



# **História do Português Brasileiro**

Ataliba T. de Castilho  
(coordenador geral)

VOLUME II

## **CORPUS DIACRÔNICO DO PORTUGUÊS BRASILEIRO**

Ataliba T. de Castilho  
(coordenador)



editora**contexto**



# SUMÁRIO

APRESENTAÇÃO .....	7
<i>Ataliba T. de Castilho</i>	
A PLATAFORMA DE CORPORA DO PHPB: UMA APRESENTAÇÃO AD INFINITUM.....	16
<i>Afranio Gonçalves Barbosa</i>	
EXPERIÊNCIAS METODOLÓGICAS EM CONSTITUIÇÃO DE CORPORA: PISTAS PARA UM PESQUISADOR INICIANTE .....	68
<i>Alexandre Xavier Lima, Leonardo Lennertz Marcotulio e Márcia Cristina de Brito Rumeu</i>	
TRATAMENTO METODOLÓGICO DAS MÃOS INÁBEIS EM CORPORA DIACRÔNICOS .....	92
<i>Huda da Silva Santiago e Zenaide Novais Carneiro</i>	
REFLEXÕES SOBRE A CONSTITUIÇÃO DE AMOSTRAS DERIVADAS A PARTIR DE UM CORPUS.....	120
<i>Nei Rocha e Leonardo Lennertz Marcotulio</i>	
O CONTROLE INDIRETO DE PERFIS SÓCIO-HISTÓRICOS EM CORPORA HISTÓRICO-DIACRÔNICOS: A IDENTIFICAÇÃO DE GRAUS DE LETRAMENTO PELA GRAFIA ETIMOLÓGICA DO SÉCULO XIX.....	168
<i>Afranio Gonçalves Barbosa e Alexandre Xavier Lima</i>	

O CORPUS DO PROJETO PARA A HISTÓRIA DO PORTUGUÊS BRASILEIRO: A CONSTITUIÇÃO DE CORPORA HISTÓRICOS BASEADA EM CRITÉRIOS DE TRADIÇÕES DISCURSIVAS.....	206
<i>Verena Kewitz e José da Silva Simões</i>	
CARACTERÍSTICAS E POTENCIALIDADES DOS CORPORA DO PORTUGUÊS PAULISTA.....	244
<i>Verena Kewitz e José da Silva Simões</i>	
REFLEXÕES SOBRE A CONSTITUIÇÃO E ANÁLISE DE CORPORA LINGÜÍSTICOS HISTÓRICOS E SOBRE A IDENTIFICAÇÃO DE PERFIS SOCIAIS DE REDADORES DO PASSADO.....	290
<i>Tânia Lobo e Zenaide Novais Carneiro</i>	
REFERÊNCIAS BIBLIOGRÁFICAS.....	313
APÊNDICE.....	350
OS AUTORES.....	363

# APRESENTAÇÃO

Ataliba T. de Castilho

Este volume 2 da série *História do Português Brasileiro* é dedicado à organização do *corpus* diacrônico desse projeto. Ele compreende oito capítulos.

Os trabalhos reunidos nesta parte se devem, maiormente, à liderança do prof. Afranio Gonçalves Barbosa, da Universidade Federal do Rio de Janeiro, que coordenou a equipe de *corpus* do projeto em seu todo. Para bem desempenhar essas atividades, o prof. Afranio visitou várias equipes regionais, preparando os pesquisadores locais para a edição filológica dos textos.

Ao mesmo tempo, as equipes regionais se dedicaram ao árduo trabalho de identificação e edição dos documentos de interesse para o projeto *Para a História do Português Brasileiro*. Devem ser lembrados aqui, entre outros pesquisadores, Marcelo Módolo, que escreveu o primeiro trabalho de grau sobre o tema: Módolo (1998).

Para garantir a uniformidade dos materiais, Afranio Barbosa e Marcelo Módolo, entre outros pesquisadores, fixaram as seguintes *Normas de Transcrição de Documentos Manuscritos e Impressos – Edição Semidiplomática*:

1. A transcrição será conservadora.
2. As abreviaturas serão desenvolvidas, marcando-se – em itálico – as letras omitidas e observando-se os seguintes casos:
  - a. A norma não se aplica às abreviaturas hoje em uso corrente ou fixadas em dicionários. Exemplos: “etc.”, “Sr.”, “Sr<sup>a</sup>”, “Ltda.”, “Cia.”, “V. Ex.” e “D.” permanecem inalteradas.
  - b. Respeitar, sempre que possível, a grafia do documento, ainda que manifeste idiosincrasias ortográficas do escriba, como no caso da ocorrência “munto”, que leva a abreviatura “m.<sup>to</sup>” a ser transcrita “munto”.

