

*Sobre o uso do texto escrito em linguística histórica:  
mitos antigos e práticas modernas*

On the use of written texts in historical linguistics:  
ancient myths and modern practices

Ana Paula Banza

Universidade de Évora, Portugal

DOI: <https://doi.org/10.5902/2176148542435>

**Resumo:** O presente artigo discute o uso de textos escritos como *corpora* em Linguística Histórica, com foco em algumas ideias que a tradição tem perpetuado, nomeadamente sobre a adequação dos dados da escrita e dos diferentes tipos de fonte e de edição em estudos sobre fases pretéritas da língua. Essas ideias são aqui revistas e analisadas tendo em conta a peculiar especificidade desta disciplina e à luz das alterações trazidas pela era digital à teoria e à prática editorial, propondo-se, como resultado, uma redefinição do objecto primário da Linguística Histórica e uma diversificação dos tipos de fonte e de edição.

**Palavras-chave:** Linguística Histórica. Humanidades digitais. Texto escrito. Edição.

**Abstract:** This article discusses the use of written texts as *corpora* in Historical Linguistics, focusing on some ideas that tradition has perpetuated, namely on the adequacy of writing data and of the different types of sources and editing in studies on past phases of the language. These ideas are here reviewed and analyzed taking into account the peculiarity of this discipline and in the light of the changes brought by the digital age to editorial theory and practice. As a result, a redefinition of the primary object of Historical Linguistics and a diversification of the types of sources and of editing are proposed.

**Keywords:** Historical Linguistics. Digital humanities. Written text. Edition.

## 1. Introdução

Se há área com uma tradição bem estabelecida desde os primórdios da Linguística, como ciência da linguagem, essa é seguramente a Linguística Histórica, que, desde o início do séc. XIX, se estabeleceu com um objecto – o estudo do passado das línguas e do seu processo de mudança ao longo do tempo – e método – o comparativo – bem definidos. Corria, então, o “século da História” e o impacto da publicação do famoso tratado de Darwin sobre *A Origem das Espécies* tornava natural a procura de um paralelo entre a evolução das espécies e a evolução da linguagem. Não é, pois, de estranhar que, no clima científico e intelectual da época, a única Linguística possível fosse a histórica.

O séc. XX, porém, traria alterações teóricas e metodológicas profundas, usualmente atribuídas ao legado de Saussure, deslocando o objecto da Linguística do estudo da mudança (diacronia) para a descrição das línguas, individualmente consideradas num ponto de tempo definido (sincronia). Na base desta alteração de foco encontrava-se a visão saussureana das línguas como estruturas – sistemas em que cada item é definido pelas relações que mantém com todos os outros – necessariamente homogêneas e, como tal, incompatíveis com as instâncias de variação, quer diacrónicas, quer sincrónicas.

Efectivamente, para Saussure, como já antes para Hermann Paul (1880), sistematicidade e variabilidade excluía-se, tornando-se necessária aos objectivos da nova Linguística descritiva a abstracção, voluntária e consciente, da real heterogeneidade das línguas. Por outro lado, a constituição do paradigma sincronia/diacronia como uma “dualidade radical”, implicando “dois caminhos absolutamente divergentes” na Linguística (SAUSSURE, 1992, p. 141), consagrou a opção da Linguística estrutural (nas diferentes formas que assumiu, ao longo da primeira metade do séc. XX, na Europa e na América), e também da Linguística generativa de Chomsky, pela sincronia, cujos processos são observáveis na actualidade, bem como o mito da homogeneidade da língua.

Como consequência destas opções, estabeleceu-se um “paradoxo infrutífero” (WEINREICH, LABOV e HERZOG, 1968, p. 150) entre funcionamento e estrutura, variabilidade e sistematicidade, que a Sociolinguística laboviana tem procurado resolver:

Vimos que, para [Hermann] Paul como para Saussure, variabilidade e sistematicidade excluem-se mutuamente. Os seus suces-

sores, que continuaram a postular mais e mais sistematicidade na língua, tornaram-se ainda mais profundamente comprometidos com uma concepção simplista do dialecto homogéneo [...] No entanto, muitos linguistas reconhecem a evidência que demonstra que a mudança da língua é um processo contínuo e o inevitável sub-produto da interacção linguística. Os paradoxos têm sido profundamente sentidos [...] mas poucos linguistas podem ficar satisfeitos com uma explicação da mudança que depende em simultâneo da inobservabilidade de todos os processos envolvidos. Sugerimos que a solução para a questão fundamental reside na direcção de quebrar a identificação entre estrutura e homogeneidade (WEINREICH, LABOV e HERZOG, 1968, p. 150-151. Sublinhados e tradução nossos<sup>1</sup>).

*Sobre o uso do  
texto escrito  
em linguística  
histórica*

---

35

Ao propor a quebra da relação entre estrutura e homogeneidade, através da abordagem sistemática de estruturas heterogéneas, e ao correlacionar variação e mudança, a Teoria da Variação e Mudança (WEINREICH, LABOV e HERZOG, 1968) veio também permitir o rompimento da “dualidade radical” (SAUSSURE, 1992, p. 141) entre sincronia e diacronia, aproximando-as e integrando-as e trouxe de volta ao panorama da Linguística a Linguística Histórica, doravante encarada como complementar em relação à Linguística Sincrónica.

