

Ler a prosa do mundo hoje [en]

Maria Clara Paixão de Sousa <mariaclara_at_usp_dot_br>, Universidade de São Paulo

Abstract

As tecnologias digitais de difusão da informação transformaram profundamente o trabalho das *Humanidades*, que está hoje inscrito na lógica digital de um modo muito mais profundo do que o rótulo *Humanidades Digitais* pode fazer crer – pois não estamos diante do surgimento de uma nova tendência, nem de uma nova linha de pesquisas, nem de um novo campo de estudo no interior das Humanidades: estamos de fato diante de *outras Humanidades*. Na raiz dessa transformação está uma nova materialidade do texto, que, como já discutiram Pédaque (2004, 2006, 2007), Crane et al. (2008), Gradmann & Meister (2008), Chaudiron et al. (2008), Baumann & Crane (2010), e Paixão de Sousa (2013), modifica a condição histórica do *documento* – e, assim, instaura novas formas de leitura e novas formas de ordenação da leitura. Nesse sentido, sendo tarefas das Humanidades a interpretação do texto e a organização da sua transmissão social, essa nova materialidade interpela diretamente os métodos, os horizontes epistemológicos, e a conformação discursiva do campo. Neste artigo discutirei alguns dos efeitos desse processo, fundada em uma abordagem conceitual da materialidade do texto digital já proposta em Paixão de Sousa (2013), e inspirada na reflexão de Pêcheux (1992[1982]) sobre a *leitura do arquivo* e nas ideias de Unsworth (2006) sobre as *formas de atenção*.

1 Preâmbulo (ou: “A Prosa do mundo”)[1]

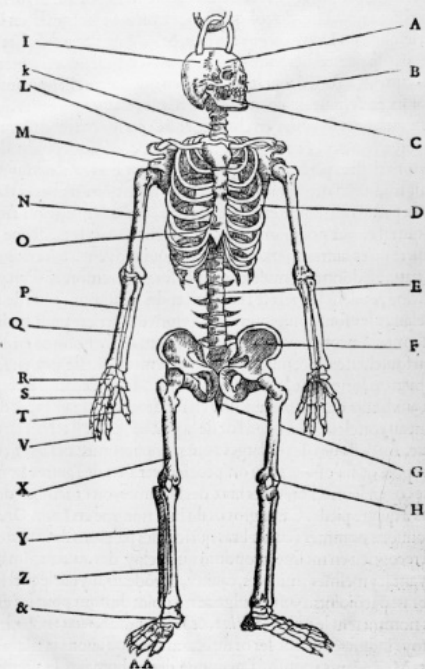
Começo essa reflexão trazendo uma anedota bastante pessoal – que, espero, em tempo se justificará como ilustração prévia para o debate que proponho. Vamos à história: a primeira vez que li *As palavras e as coisas* [Foucault 2000] foi há quase trinta anos, naqueles tempos longínquos quando, estudantes, entrávamos nas bibliotecas (aqueles edifícios feitos de concreto), íamos até as estantes pegar os livros (aqueles objetos feitos de papel), sentávamos à mesa e... líamos. E, se quiséssemos consultar um ponto obscuro na leitura, tínhamos que andar até uma outra estante, pegar um outro objeto feito de papel, e tentar achar entre suas páginas (literalmente páginas, viradas uma a uma) a informação de que precisávamos. Nessas condições antediluvianas, e nos meus vinte anos, a leitura das *Palavras* foi um desafio imenso – não apenas pela imaturidade intelectual que, inevitavelmente, me permitia compreender muito pouco ou quase nada do que estava ali, mas também (hoje concluo) porque eu não conseguia *enxergar* nada do que era descrito na complexa rede de referências discursivas e, em alguns pontos, iconográficas tramada por Foucault na obra.

Mais recentemente, em uma nova tentativa com *As Palavras*, descobri neste mesmo livro um universo repleto de imagens coloridas e fascinantes. Dessa vez, não lia na mesa escura de uma biblioteca de concreto; lia podendo consultar, numa tela de vidro brilhante, todas as obras citadas no livro – todas, simultaneamente, abertas aos meus olhos, acessíveis aos meus dedos. Um efeito particularmente forte desse alcance às referências se revelou na leitura do capítulo “A Prosa do Mundo”, aquele em que Foucault discute o poder das analogias visuais nos tratados renascentistas. Essa discussão, repleta de força imagética, tornou-se inacreditavelmente menos hermética e mais compreensível com o acesso às gravuras que ele comenta.

1

2

Portrait de l'amas des os humains, mis en comparaison
de l'anatomie de ceux des oyseaux, faisant que les
lettres d'icelle se rapporteront à ceste cy, pour
faire apparoistre combien l'affinité est
grande des vns aux autres.



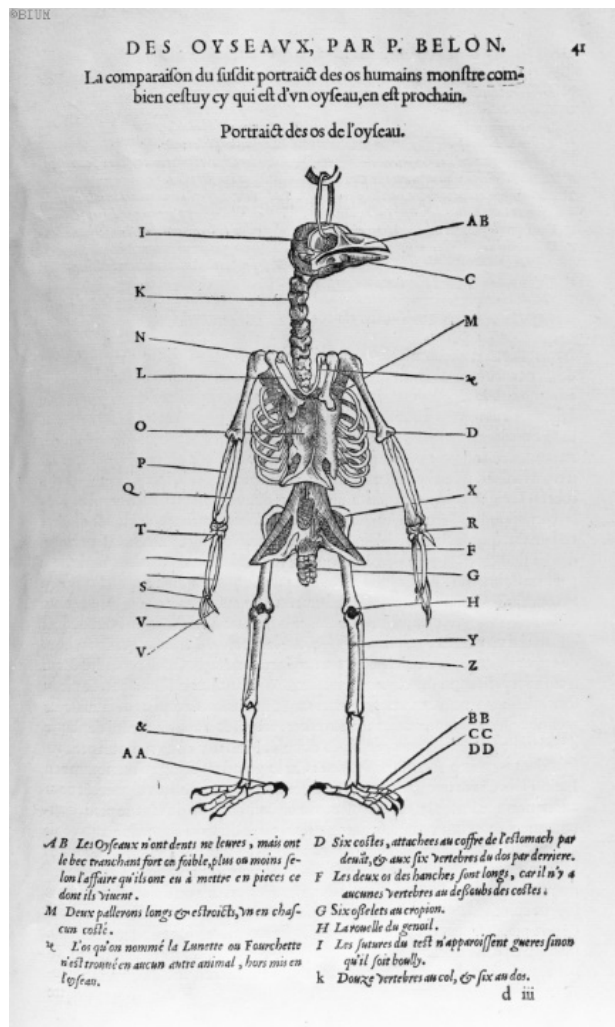


Figure 1. Portraict de lamas des os humains / Portraict des os de loyseau – In: L'Histoire de la nature des oyseaux, P. Belon, 1555

Animei-me então a elaborar para os meus alunos (para quem, afinal, estava preparando a leitura do texto) uma pequena *biblioteca virtual* com ligações para versões digitais das obras citadas por Foucault^[2]; e, depois, fui ler o capítulo com eles, absolutamente encantada com os meus achados. Eles, por sua vez, não pareceram ver nada demais naquilo. Ingenuidade minha – pois, ao contrário de nós que nos cansávamos nas andanças frustrantes pelos corredores das bibliotecas universitárias nos tempos dantanho, os jovens de vinte anos hoje consideram normal; esperado; “evidente” que “tudo pode ser encontrado”; que “tudo pode ser visto”; que tudo está “aqui”, na ponta dos nossos dedos.

