBBDD Boletim Informativo, ideal, idéia, idioma, imediata, implantação, inclui, incluirá, indicada, índice, índice Catálogo Médico Brasileiro, índice Tecnológico [da Escola Politécnica da Universidade da Bahia], infelizmente, iniciada, inicial, iniciar, instituições, intelectual, internacionais, internacionalmente, introdução, janeiro, Jorge de Andrade Maia, latino-americana, latino-americanos, levado, levantamento, levantamentos, levantaria, lhes, limita, livros, maio, maiores, marco, marcou, Maria de Lourdes Meneghezzi, mas, matéria, médica, medicina, menos, mensais, mesmo, México, morte, mostrará, mundo, nas, natural, naturalmente, nessa, Normalização da Documentação no Brasil, normalizada, nos, notação, nova, novembro, novo, objeto, obrigado, observadas, ocasião, oferece, oportuna, optaria, organizado, outra, outras,

panorama, particular, passado, passando, paulatinamente, periódicas, periódicos, permaneceu, pessoas, plano, pode, podemos, poderá, portanto, preferido, preferiu, prendeu-se, preocupação, preocupações, preocupou-se, prever, prevista, Primeira Reunião da Comissão Latino-americana da Federação Internacional de Documentação, principais, principio, propôs, publicação, publicada, publicadas, qual, química, razoável, realizado, realizou, recomendadas, reconhecendo, refere, regularidade, relação, relativa, relativo, repetição, representa, representar, resolveu, resolvido, respeito, retrospectão, Revista do Livro, sabe-se, saindo, saiu, século, seguintes, sem, semestrais, senhoras (sras.), separadamente, separado, será, serem, série, servem, serviço, seus, Sindicato Nacional das Empresas Editoras de Livros e Publicações Culturais, sinopse, sistematizada, situação, sob, Sociedade Nacional de Agricultura, subordinadas, substituição, superioridade, suplemento, tarefa, técnico, tem, tema, tentada, tentaria, terceiro, Terezine Ferraz, teria, termos, teses, tipo, tipos, títulos, todas, tomou, tratar, universal, Uruguai, uso, utilizado, valor, vários, vêm, vem, veremos, visão, viu-se, vulto.

**TABELA 4**

**CONTAGEM DE PALAVRAS DOS 3 TEXTOS**

Textos	A	B	C
Total de palavras	6395	2210	1290
Palavras diferentes	1791	839	535
Palavras que ocorrem 1 vez	1118	542	360
Palavras que ocorrem 2 vezes	281	122	78
Palavras que ocorrem 3 vezes	127	69	31
Palavras que ocorrem 4 vezes	59	32	20
Palavras que ocorrem 5 vezes	42	19	13