As teorias contemporâneas da mudança incluem também, além da variacionista, a funcionalista, na linha do estruturalismo diacrónico, e a generativista, todas elas bastante produtivas. No entanto, pode dizer-se que é graças, sobretudo, aos desenvolvimentos da Sociolinguística variacionista, a partir dos anos 1960, e, mais tarde, também do modelo generativista paramétrico, nos anos 1980 – que,

considerando a variação interlinguística, trouxe à especulação da teoria questões fundamentais em que se interrelacionam a aqui-

---

1 “We have seen that for [Hermann] Paul as well as for Saussure, variability and systematicity excluded each other. Their successors, who continued to postulate more and more systematicity in language, became ever more deeply committed to a simplistic conception of the homogeneous idiolect [...] Yet most linguists acknowledge the evidence which demonstrates that language change is a continuous process and the inevitable by-product of linguistic interaction. The paradoxes have been deeply felt [...] but few linguists can rest content with an explanation of change which depends upon the joint unobservability of all the processes involved. We have suggested that the solution to the fundamental question lies in the direction of breaking down the identification of structuredness with homogeneity.”

sição – questão sempre presente, desde as origens, nos modelos chomskyanos – à mudança lingüística, até então questão secundária nos gerativismos anteriores (MATTOS e SILVA, 1999, p. 148).

– que a Linguística Histórica, incluindo os estudos de dialectos sincrónicos, em particular os mais arcaizantes, voltou a ocupar lugar de relevo nos estudos lingüísticos.

Por outro lado, tendo em conta o grande desenvolvimento do modelo generativista paramétrico – que se centra na *I-Language*, ao contrário das vertentes mais tradicionais da Linguística Histórica e da Filologia, que se centram na *E-Language* – e os mais recentes desenvolvimentos da teoria laboviana (cf. LABOV, 2010), tornou-se pertinente, como sugere Mattos e Silva (1999, p. 151), distinguir “*Lingüística Histórica lato sensu*, que inclui descrições e interpretações sincrônicas datadas e localizadas, *Lingüística Histórica stricto sensu*, que se concentra na mudança lingüística no tempo, levando em consideração fatores intralingüísticos ou estruturais e fatores extralingüísticos ou sócio-históricos e *Lingüística Diacrônica*, que, tratando da mudança no tempo, se concentra no sistema ou na gramática”.

Em todos os casos, porém – sem prejuízo da possibilidade de “reconstrução dos estados passados baseada na comparação entre variedades modernas” (CASTRO, 2006, p. 78) e apesar de, desde Labov, se considerar que o conhecimento de processos que operaram no passado pode ser inferido pela observação de processos no presente – a fonte por excelência, quer na Linguística Histórica, quer na Linguística Diacrônica (cf. *supra*) continua a ser, sem dúvida, o texto escrito, único documento do passado das línguas até ao séc. XX.

Tomando este facto, que se mantém actual, como ponto de partida, analisam-se aqui alguns dos problemas teóricos e metodológicos levantados pelo uso de textos escritos como fontes em Linguística Histórica, com destaque para a adequação dos dados da escrita e de determinados tipos de fontes e de edição para os estudos lingüísticos. Estas questões são discutidas a partir do posicionamento das diferentes teorias da mudança, procurando-se oferecer uma nova perspectiva sobre o uso do texto escrito como fonte em Linguística Histórica com base numa reavaliação, por um lado, das relações oral/escrito e, por outro, da relevância dos diferentes tipos de fontes escritas e de edição no peculiar contexto da Linguística Histórica. Procura-se ainda evidenciar o contributo do meio digital na resolução de alguns dos problemas discutidos.

Ana Paula  
Banza

---

36

## 2. O texto escrito como *corpus*: alguns mitos

Desde a sua origem, a Linguística Histórica assumiu a inevitabilidade do uso de textos escritos como fonte. No entanto, foi sobretudo no séc. XX que este condicionalismo começou a ser encarado como um problema metodológico de base. Com o desenvolvimento das novas tendências da Linguística dita “moderna”, e tendo em conta que esta assentava numa separação radical, não apenas entre sincronia e diacronia, mas também entre os modos de representação oral e escrito das línguas e na primazia do primeiro (primário, dinâmico, não regulamentável) sobre o segundo (secundário, estático, regulamentável), a língua oral tornava-se o único objecto possível e, nessa medida, os textos escritos, apesar de também serem manifestações de língua, não poderiam, naturalmente, ser considerados fontes fiáveis, na medida em que não eram encarados como fontes para o estudo da língua escrita, mas como fontes para o estudo da língua oral. Mesmo Labov, que logrou superar a identificação entre estrutura e homogeneidade e a separação radical entre sincronia e diacronia, não superou a separação entre oralidade e escrita e, sobretudo, a primazia do oral sobre o escrito na Linguística actual. Ao considerar a Linguística Histórica como “a arte de fazer o melhor uso de maus dados” (LABOV, 1994, p. 11), Labov perpetuaria também o preconceito em relação à escrita, considerada, a todos os títulos, secundária. A esta ideia, veio juntar-se uma outra, que a tradição tem também perpetuado, a de que, no estudo do passado das línguas e dos processos de mudança, sendo todas as fontes más, as fontes primárias não literárias são, ainda assim, preferíveis às literárias e às metalinguísticas, por estarem, em princípio, mais próximas da oralidade.