Naquele dia, talvez um pouco espantada pela naturalidade com a qual meu pequeno gabinete virtual de maravilhas foi recebido, pensei em como, por contraste, devem ter sido raros os leitores que, desde a publicação de *As palavras e as coisas* em 1966, puderam acompanhar a leitura do texto com a consulta fácil e imediata às vinte e três fontes textuais escritas entre 1555 e 1674, citadas só no capítulo da “Prosa”, e às ilustrações que elas contêm. Considerando a raridade dos livros referidos no capítulo, só podemos imaginar esse leitor singular no gabinete de uma biblioteca particular excepcional, ou na sala de leitura de uma grande biblioteca europeia – por exemplo, na Biblioteca Nacional da França, que guarda a maior parte das obras citadas ali; mas mesmo esse leitor privilegiado ainda precisaria nalgum momento se deslocar para a Itália ou para a Alemanha se quisesse de fato consultar todas as referências. Hoje, entretanto, a Biblioteca Nacional da França, a Biblioteca da Universidade de Bolonha, a Biblioteca da Universidade de Munique, tornaram-se próximas, separadas apenas pelo movimento de um dedo. E um jovem de vinte anos, em São Paulo, ao ler a “Prosa” pela primeira vez, pode, com toques rápidos numa pequena tela iluminada, ver todos os animais, monstros e rostos bestiais comentados minuciosamente por Foucault.



Figure 2. Des oreilles grandes In: Giambattista della Porta, *La Physionomie humaine*, 1655, p. 147. Ch.VIII, *Des oreilles*.

Entretanto – quais os efeitos dessa leveza no levantar o olho do livro? Essa profusão de informações imagéticas contribui para um aguçamento do entendimento da obra, ou de fato impede o olhar de concentrar-se nas palavras do autor? Pois – o olho muitas vezes não perde o caminho de volta, quando o dedo toca a tela iluminada? Horas depois, não vamos perceber que já não estamos mais lendo *As Palavras*, mas sim nos maravilhando com as 11.623 imagens que chegaram pelo “Google” quando buscamos: “Aldrovandi, Dragons”?

5

Mais que isso. Quando preparamos para nossos alunos pequenas coleções de maravilhas, e de fato quando damos ao mundo nossos novos produtos de pesquisa – nossas coleções coloridas de iconografias em torno de obras clássicas, com finalidade pedagógica; nossas constelações de textos antigos, dispersos por séculos no Velho Mundo, mas agora unidos em uma *rede* que nós mesmas criamos digitando códigos numa tela, no Novo; quando fabricamos textos que se movem, edições que chamam o dedo do leitor – o que é que estamos fazendo? Não estamos (nós mesmos, os *humanistas*, os que pensamos ser os guardiões das formas que a nossa cultura inventou para ler-se) também contribuindo para essa dispersão dos olhares? Pois reordenar materialmente o texto, transformar as redes de suas referências, torná-lo poroso e aberto à manipulação instantânea – não é tecer um novo texto?

6

Borges não acrescenta nenhuma figura ao atlas do impossível; não faz brilhar em parte alguma o clarão do encontro poético; esquiva apenas a mais discreta, mas a mais insistente das necessidades; subtrai o chão, o solo mudo onde os seres podem justapor-se. Desaparecimento mascarado, ou, antes, irrisoriamente indicado pela série abecedária de nosso alfabeto, que se supõe servir de fio condutor (o único visível) às enumerações de uma enciclopédia chinesa... Numa palavra, o que se retira é a célebre “tábua de trabalho”; e, restituindo a Roussel uma escassa parte do que lhe é sempre devido, emprego esta palavra “tábua” em dois sentidos superpostos: mesa niquelada, encerada, envolta em brancura, faiscante sob o sol de vidro que devora as sombras — lá onde, por um instante, para sempre talvez, o guarda-chuva encontra a máquina de costura; e quadro que permite ao pensamento operar com os seres uma ordenação, uma repartição em classes, um agrupamento nominal pelo que são designadas suas similitudes e suas diferenças — lá onde, desde o fundo dos tempos, a linguagem se entrecruza com o espaço [Foucault 2000, xi].

7

Naquele lugar onde “desde o fundo dos tempos, a linguagem se entrecruza com o espaço” – que aspecto tem, hoje,

8

nossa “tábua de trabalho”? Por sobre que mesa ordenamos e damos sentido aos nossos escritos; que superfície “niquelada, encerada, envolta em brancura” forma a base sobre a qual podemos espalhar nossas leituras? Que quadro nos permite hoje operar uma ordenação dos seres e das “coisas”? Em que espaço lemos “a prosa do mundo”, hoje?

Aqui interrompo minha anedota, pois não tenho respostas para essas perguntas. Mas me parece certo que vemos se instaurar, no estudo e no ensino das Humanidades, uma nova forma de ler.

9

2 As humanidades e o ambiente digital de difusão do texto

Em 2001, John Unsworth apontava para a profundidade das mudanças epistemológicas em jogo nas Humanidades frente ao contato com as tecnologias computacionais destacando, entre outros fatores, seus impactos sobre a formação de novos humanistas. Inscritas na lógica digital, as novas gerações encontrariam formas inteiramente novas de relacionarem-se com a documentação deixada pelas gerações passadas, fazendo transformarem-se “as Humanidades”:

10

The assertion of this paper is that the methodology known as knowledge representation has profound implications for humanities computing, and through humanities computing, has the potential to change the way humanities scholarship is done, to change the nature of graduate education in the humanities, and to change the relationship between the humanities and other professions, let alone other disciplines. I believe that knowledge representation has already produced important new research, and will, in the future, bring us new insights into what we know about the human record, and how we know it^[3] [Unsworth 2001].

11

Em 2008, Gregory Crane, David Bamman e Alison Babeu sugeriam que as novas ferramentas digitais à disposição dos pesquisadores nas Humanidades fariam vislumbrar a criação de “um espaço dinâmico para a vida intelectual” que equivaleria a um passo revolucionário na história da leitura:

12

The tools at our disposal today, primitive as they may appear in the future, are already adequate to create a dynamic space for intellectual life as different from what precedes it as oral culture differs from a world of writing^[4] [Crane et al. 2008].