- c. No caso de variação no próprio documento ou em coetâneos, a opção será para a forma atual ou mais próxima da atual, como no caso de ocorrências “Deos” e “Deus”, que levam a abreviatura “D.” a ser transcrita “Deus”.
3. Não será estabelecida fronteira de palavras que venham escritas juntas, nem se introduzirá hífen ou apóstrofo onde não houver. Exemplos: “epor” “ser”; “aellas”; “daPiedade”; “omninino”; “dosertaõ”; “mostrandoselhe”; “achandose”; “sesegue”.
4. A pontuação original será mantida. No caso de espaço maior intercalar deixado pelo escriba será marcado [espaço]. Exemplo: “que podem perjudicar. [espaço] Osdias passãõ eninguemcomparece”. Serão observados dois casos especiais:
  - a. Em relação a trechos que demandem maior esforço para decodificação, seja pela ausência de sinais de pontuação, seja por estarem sob sistema diverso, o editor incluirá, em nota de rodapé, uma possível interpretação. Exemplo: Na edição teríamos, “Tenho uma criada que | dice que sabia fazer tudo | que eu mandace ella | fazer emtaõ perguntei | e Paõ doce voce sabe fazer | sei emtaõ mando todos | os sabados fazer.” Em nota teríamos, “Nota 1: Tenho uma criada que dice que sabia fazer tudo que eu mandace ella fazer. Emtaõ perguntei: E Paõ doce, voce sabe fazer? ‘Sei’. Emtaõ mando todos os sabados fazer”.
  - b. A sinalização [espaço] não se aplica aos espaços em cabeçalhos, títulos e/ou rótulos de seções de periódicos, fórmulas de saudação/encerramento ou na reprodução de diálogos, devendo o editor estabelecer o intervalo conforme o original.
5. A acentuação original será rigorosamente mantida, não se permitindo qualquer alteração. Exemplos: “aRepublica”; “docommercio”; “edemarcando tambem lugar”; “Rey D. Jose”; “oRio Pirahý”; “oexercicio”; “que hé munto conveniente”; “fomos a ele”; “fomos á ele”; “fomos à ele”. Os sinais de separação de sílaba ou de linha, usados pelos autores dos diversos documentos, serão mantidos como no original. Exemplos: “atira- | mos” e “atira= | mos”.
6. Será respeitado o emprego de maiúsculas e minúsculas como se apresentam no original. No caso de alguma variação física dos sinais gráficos resultar de fatores cursivos, não será considerada relevante. Assim, a comparação do traçado da mesma letra deve propiciar a melhor solução.