**TABELA 5**

**PERCENTUAIS DE OCORRÊNCIAS ÚNICAS NOS VOCABULÁRIOS DOS 3 TEXTOS**

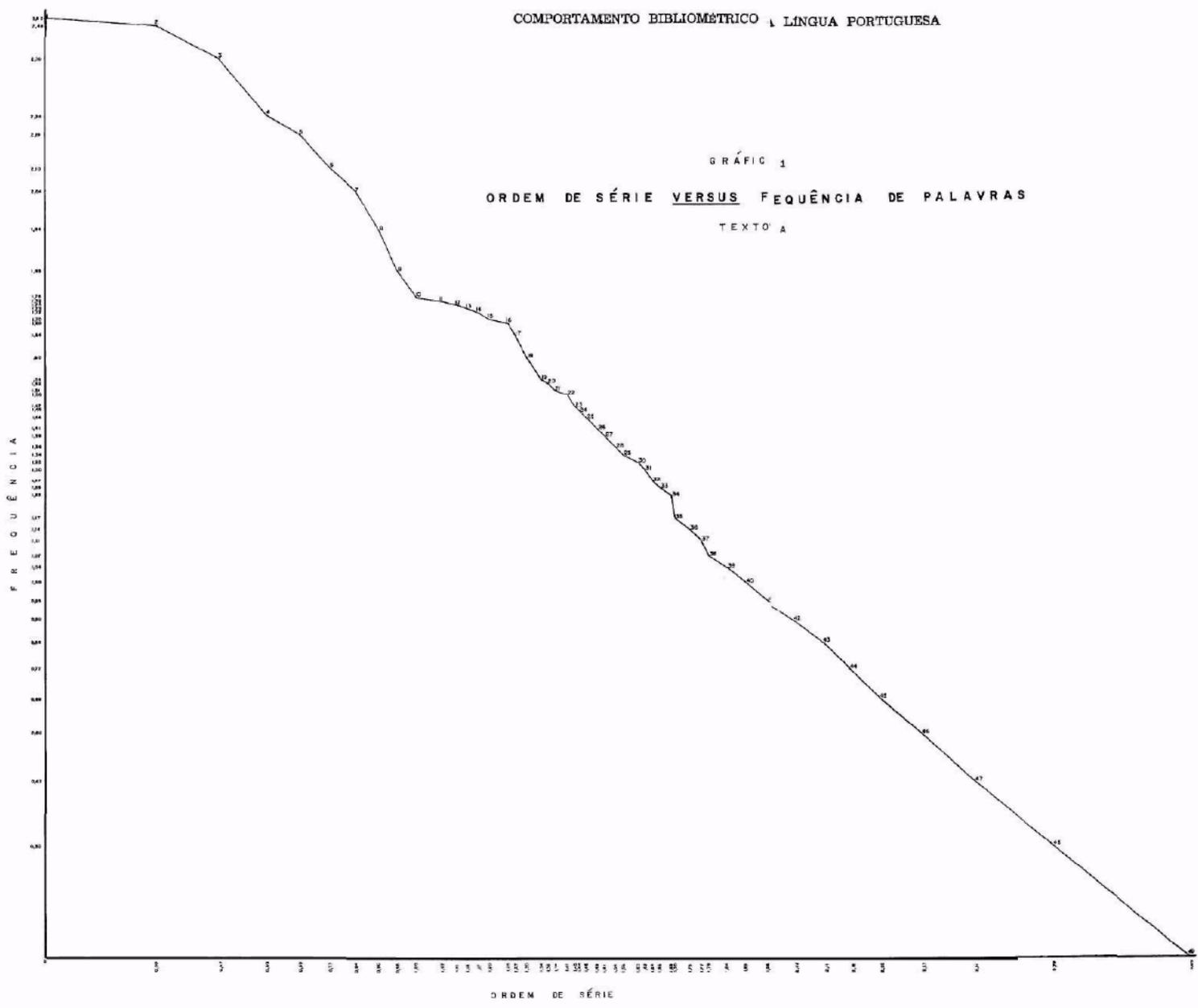
Textos	Palavras diferentes	Palavras que ocorrem 1 vez	$I_1/D$
A	1791	1118	0,62
B	839	542	0,64
C	535	360	0,67