13

Como muitos autores que enxergaram no impacto das tecnologias computacionais um divisor de águas na história das Humanidades, Crane et al (2008) e Unsworth (2001) falam a partir de um campo específico das Humanidades que, me parece, foi particularmente sensível à interferência das técnicas eletrônicas: o trabalho com o texto, na chave mais filológica ou na chave mais linguística. Baumann & Crane (2010), por exemplo, fazem explicitamente essa ligação direta entre o deparar-se com um “novo espaço intelectual” e o trabalho com *corpora* de línguas:

14

Treebanks are collections of text with extensive morphological, syntactic and similar categories of annotation and are familiar instruments for corpus and computational linguistic research. In building Treebanks for historical languages such as Greek and Latin, we found a new intellectual space that combined elements from computational and corpus linguistics and from the ancient discipline of philology^[5] #baumann2010.

15

Não é surpreendente encontrarmos as principais reflexões sobre a profundidade do impacto do digital nas Humanidades nas áreas que trabalham de modo fundamental com o *texto* – pois foi o texto, em primeiro lugar, o objeto de atenção humanística que se transfigurou da forma mais absoluta com o advento da transmissão digital da informação. É talvez por se verem cotidianamente diante desse novo objeto, materialmente distinto do objeto-texto anterior (conforme volto a comentar mais à frente), que os filólogos e linguistas envolvidos na construção de coleções estruturadas de textos em meio eletrônico têm se mostrado fascinados com as potencialidades inovadoras desse objeto. Não haverá testemunho mais eloquente desse fascínio que as palavras do jesuíta Roberto Busa sobre as primeiras versões de seu *Corpus Thomisticum* já nos anos 1940: “Digitus Dei est hic!”, “o dedo de Deus está aqui!” #busa2004. Fora do ambiente de trabalho com o texto, entretanto, o entusiasmo de linguistas e filólogos talvez pareça um tanto exagerado. Acredito que para compreender nossa perspectiva de que a difusão digital inscreve novas

16

Humanidades seja interessante compreender por que o texto digital nos parece conformar um objeto tão radicalmente novo.

2.1 Perspectiva material do texto digital

Como venho propondo em alguns trabalhos recentes, de uma perspectiva *material* o *texto digital* é uma etapa objetivamente nova na história do texto, diferindo fundamentalmente das outras formas textuais por envolver uma combinação de elementos lógicos naturais e artificiais em sua codificação e, no limite, por sua propriedade de *representar* artificialmente a linguagem [Paixão de Sousa 2013]. Por conta dessa participação da lógica artificial em sua composição, o *texto digital* é um texto *descorporificado*, cuja materialidade não se consubstancia empiricamente em um objeto palpável, mas sim se inscreve na representação do código que forma a possibilidade do texto.

Essa é uma diferença material fundante, não uma nova forma para um mesmo objeto. Vemos isso já no plano mais básico da representação linguística artificial contida no texto digital, a *codificação de caracteres* – aquilo que de fato, de uma perspectiva material e computacional, faz diferir um arquivo do tipo *texto* de outros tipos de arquivos. A codificação de caracteres funciona como uma matriz lógica de remissões entre dígitos e um inventário de símbolos convencionados para certos valores linguísticos (“a, b, c”...), e as programações computacionais operam, nos processadores de texto, manipulando essas remissões lógicas de modo a compor uma representação que é interpretada, humanamente, como um *texto*:

17

18

Isso é um texto

```
01001001 01110011 01110011 01101111 00100000 11000011
10101001 00100000 01110101 01101101 00100000 01110100
01100101 01111000 01110100 01101111
```

Figure 3. “Isso é um texto”. Texto digitado no conversor ASCII em <https://www.asciitohex.com>, fazendo explicitar-se a codificação binária de caracteres

A mágica aqui – o “dedo de Deus” no dígito – está nas possibilidades de manipulação que a codificação digital do texto abre para o trabalho filológico e linguístico; e o instrumento no qual essas possibilidades se evidenciam mais claramente é, talvez, o corpus eletrônico anotado. De partida, a propriedade *incorpórea* do texto digital acaba colocando em termos muito interessantes o problema da definição do corpus eletrônico, ou seja, dessa coleção de documentos cuja coesão material nos parecia tão forte que a denominamos *corpo*. Em Paixão de Sousa (2014) sugeri o conceito de *corpus eletrônico* como coleção de textos digitais hiper-codificados que formam uma “representação artificial da análise linguística”. A partir de Unsworth (2006), sugeri que o trabalho de anotação depositado sobre os textos na formação de um corpus eletrônico é uma das instâncias da chamada *representação do conhecimento*, *Knowledge Representation*, de modo que anotar um texto eletrônico para fins de análise linguística é tornar explícitas nossa interpretação e nossa análise. Ressalto entretanto um passo importante que poderia ficar obscuro ao se conceituar a anotação de textos como uma operação da representação do conhecimento: não estou afirmando que anotar um texto é *explicitar sua estrutura*. Do ponto de vista linguístico e do ponto de vista material, um texto não *tem* uma estrutura a ser explicitada; nós, como intérpretes, como analistas, é que atribuímos uma estrutura ao texto, e é isso que anotamos. Assim, *anotar* textos é efetivamente aplicar *camadas de representação interpretativa* sobre eles.

É em cima da nossa anotação – ou seja, em cima dessa nossa *interpretação explicitada* – que um texto poderá vir a ser processado *automaticamente* em um corpus eletrônico. Note-se o detalhe central: o processamento automático não se dá sobre um objeto natural, e sim sobre uma interpretação do texto, que inscrevemos computacionalmente na codificação e levamos ao processamento de uma máquina.

19

20

[Isso/DETERMINANTE]-SUJEITO

[é/VERBO um/DETERMINANTE texto/NOME anotado/ADJETIVO]-PREDICADO

```
01011011 01001001 01110011 01110011 01101111 00101111 01000100
01000101 01010100 01000101 01010010 01001101 01001001 01001110
01000001 01001110 01010100 01000101 01011101 00101101 01010011
01010101 01001010 01000101 01001001 01010100 01001111 00100000
00001101 00001010 01011011 11000011 10101001 00101111 01010110
01000101 01010010 01000010 01001111 00100000 01110101 01101101
00101111 01000100 01000101 01010100 01000101 01010010 01001101
01001001 01001110 01000001 01001110 01010100 01000101 00100000
01110100 01100101 01111000 01110100 01101111 00101111 01001110
01001111 01001101 01000101 00100000 01100001 01101110 01101111
01110100 01100001 01100100 01101111 00101111 01000001 01000100
01001010 01000101 01010100 01001001 01010110 01001111 01011101
00101101 01010000 01010010 01000101 01000100 01001001 01000011
01000001 01000100 01001111 00001101 00001010
```

Figure 4. Isso é um texto anotado. Texto digitado no conversor ASCII em <https://www.asciitohex.com>, fazendo explicitar-se a codificação binária de caracteres.