7. No caso dos impressos, eventuais erros de composição serão marcados com (*sic*) logo após o vocábulo e remetidos para nota de rodapé, onde se deixará registrada a lição por sua respectiva correção. Exemplo: “Nota 1: Acawado por acamado”. Se for inevitável, por limites do editor de textos, o erro será descrito. Exemplo, “Nota 2: A letra <a> inicial de ‘affirma’ está invertida.”
8. No caso dos manuscritos, eventuais grafias diferenciadas serão remetidas para nota de rodapé, onde se registrará(rão) sua(s) variante(s) mais comum(ns) e, quando possível, considerações sobre a variação em si. Exemplos: “Nota 1: ‘que eu afamado livro’ provavelmente ‘que meu afamado livro’. Talvez a escrita de eu por meu possa indicar lapso de escritura ou erro de cópia”; “Nota 2: Pirassocunda possível variante de Pirassununga: talvez a oscilação de grafia indique instabilidade para a escrita de termos Tupi”.
9. Inserções do escriba ou do copista, para não conferir à mancha gráfica um aspecto demasiado denso, obedecem aos seguintes critérios:
  - a. Se na entrelinha do documento original, entram na edição em alinhamento normal e entre os sinais: < >; <↑>, se na entrelinha superior; <↓>, se na entrelinha inferior. Por exemplo: “em dezembro recebi <↑todos> os senadores em casa”. Se houver palavra(s) riscada(s) abaixo da inserção, deverá haver menção ou, conforme sua legibilidade, transcrição em nota de rodapé. Exemplos: “Nota 1: abaixo de <↑todos> há palavra suprimida”; “Nota 2: abaixo de <↑todos> foi riscado ‘dentre’”.
  - b. Se nas margens superior, laterais ou inferior, entram na edição entre os sinais < >, na localização indicada. Exemplo: <fica definido que olugar convencionado é acasa dePedro nolargo damatriz>. Caso seja necessário, ficará em nota de rodapé a devida descrição da direção de escritura ou quaisquer outras especificidades. Exemplo: “Nota 1: Escrito verticalmente de cima para baixo”.
10. Supressões feitas pelo escriba ou pelo copista no original serão tachadas. Exemplos: “todos ~~ninguem~~ dos presentes assignarom; sahiram ~~sahiram~~ aspressas para oadro”. No caso de repetição que o escriba ou copista não suprimiu, passa a ser suprimida pelo editor que a coloca entre colchetes duplos. Exemplo: “fugi[[gi]]ram correndo [[correndo]] emdireção opaço”.
11. Intervenções de terceiros no documento original devem aparecer em nota de rodapé informando-se a localização. Exemplos: “Nota 1: À

- direita do título encontra-se escrito por outro punho: ‘copiado’; “Nota 2: Na margem inferior encontra-se escrito por outro punho: ‘página 18’”; “Nota 3: Na margem superior encontra-se o carimbado ‘Arquivo Nacional’”.
12. Intervenções do editor hão de ser raríssimas, permitindo-se apenas em caso de extrema necessidade, desde que elucidativas a ponto de não deixarem margem à dúvida. Quando ocorrerem, devem vir entre colchetes. Exemplo: “não deixe passar neste [registro] de Areas”. Quando houver dúvida sobre a decifração de alguma letra, parte de ou vocábulo inteiro, o elemento em questão será posto entre colchetes e em itálico. Exemplos: “ent[re]gue o [rapaz]”; “faça venda a duas b[arric]as de vinho”.
  13. Letra ou palavra(s) não legíveis por deterioração ou rasura justificam intervenção do editor com a indicação entre colchetes conforme o caso: [.] para letras, [ilegível] para vocábulos e [ilegível. + n linhas] para a extensão de trechos maiores. Exemplos: “É assim pe[.]r.”; “É assim [ilegível.] em Java”; “É assim [ilegível + 2 linhas] em Havana.” Caso suponha ser extremamente necessário, o editor indica em nota a causa da elegibilidade: corroído, furo, borrão, rasura, etc.
  14. Letra ou palavra(s) simplesmente não decifradas, sem deterioração do suporte, justificam intervenção do editor com a indicação entre colchetes conforme o caso: [?] para letras, [inint.] para vocábulos e [inint. + número de linhas] para a extensão de trechos maiores. Exemplos: “É assim pe[?]r.”; “É assim [inint.] em Java”; “É assim [inint. + 2 linhas] em Havana.”
  15. A divisão das linhas do documento original será preservada, ao longo do texto, na edição, pela marca de uma barra vertical entre as linhas. A mudança de parágrafo será indicada pela marca de duas barras verticais. Exemplo: “Es- | taes pois muito atrazado, ponde-vos na | pira meu ignorantão. || Seria bonito que todas as.”
  16. A mudança de fólho ou página receberá a marcação entre colchetes conforme o caso:
    - a. Se em documentos manuscritos, com o respectivo número e indicação de frente ou verso. Exemplos: [fol. 1r]; [fol. 1v]; [fol. 2r]; [fol. 2v]; [fol. 3r]; [fol. 3v]; [fol. 16r].
    - b. Se em documentos impressos, indicação de página. Exemplos: [p. 1]; [p. 2]; [p. 3]; [p. 19].
  17. Na edição, as linhas serão numeradas de cinco em cinco a partir da quinta. Essa numeração será encontrada à margem direita da mancha, à esquerda do leitor. Será feita de maneira contínua por documento.