Segundo Labov (*idem, ibidem*), os dados fornecidos pelos textos escritos, única face observável da língua em sincronias pretéritas, ainda que ricos em muitos aspectos, são necessariamente maus, na medida em que os documentos históricos estão sujeitos a todo o tipo de acidentes e as formas linguísticas que neles encontramos são, regra geral, distintas das vernaculares dos escritores, resultando, muitas vezes, de um esforço de adaptação à norma que conduz a textos onde são frequentes fenómenos como a hipercorreção e a mistura de dialectos, além de banais erros de cópia. Tudo isto é, naturalmente, verdade no que respeita à reconstrução da oralidade.

Não restam [...] dúvidas de que o texto, produto escrito de um escritor concreto, num determinado estilo, numa dada situação e num determinado momento, isto é, num determinado contexto – que a actividade filológica do editor procura reproduzir – não pode, no que à língua diz respeito, ser encarado como documento com valor absoluto de um estado pretérito da língua. Pelo menos, não pode ser encarado como documento de um estado de língua no seu modo de representação primário: o oral. Pode, sim, ser encarado, antes de mais, como um documento do sistema de escrita de uma determinada língua numa determinada época. Como documento para a reconstrução da oralidade, constitui, como não poderia deixar de ser, matéria-prima a ser usada de forma muito condicionada, devendo o linguista histórico, se quiser ser realista, renunciar a obter um retrato fiel da oralidade de fases pretéritas da língua (BANZA, no prelo. Sublinhado nosso).

Creemos, pois, que a visão negativa do uso de textos como fontes na Linguística Histórica e Diacrónica só pode ser ultrapassada se se reconhecer como utópico o objectivo de reconstituição da oralidade das línguas em estados pretéritos e, por outro lado, se reconhecer, realisticamente, que o objecto possível nestas disciplinas é, em primeiro lugar, a reconstituição da língua escrita de estados pretéritos, renunciando a “explicar” e reconhecendo que o máximo que é possível fazer é apenas “interpretar”, ou procurar “ouvir o inaudível” (LASS, 1997, p. 45ss), como lucidamente reconheceu Lass, que considera a Linguística Histórica como uma arte interpretativa. A metáfora de Lass – presente, por exemplo, na última obra de Mattos e Silva, de 2008: *Caminhos da Linguística Histórica – ouvir o inaudível* – é, aliás, modelar, na medida em que consagra a impossibilidade objectiva de “ouvir” a língua do passado, “inaudível” porque acessível apenas na sua vertente escrita, que, como reconhece o mesmo autor, embora represente a língua oral, o faz indirectamente e segundo princípios próprios, não permitindo, por isso, mais do que a construção de hipóteses sobre a oralidade de estados pretéritos da língua. A mesma ideia encontra-se, aliás, também noutros autores, como, por exemplo, Clarinda Maia, que, a propósito dos textos notariais medievais por ela analisados, escreve que,

embora os documentos notariais de carácter particular não sejam um espelho fiel, uma reprodução da linguagem local, eles deixam transparecer certos factos da linguagem falada da época. Uma interpretação crítica desses documentos [...] permit[e] compreender a verdadeira natureza da língua escrita dos documentos não literários da Idade Média e manifesta ser impossível uma plena reconstrução dos estados pretéritos da língua” (MAIA, 1986, p. 950. Sublinhado nosso).

*Sobre o uso do  
texto escrito  
em linguística  
histórica*

---

39

Importa, assim, reconhecer a escrita como o primeiro objecto da Linguística Histórica, devendo, neste caso, a oralidade ser assumida como um objecto necessariamente secundário, uma vez que só pode ser accedido de forma indirecta e necessariamente imperfeita. Nesta perspectiva, os dados da escrita são naturalmente bons quando usados para a descrição da escrita e não da oralidade e os diferentes tipos de texto igualmente relevantes para diferentes registos, que, no seu conjunto, embora não representem a língua oral, representam a língua escrita das diferentes épocas históricas e, indirectamente, reflectem a língua oral, por via da relação que se estabelece entre estes dois modos de representação das línguas, em particular nas que usam sistemas alfabéticos de escrita.

Nesta perspectiva, a própria construção de hipóteses sobre o que poderá ter sido a oralidade de fases pretéritas de uma língua, na sua diversidade, pode, sem dúvida, ser mais conseguida pela utilização de diferentes tipos de fonte e não apenas das fontes primárias não literárias; nomeadamente, também das fontes literárias pertencentes a diferentes géneros textuais (BERLINCK, BARBOSA e MARINE, 2008) e, sobretudo, das fontes metalinguísticas, cujo valor, apesar de reconhecido em alguns casos (CASTRO, 2006, p. 83), tem sido muito, e injustamente, subestimado (SÁEZ RIVERA, 2006; BANZA, 2019).