Significa dizer, fundamentalmente, que preparar um texto para ser processado automaticamente por uma máquina é construir a possibilidade do processamento computacional de sua interpretação (linguística, estrutural, estética). Nesse ponto já podemos compreender um pouco do fascínio de um linguista diante dessa metodologia: como sugeri em Paixão de Sousa (2014), a anotação sintática, por exemplo, nos permite *analisar nossas próprias análises*, por meio da construção de sua *explicitação matemática*. Nisso opera um passo pouco discutido, mas fundamental: pois deparamo-nos, nesse processo, com a restrição lógica da máquina – a máquina, esse construto humano que não compreende o ambíguo, que não compreende o fluido, que precisa de categorias muito claras e recursivamente aplicáveis. Ou, nas palavras de Unsworth (2006), deparamo-nos com a “consistência tola” exigida pelo computador (“the ‘foolish consistency’ that the computer requires”). No que toca especificamente o trabalho do linguista, essa exigência de uma consistência absoluta torna o trabalho de anotação particularmente interessante – já que, lembro, a linguística é uma área das humanidades na qual a *intuição* do analista é uma das mais legítimas e centrais ferramentas de análise; e traduzir uma análise intuitiva em uma sequência a de instruções de processamento inequívocas e inambíguas pode ser uma tarefa extremamente desafiante. O efeito da “consistência tola” afeta a análise linguística pelo seu avesso: *ensinar* uma máquina a processar computacionalmente padrões linguísticos pode ser interessante justamente pelo fato de a análise linguística intuitiva nos ser tão acessível, e porque sua facilidade esconde uma série de passos normalmente implícitos. Ou seja, esconde interpretações que fazemos só porque somos humanos e falantes de uma língua natural (não por causa das nossas sofisticadas teorias; de fato, muitas vezes até apesar delas). Essa é talvez uma especificidade que diferencia o trabalho computacional em linguística do trabalho computacional em outras áreas das Humanidades: o que estamos analisando computacionalmente, na linguística, é algo que (a rigor) *já sabemos*. Mas o computador, com sua exigência de uma “consistência tola”, nos impede essa intuição, e nos faz, pelo avesso, ver onde ela atua. Assim, na batalha cotidiana frente à burrice fundamental do autômato, terminamos questionando radicalmente nossas categorias naturais, sempre tão ricas em compreender o ambíguo, o fluido, o que não é claro, o que não se repete jamais.

E com isso chegamos ao coração do problema – pois essa incursão pelo mundo da lógica fria de uma máquina que calcula sequências de números termina por transformar, indelevelmente, nosso olhar sobre o texto. Os *corpora eletrônicos* anotados ressignificam profundamente o trabalho de análise linguística, ao obrigar o linguista a uma nova hermenêutica, mediada por procedimentos artificiais de processamento simbólico. Assim é que a construção de bases de dados codificadas quanto à estrutura linguística cria não só uma nova ferramenta, mas sim um novo objeto de análise para a área, e funda, consequentemente, um novo olhar.

Saindo da linguística, mas ainda permanecendo no território dos estudos textuais, esse reajuste hermenêutico no uso das tecnologias computacionais está também presente em outras áreas de estudo. No campo dos estudos literários, a possibilidade de se reconhecerem padrões que não são possíveis a *olho nu* pelo intermédio do processamento automático é salientada, por exemplo, por Berry (2011), lembrando a experiência de Tanya Clement com a obra de Gertrude Stein, em Clement, Steger, Unsworth & Uszkalo (2008). O autor relaciona esse aspecto ao conceito de *leitura distante* versus *leitura próxima* (*distant reading* e *close reading*). O trabalho computacional possibilita a visão da repetição de padrões, coisa que não vemos a olho nu na língua e no texto – porque, na língua e no texto, fora do espaço artificial que criamos como um *corpus* (corpo!), o universo é amplo, fluido, e ambíguo demais.

23

Dessa forma, o trabalho computacional, nas disciplinas centradas no texto, cria uma nova *leitura*. As grandes bases de dados não nos ajudam apenas a arquivar grandes volumes de informação e processá-las mais rapidamente, mas sim, fundamentalmente, a depreender padrões, e no limite, *interpretá-los*. Como salientei de início, essa questão dos efeitos do processamento automático do texto parece atingir com mais centralidade as disciplinas humanísticas mais tradicionalmente dedicadas ao texto como objeto de estudo, como a linguística, a filologia, e (talvez de uma forma diferente, mas ainda assim intensa) os estudos literários; não poderei aqui visitar com propriedade seus efeitos sobre as disciplinas irmãs, como a história e as áreas voltadas para os estudos sociais e culturais. Entretanto, talvez se mostre igualmente relevante nesses campos próximos a segunda questão que gostaria de debater aqui: a proliferação e dispersão das fontes.

24

2.2 O texto digital e sua difusão

A nova forma de interpretação e leitura aberta pelo meio digital constrói-se em meio a uma tensão entre a restrição colocada pelos limites lógicos e *tolos* da máquina, que comentei até aqui, e algo que eu agora chamaria de uma *explosão* de textos a serem *lidos* – na realidade, muitas vezes, lidos indiretamente, com a ajuda da extensão de olhos eletrônicos.

25

Para justificar o meu uso do termo “explosão”, vejamos três dados brutos. Primeiro: hoje, 1º de março de 2018, devem existir na internet 4.24 bilhões de páginas [Kunder 2018]^[6] – e estamos falando em “páginas” no sentido (emulatório) de “páginas-web” (considere-se que, segundo um curioso experimento feito em 2015, para se imprimir todo o conteúdo da “web” naquele ano gastar-se-iam 305.5 bilhões de páginas de fato, ou seja, páginas de papel; cf. Dewey, 2015). Em meio a esse volume, existiam, em 2014, 114 milhões de documentos que podem ser considerados “acadêmicos”, segundo Khabza (2014), apenas na língua inglesa (desconheço pesquisa semelhante que vá além da documentação em inglês). Para tornar a questão um pouco mais próxima da nossa reflexão sobre as Humanidades, até 2012, 23 milhões de documentos (entre livros, gravuras e manuscritos) haviam sido digitalizados apenas nas bibliotecas institucionais europeias, segundo o levantamento de Peckel (2012) – o que, por sinal, ainda correspondia então a cerca de 10% do patrimônio físico desses acervos (levando a crer que, a permanecerem as condições atuais, esse número só tende a crescer nos próximos anos).