18. Os sinais públicos, diferentemente das assinaturas e rubricas simples, serão sublinhados e indicados entre colchetes. Exemplos: assinatura simples, Bernardo Jose de Lorena; sinal público, [Bernardo Jose de Lorena].
19. Informações que o editor julgar significativas sobre a diagramação e *layout* do texto em impressos devem aparecer em nota de rodapé.

Essas Normas reproduzem o texto “Normas para transcrição de documentos manuscritos para a história do português do Brasil”, publicado em Mattos e Silva (Org. 2001a, tomo 2, pp. 553-555), tendo-se agregado pequenas alterações, resultantes da prática de edição.

Tais Normas vieram encimadas pelas seguintes informações:

A Comissão de estabelecimento de Normas para transcrição de documentos manuscritos para a História do Português do Brasil, formada pelos pesquisadores César Nardelli Cambraia (USP), Gilvan Muller de Oliveira (UFSC), Heitor Megale (USP), Marcelo Módolo (Mestrando – USP), Permínio Souza Ferreira (UFBA), Sílvio de Almeida Toledo (USP), Tânia C. Freire Lobo (UFBA), Valdemir Klamt (UFSC), após apresentação e discussão dos subsídios colhidos nos trabalhos individuais, levou a plenário um elenco de normas que, submetidas a ampla discussão, tiveram como resultado a aprovação das seguintes *Normas para transcrição de documentos manuscritos para a história do português do Brasil* (texto acima).

Posteriormente, realizou-se em João Pessoa, por ocasião do VIII Seminário Nacional do PHPB uma reunião da Comissão de *Corpus*. A ata respectiva foi transcrita no “Apêndice” deste volume.

\*

Seguem-se os resumos dos capítulos.

No primeiro capítulo, “A Plataforma de *corpora* do PHPB: uma apresentação *ad infinitum*”, Afranio Gonçalves Barbosa oferece a memória da construção do *corpus* do Projeto Nacional Para a História do português brasileiro (PHPB) de modo a servir como referência de etapas, questões e operacionalidade para outros empreendimentos de construção de



*corpora* histórico-linguísticos com grande cobertura diatópica. São apresentadas as opções metodológicas de base que definiram os controles externos na constituição do *corpus mínimo comum* do projeto e do *corpus diferencial*, material contrastivo inserido na *Plataforma de corpora do PHPB* na internet, que cataloga o conjunto total de edições organicamente realizadas sob a chancela do PHPB.

No segundo capítulo, “Experiências metodológicas em constituição de *corpora*: pistas para um pesquisador iniciante”, Alexandre Xavier Lima, Leonardo Lennertz Marcotulio e Márcia Cristina de Brito Rumeu trazem várias reflexões sobre a árdua tarefa da constituição de *corpora* diacrônicos. Com base na experiência em trabalhos anteriores (Lima, 2010; 2014; Marcotulio, 2010; Rumeu, 2004; 2013), os autores pretendem contribuir para a otimização do trabalho de quem inicia uma pesquisa em arquivos, constituindo um *corpus* para a sua investigação.

No capítulo seguinte, “Tratamento metodológico das *mãos inábeis* em *corpora* diacrônicos”, escrito por Huda da Silva Santiago e Zenaide Novais Carneiro, são tratadas questões tais como a identificação da inabilidade em escrita alfabética nos planos supragráfico e paleográfico, da segmentação gráfica, da escriptualidade, da escrita fonética e da repetição lexical. A seguir, são examinadas as cartas de sertanejos no século xx.

Em “Reflexões sobre a constituição de amostras derivadas a partir de um *corpus*”, os autores Nei Rocha e Leonardo Lennertz Marcotulio centram-se no *corpus histórico do português*, propondo uma reflexão sobre a constituição de amostras derivadas a partir de uma amostra base, tanto do ponto de vista da disponibilidade e comparabilidade de materiais quanto do ponto de vista estatístico, com o intuito de discutir a possibilidade de obtenção de amostras mínimas significativas que conservem o potencial de representatividade do *corpus*.

O capítulo “O controle indireto de perfis sócio-históricos em *corpora* histórico-diacrônicos: a identificação de graus de letramento pela grafia etimológica do século xix”, escrito por Afranio Gonçalves Barbosa e Alexandre Xavier Lima, oferece caminhos metodológicos para controle indireto de perfis sociais de redatores de manuscritos. Para o século xix, adotou-se