Em todo o caso, ultrapassado o mito dos “maus dados” por via de uma reformulação realista dos objectivos a atingir/atingíveis, a Linguística Histórica lida ainda com um outro problema: o do tratamento das fontes disponíveis, que envolve aquilo a que Ivo Castro chamou “filtração filológica” e na qual participam, além da Crítica Textual, outras disciplinas auxiliares, sobre as quais a Crítica Textual se constrói, tais como a Paleografia, a Diplomática, a Codicologia e a Bibliografia Material (CASTRO, 2006, p. 81-82), ao serviço da produção de edições que possam ser consideradas confiáveis para estudos de natureza linguística.

Não restam dúvidas de que todas as edições de textos do passado, e não apenas aquelas que se destinam a estudos linguísticos, devem assumir o princípio filológico básico do respeito pelo texto, o que nem sempre acontece, por erro humano ou por limitações práticas. No entanto, sendo a edição, também ela, por natureza, uma prática interpretativa, são possíveis, sem prejuízo do respeito pelo texto, diferentes opções “estratégicas e táticas” (CASTRO e RAMOS, 1986) no que toca ao tipo e critérios de edição. Assim, da edição diplomática ou paleográfica às edições interpretativa, crítica e críticogenética e das transcrições mais conservadoras às mais modernizadoras, são possíveis diferentes opções que o editor, desejavelmente um linguista, deve analisar. No entanto, e apesar da legitimidade dos diferentes tipos de edição e de transcrição, tem prevalecido a ideia de que a edição diplomática ou paleográfica é a única adequada a textos usados como *corpora* linguísticos, devendo os critérios de edição ser sempre maximamente conservadores.

É esta também uma ideia que importa desmistificar, na medida em que, dependendo do tipo e época dos textos e dos objectivos do estudo, são possíveis, e mesmo desejáveis, diferentes tipos de edição e diferentes critérios de transcrição. Mesmo nos textos antigos, que colocam problemas muito particulares, é possível e desejável dispor de diferentes tipos de edição e transcrição, que poderão ser escolhidos em função dos objectivos do estudo. Neste aspecto, é modelar a proposta de Emiliano (2002a), que sugere, para o texto medieval, uma edição de base maximamente conservadora, com tipo medieval e transcrição estreita e, a partir dessa, edições progressivamente mais modernizadoras: tipo medieval com transcrição larga, tipo normal com transcrição larga e edição interpretativa (modernizadora), garantindo, assim, a utilidade destas transcrições para todos os tipos de análise, linguística ou outra, desde as de natureza fonético-fonológica e/ou grafemática às de natureza sintáctica ou histórica, por exemplo.

Finalmente, um último “mito” que importa à Linguística Histórica ultrapassar definitivamente é aquele a que poderíamos chamar a “morte da Filologia”, que fez caminho em alguns círculos, ao longo do séc. XX, segundo o qual, em Linguística, apenas a forma do texto interessaria, devendo a edição resumir-se a uma mera transcrição, independente do contexto histórico-cultural da sua produção.

Em 1911, nas *Lições de Philologia Portuguesa*, Leite de Vasconcellos definia ainda Filologia como “o estudo da nossa lingua em toda a sua amplitude, e o dos textos, em prosa e em verso, que servem para a docu-

mentar” (p. 9), perspectiva alargada em que a Crítica Textual é apenas uma das vertentes, a mais clássica, da Filologia. Porém, a progressiva especialização dos estudos linguísticos e literários em disciplinas separadas e estanques, ao longo do séc. XX, levou a um esvaziamento da Filologia, que progressivamente se foi restringindo à Crítica Textual, cuja “orientação última” seria a de “preparar a edição do texto” (MARQUILHAS, 2010). Uma das consequências mais dramáticas deste afunilamento foi o afastamento dos linguistas do trabalho de edição.

Esta tendência, no entanto, tem vindo a ser revertida, nos últimos anos, sendo possível afirmar que, actualmente, “a Filologia caracteriza-se por sustentar uma investigação de natureza interdisciplinar, que envolve a Crítica Textual, a Crítica Genética, a Bibliografia Textual, a Sociologia do Texto, a Linguística, os Estudos Literários e a História Cultural.” (BORGES *et alii* 2012, p. 25) e, por outro lado, a Crítica Textual assume também uma interdisciplinaridade que renova e actualiza a antiga Filologia. Como bem referiu Pérez Priego (1997, p. 36. Tradução nossa<sup>2</sup>),

...os testemunhos são efectivamente indivíduos históricos, com uma fisionomia própria, portadores, no seu seio, muitas vezes, de eloquentes marcas e dados sobre onde foram compostos, quem os encomendou ou possuiu, quem foram os copistas, os impressores, os leitores, que tipo de papel e de letra foi utilizado, qual a oficina tipográfica, etc. Tudo isto nos proporciona uma informação muito interessante, seguramente, para a história cultural, mas também muito rica e aproveitável do ponto de vista da pura crítica textual.