26

Do muito que se poderia discutir a partir desses dados bastante objetivos, faço apenas uma pergunta: como vamos chegar a *ler* isso tudo? Alguém sentará frente à tela de um computador e consultará as 4 bilhões de páginas-web disponíveis na internet em 2018, ou mesmo (para sermos menos ambiciosos) os 114 milhões artigos acadêmicos disponíveis em 2014, até encontrar um assunto que lhe interesse? Algum filólogo ou historiador consultará os 23 milhões de documentos primários disponíveis em formato digital só no contexto europeu em 2012 para selecionar os que lhe interessem? Evidente que não. Pois não somos nós que organizamos os textos colocados nas máquinas – são as próprias máquinas, que indexam, organizam, e fazem os textos revelarem-se para nossas “buscas”.

27

Noutros termos: o arquivo textual disponível hoje “na internet” se diferencia das outras formas de arquivo por se organizar fundamentalmente a partir de uma lógica artificial (na qual, a partir de sua concepção, um algoritmo ordena *automaticamente* os textos a partir de instruções que podem se repetir indefinidamente), enquanto as formas de arquivo construídas em épocas passadas eram organizadas por diferentes formas de interferência humana. Essa interferência humana, certamente, contou com instrumentos auxiliares (o glossário; o índice; o tombo); tais

28

instrumentos indexadores, entretanto, funcionavam sempre a partir da vigilância cotidiana de *alguém* – um bibliotecário, um arquivista, um filólogo – que, em algum ponto do processo, *lia* o texto a ser indexado. A forma digital do arquivo, em contraste, pode prescindir da leitura humana ao longo de todo o processo de indexação. Assim, como sugeria acima, é só com a ajuda de *olhos eletrônicos* que podemos pensar em dar conta do volume absolutamente acachapante de textos a serem ordenados e “lidos” no mundo hoje. Esta contingência, me parece, se soma àquela outra, mais específica (da formação da leitura *distante* em contraponto à leitura *próxima*), para conformar uma relação diferente entre leitor e texto.

Desconfio que vivemos um momento paradoxal, em que nunca tivemos tantos textos facilmente acessíveis para ler, mas que tornou impossível *ler*. E isso é também um fator de transformação da nossa leitura, e das nossas disciplinas. Pois, como humanistas, nosso ofício sempre foi o de ler o mundo – mas o mundo não pode mais ser lido como antes; tanto porque estamos precisando abandonar o sonho de ler tudo o que há para ler, como sugeri – mas também por uma segunda razão, mais delicada: porque os muros que separavam os diferentes lugares de leitura se dissolveram, de forma que não temos mais a clareza do que, diante daquele imenso volume de documentos dispersos, é importante ler, ou (sequer) indexar. O que é “da nossa área”? Quais são os textos “fundamentais”? Quais são as boas fontes? Em outras palavras – o que é, hoje, o “Arquivo”?

Podemos compreender “Arquivo” em diversos sentidos – aqui, estou centralmente ocupada com um conceito discursivo, do “Arquivo” compreendido como a construção social de uma ordenação da memória e da leitura, e, portanto, sujeito à historicidade; em particular, remeto ao conceito trabalhado por Michel Pêcheux em “ Ler o arquivo hoje” [Pêcheux 1982]. Entendendo o “arquivo” em sua dimensão discursiva mais ampla de “campo de documentos pertinentes e disponíveis sobre uma questão”, o autor discute os procedimentos sociais e teóricos envolvidos na construção (e na “leitura”) dos arquivos em diferentes momentos da história e na atualidade – neste caso, em particular frente aos desafios trazidos pelas tecnologias computacionais. Para ele, então,

...nos encontramos diante de uma nova divisão do trabalho de leitura, uma verdadeira reorganização social do trabalho intelectual, cujas consequências repercutirão diretamente sobre a relação de nossa sociedade com sua própria memória histórica [Pêcheux 1982].

A ideia do “Arquivo” como um objeto social construído a partir de uma “divisão do trabalho” de leitura inspira fundamentalmente a reflexão aqui proposta: preocupa-me centralmente a posição dos humanistas na divisão do trabalho de construção do Arquivo hoje (nos termos de Pêcheux), e é nesse sentido que se colocam as perguntas, aqui sugeridas, sobre o modo como as tecnologias computacionais de difusão do texto interpelam os humanistas em seus ofícios tradicionais de “leitores” do Arquivo. Considerando, ainda, que a dispersão e a porosidade dos objetos-texto é parte fundante da natureza material do texto digital, não nos distanciamos, nesse segundo problema, das questões em torno da materialidade do texto digital, que condicionavam o primeiro). Como argumentei em Paixão de Sousa (2013), a condição de transmissão e circulação de uma forma de texto faz parte de sua definição material, e a condição de circulação do texto digital é essa: volumosa; dispersa; porosa. Isso traz consequências para a leitura – em particular se entendemos “leitura” no sentido proposto por Pêcheux, pois

a difusão digital, pulverizando o primado da autoridade sobre o objeto, desintegrou o portal regulador da circulação dos corpos. Nesse novo ambiente, o erudito pode construir novas esferas de circulação do saber; mas outros construirão outras esferas, em que circularão outros saberes. Assim, os estudiosos especializados da leitura e da construção do “arquivo” (no sentido de Pêcheux, 1994) podem passar a se ver acompanhados de novos leitores e construtores do arquivo. Pois o “saber escrever” (aqui no sentido expandido, não de codificar a escrita, mas de escrever o arquivo, inscrevendo a memória) saiu das nossas mãos, mais radicalmente do que saíra das mãos dos escribas diligentes a serviço dos eruditos medievais para passar para as mãos (máquinas) dos fabricantes de livros. Saiu de nossas mãos, de todas as mãos, e de todas as máquinas: descorporificou-se [Paixão de Sousa 2013].

Assim, essa nova forma de difusão do texto – nova forma de sua produção, circulação e leitura – ao conformar um

novo objeto-texto, incorpóreo, transforma nossa forma de estarmos *atentos ao texto*, como discuto a seguir.

3 Formas de atenção

I think we are arriving at a moment when the form of the attention that we pay to primary source materials is shifting from digitizing to analyzing, from artifacts to aggregates, and from representation to abstraction^[7] [Unsworth 2006].