COMPORTAMENTO BIBLIOMÉTRICO 1 LINGUA PORTUGUESA

GRÁFICO 1

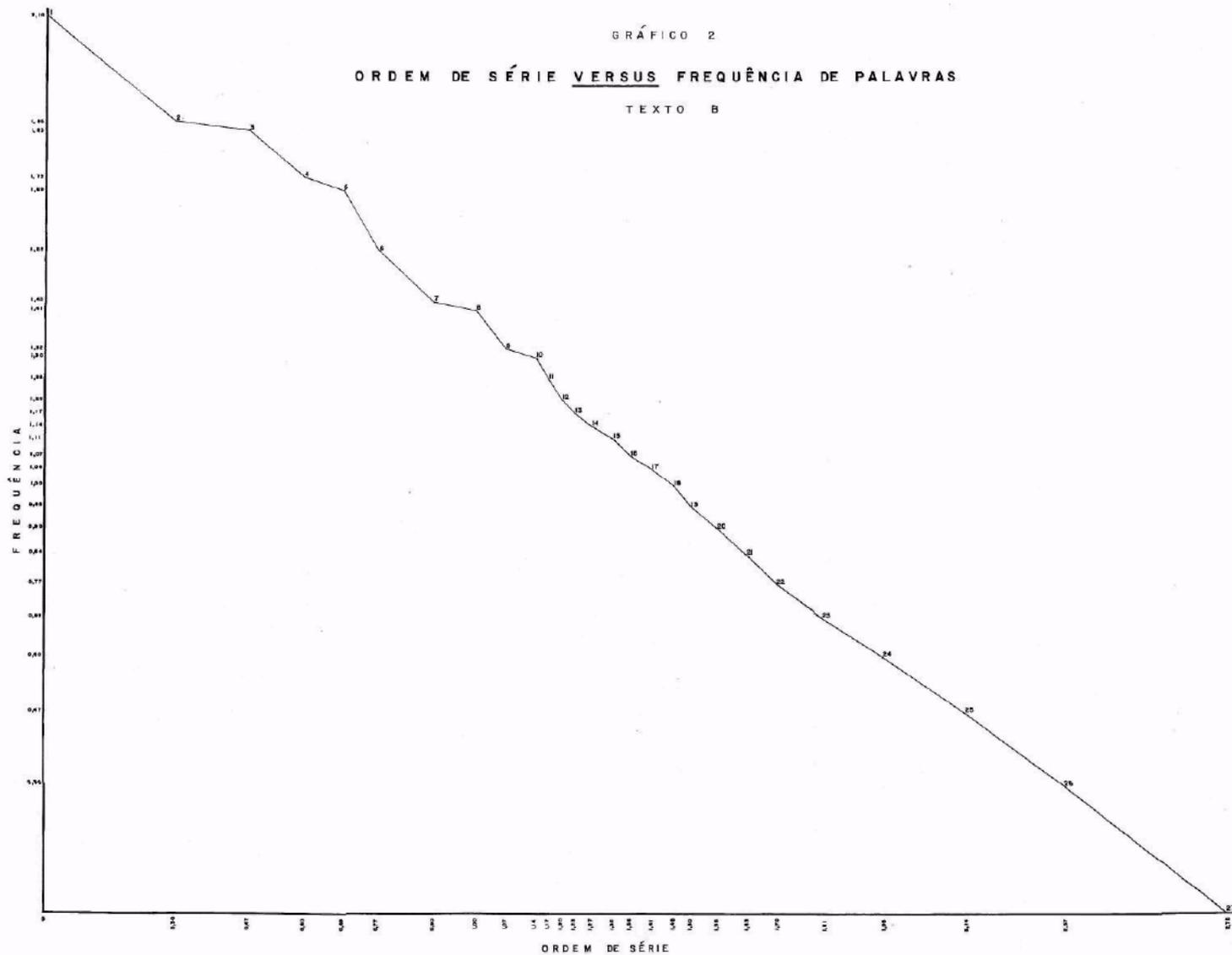
ORDEM DE SÉRIE VERSUS FREQUÊNCIA DE PALAVRAS