Por essa razão, uma boa edição, adequada a estudos linguísticos, deve ser o resultado de um trabalho interdisciplinar, no qual avulta a importância do linguista, que possa devolver ao investigador, bem como também ao simples leitor, o texto enquanto objecto cultural, resultado de uma língua, de uma cultura e de um conjunto de particularidades específicas do género e do autor, bem como de determinadas circunstâncias históricas e sociais.

---

2 “...los testimonios son efectivamente individuos históricos, con una fisionomía propia, portadores en su seno muchas veces de elocuentes huellas y datos respecto de dónde se compusieron, quién los encargó o poseyó, quiénes fueron los copistas, los impresores, los lectores, qué tipo de papel y de letra fue utilizado, qué taller tipográfico, etc. Todo ello nos proporciona una información muy interesante, por supuesto, para la historia cultural, pero también muy rica y aprovechable desde la pura crítica textual.”

### 3. Ultrapassando mitos antigos: práticas de edição na era digital

A chegada da era digital ao domínio da actividade editorial, por volta dos anos 1960 do século passado, desencadeou uma mudança de paradigma que continua em curso e cujos desenvolvimentos futuros não são totalmente previsíveis. No entanto, é certo que, apesar das naturais dificuldades de diálogo entre áreas tradicionalmente não comunicantes, a “revolução digital” tem trazido enormes vantagens, hoje geralmente reconhecidas, entre outros, pelos linguistas, permitindo, no caso do tratamento de fontes para a história das línguas, ultrapassar alguns dos mitos acima referidos.

Assim, antes de mais, tornou-se possível a constituição de *corpora* digitais, que permitem, por exemplo, o acesso fácil a uma grande quantidade de textos, de diferentes géneros, de um mesmo período histórico, facilitando, assim, a diversificação das fontes; mas também a constituição de arquivos, que permitem armazenar um enorme manancial de informação sobre cada um deles, bem como diferentes tipos de transcrição, que, por seu turno, podem também, agora, beneficiar de recursos como o “tipo medieval” (fonte *Notator Mono*) – proposto por Emiliano (2002a, 2002b, 2005, 2007) para reduzir ao mínimo as operações de transliteração nas transcrições mais conservadoras de documentos medievais – ferramentas de leitura automática, como o OCR (*Optical Character Recognition*) – que, apesar das limitações que ainda apresenta, nomeadamente no que respeita à leitura de manuscritos, auxilia a tarefa de transcrição, ou o “software *eDictor* [...] que inclui tanto a etapa da edição filológica, como a etapa de anotação morfossintática automática [...] que [...] permite a preservação das diversas etapas de interferência sem prejuízo da possibilidade de inúmeras formas de representação” (PAIXÃO de SOUSA, 2013, p. 126), entre muitas outras actualmente disponíveis.

Por outro lado, no caso de textos particularmente complexos, na sua génese e/ou tradição, as edições de hipertexto (FAULHABER, 1991; LAVAGNINO, 1995; DUARTE 1997) facilitaram ou tornaram possível a elaboração de edições críticas, crítico-genéticas e genéticas, que, no meio impresso, devido ao número frequentemente elevado de documentos a considerar (diferentes versões, cópias, notas, aditamentos, etc.) eram impossíveis ou, quanto muito, concretizáveis apenas sob a forma de pesadas e dispendiosas edições sinópticas, com extensos e complexos aparatos críticos. Todas estas alterações, resultantes da ins-

trumentalização da informática ao serviço da Filologia, repercutiram-se, pois, antes de mais, no modo de fazer, prevenindo o erro humano e o abandono de tarefas demasiado complexas ou pesadas.

No entanto, a chegada do digital trouxe também alterações mais profundas, ao conceito de edição, mas também ao próprio conceito de texto. Desde os anos 1960, as edições de hipertexto “coloca[m] em cena conexões e relações e, assim fazendo, muda[m] a maneira como os textos existem, mas também aquela como lemos” (FAULHABER, 1991, p. 129) e, actualmente, as novas “edições académicas digitais”, entendidas como recursos de informação que oferecem uma representação crítica dos textos (SAHLE, 2014), permitem visualizar todos os materiais disponíveis, incluindo diferentes tipos de transcrição e imagens de alta resolução dos textos, bem como as relações entre eles. Como resultado, a edição deixa de poder ser pensada como mero resultado da interpretação do editor, apresentada ao público sob uma forma acabada; torna-se colaborativa, flexível e evolutiva. Por outro lado, o próprio conceito de texto é também alterado, na medida em que a materialidade que tradicionalmente o define se torna dinâmica. Como refere Paixão de Sousa (2013, p. 115, 127ss), o texto é “descorporificado”, isto é, verifica-se uma “desvinculação entre o texto lógico e seu suporte material” (PAIXÃO de SOUSA 2013, p. 129. Ver também Emiliano, 2005; Paixão de Sousa, 2006).