35

Valendo-se do conceito de *formas de atenção* originalmente formulado por Kermode (2016[1985]) para explicar a formação do cânone artístico e literário, Unsworth (2006) propõe um breve, mas denso, histórico das diferentes etapas da relação entre as humanidades e as formas eletrônicas do texto. Kermode propunha, em seu trabalho clássico, que as diferentes formas da nossa atenção sobre os objetos da nossa cultura mudam com o tempo, por conta de forças sociais e históricas, e, com isso, trazem sempre novos objetos à vista, ao mesmo passo em que “escondem” outros – formando, assim, o “cânone” de cada momento. Mais que a ideia de cânone, ou de valoração cultural, a Unsworth interessa debater as forças tecnológicas que moldam nossa relação com as obras que, graças ao fato de terem sobrevivido ao processo descrito por Kermode, “chegaram até nós” como fontes primárias (de fato: as obras que seguem merecendo alguma forma da nossa atenção como fontes de estudo). É nesse contexto que o autor coloca a ideia de *text-as-tools*, o *texto-ferramenta*, como o objeto a ter merecido atenção primordial da computação em humanidades desde os anos 1940 (um corpus eletrônico anotado seria, talvez, o exemplo mais bem-acabado do “texto-ferramenta”). Entretanto, para ele, nossa forma de atenção ao texto digital estaria hoje se deslocando de modo a enxergar para além do texto-ferramenta, passando a ver (e criar) o *texto-jogo*, *text-as-game*, um objeto lúdico em estado latente, que ele já antevê (embora desconfie que trará algum incômodo ao ambiente acadêmico tradicional). É como entendo a passagem da *representação* para a *abstração*, o último passo do deslocamento da forma de atenção das humanidades ao texto digital segundo Unsworth, e que será essencial para que eu termine, mais à frente, a reflexão aqui proposta sobre nossas novas formas de leitura.

36

Entretanto, esse breve resumo da reflexão proposta por Unsworth (2006) em torno do *texto-jogo* parecerá remeter a um mundo incredivelmente utópico, quando a maioria de nós, humanistas, ainda acreditamos que um livro em uma tela de computador não passa de uma instanciamento de um livro feito de papel, apenas visto *em outro lugar*. De fato, é tão profundo o grau de elaboração artificial em torno do objeto-texto no ambiente digital, que poderíamos dizer que o que temos hoje, no texto digital (e em seu ambiente de circulação), mais que nada, uma emulação de texto (e de um ambiente de circulação). Essa contingência da difusão digital do texto como uma emulação do ambiente não-digital faz com que seja pouco surpreendente que ele opere, quase sempre, dissimulando o digital, simulando o livro físico e o ambiente de leitura do mundo físico. Essa emulação, esse *não revelar de um simulacro* de leitura, fazem parte do caráter alienante das tecnologias digitais, que nos fazem sentir, a cada passo, que estamos diante de bibliotecas, de livros, com páginas que podemos “folhear” na tela – quando estamos diante de uma máquina que apenas nos envia sinais para nos convencer de que são textos. É nesse sentido que uso, acima, o termo “alienante”: estamos imersos em uma lógica de produção textual cujas etapas fundamentais, em termos materiais, escapam à nossa percepção; e essa dissimulação faz parte da lógica interna do sistema de produção material do texto.

37

É razoável propor que enquanto estivermos presos a essa emulação, não conseguiremos entender a mudança na forma de atenção ao texto, e nos sentiremos perdidos, enquanto leitores, enquanto organizadores, e enquanto produtores do “texto”. Como leitores, aqui estamos diante daquilo que nos é feito ver como “textos”; assim é que temos por diante 114 milhões de artigos em periódicos eletrônicos, e nos comportamos como quem tem diante de si algumas tantas revistas em papel, a serem lidas concentradamente, uma após a outra, com uma calma que ficou perdida no tempo. O fato é que seguimos na emulação/simulação de que “vamos ler as coisas”, quando o que mais fazemos é selecionar, das telas iluminadas dos nossos computadores, pequenos trechos, aqui e ali, do que outros escreveram e que conseguimos *salvar* (termo, em si, muito interessante) em meio ao dilúvio de palavras que “a internet” nos faz chegar em casa todos os dias.

38

É também em meio a esse dilúvio informacional que ainda tentamos nos equilibrar como organizadores e editores do texto, na nossa tradicional função de *levar os textos ao mundo* (na etimologia, “editor”, de “edere”, “expulsar, colocar

39

para fora”). Não os textos que nós mesmo escrevemos – mas sim, a tradição da nossa cultura escrita, de que nós, como humanistas, nos arvoramos como guardiões e transmissores. Vamos lembrar que os humanistas construímos os instrumentos para a humanidade *se ler* – tanto aqueles entre nós cujo trabalho é dar textos à leitura, como aqueles cujo trabalho é explicar o texto, como aqueles cujo trabalho é ordenar o mundo dos textos. É isso que fazemos, tradicionalmente – mas diante dos milhões de documentos históricos digitalizados hoje, por exemplo, o que estamos fazendo, quando sequer estamos participando da organização desse volume insondável de palavras escritas soltas na rede?

Volto neste ponto à minha anedota inicial. Quando, em 2012, preparava meu gabinete de maravilhas para acompanhar os trabalhos de um grupo de leituras em torno da “Prosa do mundo”, eu não conhecia o projeto “La Bibliothèque Foucauldienne” [Artières and Berg 2010], no qual pesquisadores da Universidade de Lyon haviam reunido, em um banco de dados, todas as notas de trabalho de Foucault e toda a bibliografia usada pelo autor em suas obras, com resultados oferecidos à consulta pública. Assim, tudo o que eu fiz sozinha, garimpando “a internet” por dias a fio, estava na “Foucauldienne” – organizado, sistematizado, curado. Ter conhecido este trabalho antes de organizar minha pequena “Foucauldiana” para os alunos teria me poupado imensos trabalhos; meu garimpo, portanto, não parece ter sido muito eficiente. Em minha defesa, em experimento recente, ao buscar pelo termo “Foucault” no buscador Google, a máquina me respondeu com mais de 26 milhões de resultados – e nenhum deles era a “Foucauldienne” de Lyon (ao menos, nas 44 páginas que ela me deu a ler). Assim, não é apenas para lamentar profundamente esta falha que puxo o caso para esta conversa – e sim para ilustrar a questão do deslocamento, ou movimento, das formas da nossa atenção.

40

Pois – qual é, afinal, a *janela* pela qual os leitores enxergam o dilúvio de textos digitais? Nesse ponto não podemos nos iludir: a organização do texto digital, quer estejamos falando de uma notícia de jornal, de um documento medieval cuidadosamente digitalizado por uma biblioteca universitária, ou do conjunto dos cadernos de notas de Michel Foucault, está a cargo do “Google”. É “no Google”, não nas bibliotecas, que os leitores procuram textos.^[8] Por ora, surpreendentemente, quase todos parecemos estar bastante tranquilos em saber que algoritmos produzidos por grandes corporações multinacionais são, hoje, os principais responsáveis por ordenar as palavras legíveis no mundo.