TEXTO A



COMPORTAMENTO BIBLIOMÉTRICO DA LÍNGUA PORTUGUESA

GRÁFICO 2  
 ORDEM DE SÉRIE VERSUS FREQUÊNCIA DE PALAVRAS  
 TEXTO B

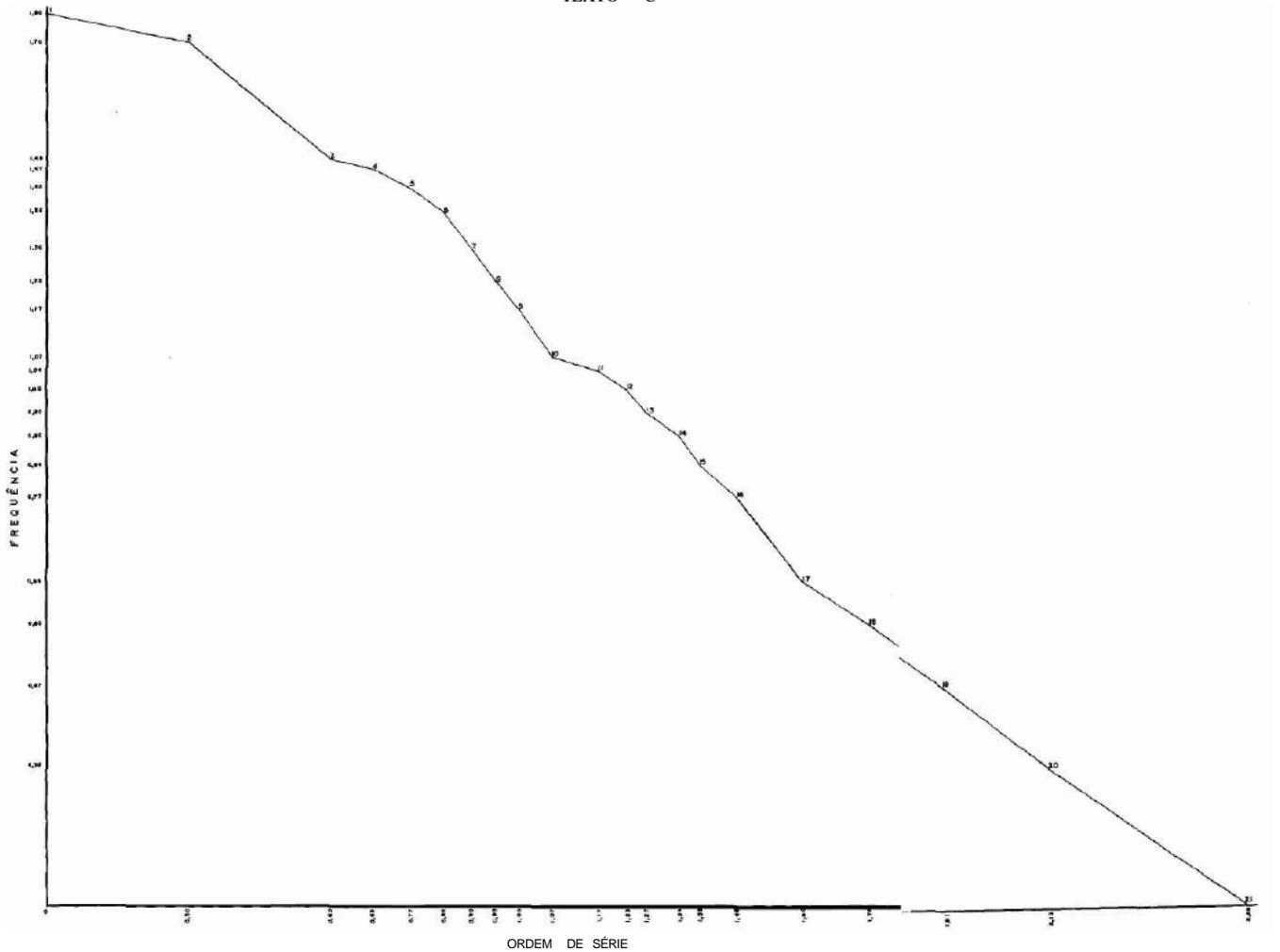


COMPORTAMENTO BIBLIOMÉTRICO DA LÍNGUA PORTUGUESA

GRÁFICO 3

ORDEM DE SÉRIE VERSUS FREQUÊNCIA DE PALAVRAS

TEXTO C



**FÓRMULA F1**

**ZIPF MODIFICADA COM CONSTANTE 2**

**TEXTO A**

$$I_2 = 1118 \frac{2}{2(2+1)} = \frac{2236}{6} = 372$$

$$I_3 = 1118 \frac{2}{3(3+1)} = \frac{2236}{12} = 186$$

$$I_4 = 1118 \frac{2}{4(4+1)} = \frac{2236}{20} = 111$$

$$I_5 = 1118 \frac{2}{5(5+1)} = \frac{2236}{30} = 74$$

**FÓRMULA F1A**

**ZIPF MODIFICADA COM CONSTANTE 1,50**

**TEXTO A**

$$I_2 = 1118 \frac{1,50}{2(2+1)} = \frac{1677}{6} = 279,5$$

$$I_3 = 1118 \frac{1,50}{3(3+1)} = \frac{1677}{12} = 139$$

$$I_4 = 1118 \frac{1,50}{4(4+1)} = \frac{1677}{20} = 83$$

$$I_5 = 1118 \frac{1,50}{5(5+1)} = \frac{1677}{30} = 55$$

**FÓRMULA F 2**

**ZIPF MODIFICADA COM CONSTANTE 2**

**TEXTO B**

$$I_2 = 542 \frac{2}{2(2+1)} = \frac{1084}{6} = 180$$

$$I_3 = 542 \frac{2}{3(3+1)} = \frac{1084}{12} = 90$$

$$I_4 = 542 \frac{2}{4(4+1)} = \frac{1084}{20} = 54$$

$$I_5 = 542 \frac{2}{5(5+1)} = \frac{1084}{30} = 36$$

**FÓRMULA P 2 A**

**ZIPF MODIFICADA COM CONSTANTE 1,50**

**TEXTO B**

$$I_2 = 542 \frac{1,50}{2(2+1)} = \frac{813}{6} = 135$$