Com a formação de “comunidades de práticas abertas” (ALVES, 2016) – como a do TEI (*Text Encoding Initiative*), “um consórcio que desenvolve e mantém coletivamente um padrão para a representação de textos em formato digital” (<https://tei-c.org/>. Tradução nossa<sup>3</sup>) cujas “diretrizes [...] definem e documentam uma linguagem de marcação para representar as características estruturais, de interpretação e conceituais dos textos” (<https://tei-c.org/guidelines/>. Tradução nossa<sup>4</sup>) – as edições académicas apresentam hoje, necessariamente e cada vez mais, uma “codificação explícita e consistente [que permite a] legibilidade por máquina, ao mesmo tempo que permitem, fruto da linguagem de marcação e da anotação rica que adotam, uma crescente manipulação das suas representações computacionais” (MARQUILHAS e HENDRICKX, 2016, p. 252).

---

3 “...a consortium which collectively develops and maintains a standard for the representation of texts in digital form”.

4 “guidelines [...] define and document a markup language for representing the structural, rendition, and conceptual features of texts”.

Neste contexto, os novos modelos de edição não lidam já, hoje, com os tradicionais problemas de fixação e de opção entre diferentes normas de transcrição, graças à “fidelidade e [...] reversibilidade nas edições assistidas por computador” (PARKINSON, 1983, p. 242–243); não lidam também já com as dificuldades de representação de determinadas características dos textos, nomeadamente dos textos medievais, como o tipo medieval ou as abreviaturas (EMILIANO, 2002b; PARKINSON, 1983); lidam, sim, com o imperativo da legibilidade computacional e da manipulabilidade, devendo, por isso, ser explícitos e consistentes e, por outro lado, com a necessidade de “converter as edições em abordagens integradas, onde se codifique o que no texto é físico, o que é conceptual e o que é social (MARQUILHAS e HENDRICKX, 2016, p. 263). Tal objetivo só pode ser conseguido pela articulação interdisciplinar; mas isso não é uma novidade, porque, desde sempre, a edição, entendida como actividade filológica por excelência, resultou da conjugação de diferentes saberes, de diversas áreas disciplinares, sendo a única novidade a entrada das ciências da computação no conjunto destes saberes; e é naturalmente aí que reside o grande desafio dos filólogos modernos: o de tentar conciliar a velocidade inerente às novas tecnologias “com o ritmo pausado da interpretação, ou exegese, método incontornável na construção de conhecimento em humanidades” (MARQUILHAS e HENDRICKX, 2016, p. 275).

#### 4. Conclusões

Como temos vindo a observar, a Linguística Histórica tem padecido de alguns mitos, que têm sido vistos como fragilidades teóricas e/ou metodológicas, pondo em causa o seu lugar na Linguística. Tais mitos resultam, em grande parte, da peculiar especificidade envolvida no estudo do passado das línguas e da sua evolução. Como sublinhou Neto (1992, p. 52),

a história de uma língua não é um esquema rigorosamente pre-estabelecido, não é um problema algébrico. [...] A evolução [...] é complexa e melindrosa, relacionada com mil e um acidentes, cruzada, recruzada e entrecruzada – porque não representa a evolução de uma coisa feita e acabada, mas as vicissitudes de uma atividade em perpétuo movimento.

A afirmação do autor é, na verdade, aplicável à própria língua, em todos os seus aspectos, sincrónicos e diacrónicos, e não apenas à sua história. No entanto, quando estão em causa estádios antigos e a própria evolução das línguas, esse facto é ainda mais sensível. Não pode, por isso, a Linguística Histórica, como reconheceu Labov, conformar-se com o domínio exclusivo de visões sistematicistas e homogeneizadoras, por muito produtivas que estas possam ser, nem, pela enorme complexidade que envolve, com uma única teoria ou método. Não pode, finalmente, conformar-se com a aceitação de que os dados disponíveis no seu campo de actuação são maus, devendo antes perspectivá-los no sentido de que, tendo as línguas, em geral, dois modos de representação, oral (ou gestual) e escrito, a língua escrita é um objecto em si, que, pela relação que mantém com a oralidade, nos sistemas alfabéticos, permite também, a partir da sua descrição e análise, a elaboração de hipóteses, e não mais do que isso, sobre o que teria sido a língua oral em estados pretéritos e sobre a sua evolução.

Nesta perspectiva, a questão da disponibilização de textos fonte para o estudo da língua, escrita e oral, e da sua evolução mantém toda a sua relevância; mas, também aqui, importa ultrapassar alguns mitos, diversificando as fontes, mas também os tipos de edição e os critérios de transcrição e contextualizando os textos, enquanto objectos de língua, mas também de cultura, em sentido lato.

Neste aspecto, os rápidos avanços das Humanidades Digitais permitem, cada vez mais, encontrar no meio digital um precioso aliado, que não pode já ser ignorado. Como sublinham Marquilhas e Hendrickx (2016, p. 259-260), depois de uma primeira fase em que o digital veio trazer novas ferramentas, essencialmente agilizadoras do processo de edição, permitindo tarefas até então impensáveis e bastante mais sujeitas ao erro humano, está já em marcha uma segunda fase, na qual “estão a ser os próprios académicos a harmonizar-se com o digital, a envolver-se nele de forma a compreenderem quantos conceitos é preciso revolucionar para que a filologia possa continuar a cumprir a responsabilidade de disciplina que se ocupa da peritagem dos textos e do seu diálogo com a cultura e a língua”. Entre esses conceitos, estão o conceito de edição e o próprio conceito de texto, na medida em que este, liberto da sua tradicional materialidade e conseqüente linearidade, “descorporificado”, não pode já ser contido naquilo a que tradicionalmente chamamos edição.