41

Na outra ponta do dilúvio informacional estamos nós como “produtores” desses estranhos objetos que emulam textos, que se multiplicam como vírus pelo mundo em forma de sinais elétricos, e que enviamos na esperança de que algum dia alguém vai sentar em uma mesa e, efetivamente, *ler*. O preocupante nesta ponta do processo, e que trago como ponto final para esta conversa, é algo a que G. Crane, D. Bamman e A. Babeu já se referiram de modo muito preciso ao comentar os impactos do digital na área dos estudos clássicos:

42

Digital technology is hardly new in classics: there are full professors today who have always searched large bodies of Greek and Latin, composed their ideas in an electronic form, found secondary sources on-line and opportunistically exploited whatever digital tools served their purposes. Nevertheless, the inertia of prior practice has preserved intact the forms that evolved to exploit the strengths and minimize the weaknesses of print culture: we create documents that slavishly mimic their print predecessors; we send these documents to the same kinds of journals and publishers; our reference works and editions have already begun to drift out of date before they are published and stagnate thereafter; even when new, our publications are static and cannot adapt themselves to the needs of their varying users; while a growing, global audience could now find the results of our work, we embed our ideas in specialized language and behind subscription barriers which perpetuate into the twenty-first century the miniscule audiences of the twentieth^[9] [Crane et al. 2008, 1–2].

43

“Estamos perpetuando, nó século 21, o minúsculo público leitor do 20” – este alarme de Crane et al. (2008) tem consequências mais importantes do que pode parecer à primeira vista. Pois se há de um lado, como vimos, o problema de estarmos afogados no mundo da informação e não conseguirmos nos mover direito dentro dele (como se, na analogia da biblioteca de concreto e do livro de papel, estivéssemos caminhando em círculos, exaustos, pelos corredores de uma biblioteca grande demais para as nossas pernas), de outro lado, e de modo correlato, *nós não*

44

estamos sendo lidos. Pois tudo o que escrevemos acaba também se dispersando antes de encontrar os leitores perdidos em seu próprio afogamento. O que há de errado?

O que nos falta, desconfio, é darmos o passo final na radical transformação da nossa maneira de enxergar esses nossos objetos de trabalho, os textos – e que nos leva de volta ao *texto-jogo* de Unsworth (2006). Tomo o *texto-jogo* de Unsworth como o exemplo mais bem acabado do que eu chamei anteriormente de texto *descorporificado* – pois esse seria o texto que não precisa mais se limitar a *aparecer* como texto: pode ser imagem, pode ser cor, pode ser som – pode, efetivamente, “falar com seus leitores”, nos termos de Crane et al. (2008). Esses autores defenderam que há uma diferença radical a separar o mundo da difusão digital e o mundo do papel, como já mencionamos; e trazem também um ponto central para este momento final da nossa conversa: a diferença entre o texto como objeto “fixo” (no papel) e o texto como objeto “móvel” (no ambiente digital). É essa, para eles, a singularidade do texto no ambiente digital – e a exploração ativa dessa singularidade deveria ser a prioridade da produção textual acadêmica hoje:

45

The question that we face is much deeper than the challenge of producing more or, preferably, better articles and monographs. We must more generally ask what kind of space we wish to produce in which to explore the linguistic record of humanity – whether we are contemplating the Odyssey, administrative records from Sumer, or tracing mathematical thought through Greek and Arabic sources. More important perhaps than the question of what we can do may be the opportunity to redefine who can do what – to open up intellectual life more broadly than ever before and to create a fertile soil in which humanity can cultivate the life of the mind with greater vigor and joy.

[10] [Crane et al. 2008, 4] (meu grifo)

Assim, seria interessante que fossemos capazes de absorver as novas formas de circulação do texto na nossa produção acadêmica, abandonando os limites da emulação do impresso, para produzir formas textuais que possam ser efetivamente organizadas, encontradas, e *lidas*. Mas isso nos leva a um último alerta – pois essas novas formas textuais fugirão, inevitavelmente, ao nosso controle. Ou, nas palavras otimistas de Crane et al. (2008), esse processo irá “abrir a vida intelectual como nunca antes”. O problema, como coloca Eggert (2010) de modo extremamente contundente em discussão correlata, é que, nesse processo, tememos ver os bárbaros invadirem nossa cidadela:

46

...this vision of a common, interactive type of scholarship and readership that democratically puts the reader in a box-seat while also empowering the scholar to make and sign more expert editions, doing much of the discovery work for us, is a very attractive prospect. We should not fear the barbarians entering the gates of the scholarly city. They have usually got less arduous things to do anyway, and even when they do decide to interfere (as in contentious passages from books of the Bible whose wordings they object to) or even if they are empowered to make their own editions, my response is: Let them!

[11] #eggert2010 (meu grifo).

Enquanto temermos “os bárbaros” – debatendo nossas premissas, e (o horror!) apontando lapsos nas nossas transcrições – não nos aventuraremos nos experimentos da textualidade móvel vislumbrados por Crane et al. (2008), nem a qualquer tipo de “leitura e produção acadêmica pública e interativa” tal como proposto por Eggert (2010), e, muito menos, ao *texto-jogo* de Unsworth (2006). Continuaremos seguros; mas, talvez, ignorados do lado de fora do muro – continuaremos, de fato, não sendo *lidos*. E no avesso, ao fecharmos os olhos para a barbárie do outro lado do muro, nos distanciaremos da nossa tarefa milenar de organizar a palavra escrita; ficaremos seguros, mas não seremos mais os leitores e escritores do “Arquivo” no sentido de Pêcheux (1994[1982]) – ou seja: já não seremos aqueles que inscrevem a *Memória*.

47

4. Por fim

No começo desta conversa, entre muitas perguntas, questionava-me como podíamos ler a “Prosa do Mundo” hoje, e

48

que quadro nos permitiria operar uma ordenação dos seres e das “coisas”, inspirada pela experiência marcante da leitura de *As Palavras e as Coisas* pelo olho eletrônico de um navegador na rede mundial de computadores. Chego ao final da conversa sem a resposta, mas com a desconfiança de que o espaço que ordena a “Prosa do mundo” hoje já não é um espaço fisicamente delimitado – em contraste com os espaços da leitura até pouco tempo, limitados, por exemplo, pela nossa possibilidade física de percorrer bibliotecas feitas de concreto para encontrar o que ler. Nossa “tábua de trabalho” agora não é um espaço no mundo físico: nela, potencialmente, podem se encontrar todos os livros, todas as imagens, todas as palavras, e todas as coisas do mundo.

De um ponto de vista discursivo, como argumentei, vemos instaurarem-se assim novas formas de leitura que transformam as *formas de atenção* que podemos dirigir aos nossos objetos de trabalho. Busquei aqui salientar a importância de compreendermos a nova *realidade material* do texto (esse instrumento e objeto da atenção privilegiada das Humanidades); e tentei considerar como parte dessa realidade material todo o processo que envolve a sua produção, difusão e leitura. Defendi, assim, que essa dissolução dos limites físicos da leitura e do texto conforma um novo estado discursivo para as Humanidades.