$$I_3 = 542 \frac{1,50}{3(3+1)} = \frac{813}{12} = 67$$

$$I_4 = 542 \frac{1,50}{4(4+1)} = \frac{813}{20} = 40$$

$$I_5 = 542 \frac{1,50}{5(5+1)} = \frac{813}{30} = 27$$

**FÓRMULA F 3**

**ZIPF MODIFICADA COM CONSTANTE 2**

**TEXTO C**

$$I_2 = 360 \frac{2}{2(2+1)} = \frac{720}{6} = 120$$

$$I_3 = 360 \frac{2}{3(3+1)} = \frac{720}{12} = 60$$

$$I_4 = 360 \frac{2}{4(4+1)} = \frac{720}{20} = 36$$

$$I_5 = 360 \frac{2}{5(5+1)} = \frac{720}{30} = 24$$

**FÓRMULA F 3 A**

**ZIPF MODIFICADA COM CONSTANTE 1,50**

**TEXTO C**

$$I_2 = 360 \frac{1,50}{2(2+1)} = \frac{540}{6} = 90$$

$$I_3 = 360 \frac{1,50}{3(3+1)} = \frac{540}{12} = 45$$

$$I_4 = 360 \frac{1,50}{4(4+1)} = \frac{540}{20} = 27$$

$$I_5 = 360 \frac{1,50}{5(5+1)} = \frac{540}{30} = 18$$

## FÓRMULA DE TRANSIÇÃO DE GOFFMAN

### TEXTO A

Fórmula T 1 : Constante 2

$$T = 47$$

Fórmula T 1 A : Constante 1,50

$$T = 62,6$$

Fórmula T 1 B : Constante 7

$$T = 13$$

### TEXTO B

Fórmula T 2 : Constante 2

$$T = 32,5$$

Fórmula T 2 A : Constante 1,50

$$T = 43,3$$

Fórmula T 2 B : Constante 7

$$T = 9,2$$

### TEXTO C

Fórmula T 3 : Constante 2

$$T = 26,5$$

Fórmula T 2 A : Constante 1,50

$$T = 35,3$$

Fórmula T 2 B : Constante 7

$$T = 7,57$$