Nas novas edições da era digital, paradoxalmente, ou talvez não, o texto reaproxima-se do sentido original de “tecido”, enquanto objecto plural, rede de sentidos que se cruzam, por vezes de diferentes formas possíveis, e a que actualmente é possível aceder como um todo, de forma colaborativa, dinâmica e, sobretudo, aberta, o que, sem dúvida, permite maximizar as potencialidades das fontes escritas em Linguística Histórica.

Ana Paula

Banza

---

46

## REFERÊNCIAS

ALVES, Daniel. As Humanidades Digitais como uma comunidade de práticas dentro do formalismo académico: dos exemplos internacionais ao caso português. **Ler História**, n. 69, p. 91-103, 2016.

BANZA, Ana Paula. Dicionários e História da Língua: ‘vozes antiquadas’ na Produção Lexicográfica e Paralexicográfica do séc. XVIII Português. **Revista Portuguesa de Humanidades**, Vol. 23 (1-2), p. 2944, 2019.

\_\_\_\_\_. Crítica Textual e História da Língua: problemas de sempre e novos desafios. In: GONÇALVES, Maria Filomena; FACHIN, Phablo (orgs.). **Fios e nós - Entre Filologia e História da Língua**. São Paulo: NEHiLP, Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa, no prelo.

BERLINCK, Rosane de Andrade; BARBOSA, Juliana Bertucci; MARINE, Talita de Cássia. Reflexões teórico-metodológicas sobre fontes para o estudo histórico da Língua. **Revista da ABRALIN**, v. 7, n. 2, jul./dez., p. 169-195, 2008.

BORGES, Rosa *et alii*. **Edição de Texto e Crítica Filológica**. Lisboa: Quarteto, 2012.

CASTRO, Ivo. **Introdução à História do Português**. Lisboa: Edições Colibri, 2006.

CASTRO, Ivo; RAMOS, Maria Ana. *Estratégia e tática da transcrição (tiré à part du volume Critique Textuelle Portugaise — Actes du Colloque)*, Centre Culturel Portugais. Paris: Fondation Calouste Gulbenkian, 1986.

DUARTE, Luiz Fagundes. Para uma edição interactiva de textos antigos. In: CASTRO, Ivo (org.). **Actas do XII Encontro Nacional da Associação Portuguesa de Linguística** (Braga - Guimarães, 30 de Setembro a 2 de Outubro de 1996), Lisboa: Associação Portuguesa de Linguística, II (Linguística História, História da Língua), 1997, p. 411-417.

EMILIANO, António. **Critérios e normas de transcrição e transliteração de textos medievais – V. 1.0**. Disponível em: [https://www.academia.edu/3777459/Crit%C3%A9rios\\_e\\_normas\\_para\\_transcri%C3%A7%C3%A3o\\_e\\_translitera%C3%A7%C3%A3o\\_de\\_textos\\_medievais](https://www.academia.edu/3777459/Crit%C3%A9rios_e_normas_para_transcri%C3%A7%C3%A3o_e_translitera%C3%A7%C3%A3o_de_textos_medievais), 2002a. Acesso em: Abril de 2020.

\_\_\_\_\_. Problemas de transliteração na edição de textos medievais. **Revista Galega de Filoloxía**, n. 3, p. 29-64, 2002b.

\_\_\_\_\_. Tipo Medieval para Computador: uma ferramenta informática para linguistas, historiadores da língua e paleógrafos”. **Signo. Revista de Historia de la Cultura Escrita** (Universidade de Alcalá de Henares), 15, p. 139 -176, 2005.

\_\_\_\_\_. Convenções de utilização dos caracteres editoriais e críticos do tipo medieval Notator Mono para a realização de edições paleográficas de textos medievais portugueses e latino portugueses, vol. 2. In: **Documentação do Projecto Origens do Português: Digitalização, Edição e Estudo Linguístico de Documentos dos séculos IX-X**. Ref.<sup>a</sup> POCI/LIN/58815/2004, 2007. Disponível em: [https://www.academia.edu/3777486/Conven%C3%A7%C3%B5es\\_de\\_utiliza%C3%A7%C3%A3o\\_dos\\_caracteres\\_editoriais\\_e\\_cr%C3%ADticos\\_do\\_tipo\\_medieval\\_Notator\\_Mono\\_para\\_a\\_realiza%C3%A7%C3%A3o\\_de\\_edi%C3%A7%C3%B5es\\_paleogr%C3%A1ficas\\_de\\_textos\\_medievais\\_portugueses\\_e\\_latino\\_portugueses](https://www.academia.edu/3777486/Conven%C3%A7%C3%B5es_de_utiliza%C3%A7%C3%A3o_dos_caracteres_editoriais_e_cr%C3%ADticos_do_tipo_medieval_Notator_Mono_para_a_realiza%C3%A7%C3%A3o_de_edi%C3%A7%C3%B5es_paleogr%C3%A1ficas_de_textos_medievais_portugueses_e_latino_portugueses). Acesso em: Abril de 2020.

FAULHABER, Charles B. Textual Criticism in the 21st Century. **Romance Philology**, XLV, n. 1, p. 123-148, 1991.

*Sobre o uso do  
texto escrito  
em linguística  
histórica*

---

47

LABOV, William. **Principles of Linguistic Change. Internal Factors.** Oxford: Blackwell, 1994.

\_\_\_\_\_. **Principles of Linguistic Change. Cognitive and Cultural Factors.** Oxford: Blackwell, 2010.

LASS, Roger. **Historical linguistics and language change.** Cambridge: Cambridge University Press, 1997.

Ana Paula  
Banza

---

LAVAGNINO, John. Reading, Scholarship, and Hypertext Editions, In: **TEXT: Transactions of the Society for Textual Scholarship**, vol. 8, p.109-124, 1995.

48

MAIA, Clarinda de Azevedo. **História do Galego-Português. Estado linguístico da Galiza e do Noroeste de Portugal desde o século XIII ao séc. XVI (Com referência à situação do galego moderno).** Coimbra: Instituto Nacional de Investigação Científica, 1986.

MARQUILHAS, Rita; HENDRICKX, Iris. Avanços nas humanidades digitais. In: MARTINS, Ana Maria; CARRILHO, Ernestina (eds.). **Manual de linguística portuguesa**, series: Manuals of Romance Linguistics, n. 16. De Gruyter Mouton, 2016, p. 252-277.

MARQUILHAS, Rita. Filologia. In: CEIA, Carlos. **E- Dicionário de Termos Literários**, 2010.

MATTOS E SILVA, Rosa Virgínia. Orientações Atuais da Linguística Histórica Brasileira. **DELTA**, vol.15, special issue, São Paulo, p. 147-166, 1999. Disponível em: <https://revistas.pucsp.br/delta/article/view/40329/27087>. Acesso em: Abril de 2020.

\_\_\_\_\_. **Caminhos da Linguística Histórica – ouvir o inaudível.** São Paulo: Editorial Parábola, 2008.

NETO, Serafim da Silva. **História da Língua Portuguesa.** 6ª ed. Rio de Janeiro: Presença; Lisboa: Dinalivro, 1992.

ORDUNA, Gérman. **Fundamentos de Crítica Textual**. Ed. de Leonardo Funes y José Manuel Lucía Megías, Madrid: Arco/Libros, S. L., 2005.

PAIXÃO DE SOUSA, Maria Clara. Memórias do Texto. **Texto Digital** (UERJ), vol. 1, 10, 2006.

\_\_\_\_\_. A Filologia Digital em Língua Portuguesa: alguns caminhos. In: GONÇALVES, Maria Filomena; BANZA, Ana Paula (ed.). **Património Textual e Humanidades Digitais**: da Antiga à Nova Filologia. Évora: CIDEHUS, p. 113-138, 2013. Disponível em: <https://books.openedition.org/cidehus/1073>. Acesso em Abril de 2020.

*Sobre o uso do  
texto escrito  
em linguística  
histórica*

PARKINSON, Stephen. Um Arquivo Computorizado de Textos Medievais Portugueses. **Boletim de Filologia** 28, p. 241-252, 1983.

49

PAUL, Herman. **Prinzipien der Sprachgeschichte**. Halle: Niemeyer, 1880 (2ª edição). Disponível em: <https://archive.org/details/prinzipien-der-sp01paulgoog>. Acesso em: Abril de 2020.

PÉREZ PRIEGO, Miguel Ángel. **La edición de textos**. Madrid: Síntesis, 1997.

SÁEZ RIVERA, Daniel M. Crítica textual, historiografia linguística e historia de la lengua. In: PONS RODRÍGUEZ, Lola (ed.). **Historia de la Lengua y Crítica Textual**, Madrid/Frankfurt am Main: Iberoamericana/Vervuert, 2006, p. 267-302.

SAHLE, Patrick. **Criteria for Reviewing Scholarly Digital Editions. Version 1.1**. IDE, Institut für Dokumentologie und Editorik, 2014. Disponível em: <http://www.i-d-e.de/reviews/criteria-version-1-1>. Acesso em: Abril de 2020.

SAUSSURE, Ferdinand. **Curso de linguística geral** [1916]. Lisboa: Dom Quixote, 1992.

TEXT ENCODING INITIATIVE (TEI). Disponível em: <https://tei-c.org>. Acesso em: Abril de 2020.

VASCONCELLOS, José Leite de. **Lições de Philologia Portuguesa**. Lisboa: Clássica Editora, 1911. Disponível em: <https://archive.org/details/liesdephilol00vascuoft/page/n4>. Acesso em: Abril de 2020.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical Foundations for a Theory of Language Change. In: LEHMANN, W. P; MALKIEL, Y. (eds.). **Directions for Historical Linguistics**. Austin: University of Texas Press, 1968, p. 95-195.

Ana Paula

Banza

---

50