49

Os desafios colocados por esta nova conformação das Humanidades frente às tecnologias de difusão digital da informação vão muito além do que o debate sobre as chamadas *Humanidades Digitais* pode fazer crer. Esse debate não é inteiramente alheio às questões discutidas aqui – mas é, talvez, apenas uma faceta particularmente visível do problema maior. Como tenho proposto em outras oportunidades (particularmente, em Paixão de Sousa, 2015), o debate em torno das humanidades digitais tem estado, muitas vezes, preso ao âmbito terminológico em torno das *Digital humanities*. Entretanto, como mostramos em Priani Saisó et al. (2014), esse termo não apresenta uma consistência de uso no espaço – sendo as *Humanidades digitais*, bem como as *Humanidades digitais*, um significante muito menos presente no mundo lusófono e “hispanohablante”, mesmo entre pesquisadores cujos trabalhos envolvem intenso uso de tecnologias computacionais em humanidades. Assim, poderíamos aproveitar alguns dos apontamentos críticos sobre as humanidades digitais para essa reflexão mais ampla, em particular os que têm levantado a importante questão do anglocentrismo e da desigualdade dentro do campo (Fiormonte, 2012, 2014; Dacos, 2013). Quanto àqueles que tem abordado as humanidades digitais como uma tendência passageira [Marche 2013], se formos além do que se oculta no rótulo *Humanidades Digitais*, veremos que as questões não são tão facilmente dispensáveis – pois o problema não é de tendências, e sim é trazido pela inserção do trabalho em Humanidades na lógica da difusão digital do texto, uma realidade presente e agente, quer gostemos dela, quer não. Por isso tudo, não me parece exagerado afirmar que a inclusão das tecnologias digitais de difusão da informação está transformando as Humanidades de uma forma radical e profunda, como já afirmaram Unsworth (2001, 2006), Crane et al. (2008), e Baumann & Crane (2010).

50

Para finalizar, sugiro que precisamos compreender essa nova forma da ordem do mundo dos textos, para sermos – os *humanistas* – um pouco mais sujeitos e um pouco menos objetos dessas transformações. Assim, se estamos diante de uma nova forma de leitura e de ordenação do texto, precisamos resolver que papel teremos nessa nova ordem – nessa nova forma do “Arquivo”, nos termos de Pêcheux (1994[1982]). Esta é a encruzilhada em que nos encontramos, como resultado de uma transformação profunda na natureza do nosso instrumento principal de trabalho, o objeto da nossa mais dedicada atenção – o texto. Em alguns casos, não só o texto, mas o texto antigo, preservado dos efeitos do tempo e trazido até nós por diferentes modos de difusão (e formas de atenção) por milênios – e que ainda hoje, me parece, é nossa tarefa preservar (e explicar, e organizar, e trazer à luz). Independente dos rumos de cada campo de estudos específico no interior das Humanidades, o texto nos está chegando, e continuará a nos chegar, na forma digital – nossos documentos, nossos “objetos de atenção”, estão sendo rematerializados em arquivos eletrônicos que armazenam sequências de números. Isso levou alguns autores a enxergar, de fato, uma revolução documental, num processo de *desconstrução* da própria noção de documento [Gradmann and Meister 2008], ou *fragmentação da unidade documental* [Chaudiron and Ihadjadene 2008], ou *redocumentarização* [Pédauque 2004]^[12], em propostas que seriam plenamente desenvolvidas em Pédauque, 2006, 2007.

51

Cabe a nós decidir que postura queremos ter diante disso: se queremos ser afogados pelo dilúvio da transformação material (aparentemente ainda crescente) dos milhões de “documentos” (i.e., arquivos de informações numéricas) dispersos em bancos de dados codificados, ou se queremos compreender as condições tecnológicas e históricas dessa rematerialização, e ressignificá-la dentro das nossas disciplinas – e talvez com isso, reversamente,

52

ressignificarmos nossas disciplinas, absorvendo as novas formas de trabalhar o texto. A incursão ao reino da técnica computacional, assim, demanda das Humanidades uma autoanálise crítica, já que nos obriga a fiar com novas fibras aquele fio com que se forma o tecido do trabalho das “ciências humanas”: o fio dos nossos olhares de leitura e dos nossos mecanismos interpretativos.

Resumo. Nesse texto proponho uma reflexão sobre alguns efeitos trazidos pelas condições materiais de difusão do texto digital para as Humanidades, em particular no que remete a novas formas de leitura e de ordenação da leitura. Partindo de observações sobre a dispersão do texto no meio digital, investigo materialmente as formas de produção e ordenação do texto nesse meio pontuando como singularidade fundante a inclusão de uma etapa lógica artificial em seu processo de difusão (a partir de propostas já delineadas em Paixão de Sousa, 2017). Essa contingência, argumento, permite considerar a difusão digital como etapa radicalmente nova na história da escrita e da leitura, em sentido semelhante aos apontados por Pédaque (2004, 2006, 2007); Crane et al. (2008,) Gradmann & Meister (2008), Chaudiron et al. (2008), e Baumann & Crane (2010). Sugiro ainda (a partir do conceito de *formas de atenção* trabalhado por Unsworth, 2006) que, ao moldar novas formas de leitura, essa mudança no sentido material do texto altera profundamente o trabalho tradicional nas humanidades, instaurando uma nova conformação discursiva para o campo. Defendo, por fim, que essa nova condição deve ser motivo de reflexão crítica nos diversos âmbitos das humanidades, mas especialmente naqueles que têm, no texto, seu objeto central de atenção.

Notes

[1] Este “preâmbulo” parte de algumas ideias delineadas originalmente em um post publicado em <https://humanidadesdigitais.org/2012/04/10/ler-a-prosa-do-mundo-hoje>, do qual também retomei o título para o presente artigo, “Ler a prosa do mundo hoje”. Como já apontava ali, esse título faz homenagem ao famoso ensaio de Michel Pêcheux, “Ler o arquivo hoje”, incluído no livro *Gestos de Leitura*, de Eni Puccinelli Orlandi.

[2] Ainda disponível, em <https://humanidadesdigitais.org/pequena-biblioteca-virtual-as-palavras-e-as-coisas>

[3] “A proposta deste artigo é a de que a metodologia conhecida como representação do conhecimento tem profundas implicações para a computação em humanidades, e por meio da computação em humanidades, tem o potencial de mudar o modo como o fazer acadêmico nas humanidades é realizado, mudar a natureza da formação nas humanidades, e mudar a relação entre as humanidades e outras profissões, que dirá outras disciplinas. Acredito que a representação do conhecimento já produziu pesquisas novas importantes, e que trará, no futuro, novos olhares sobre o que sabemos do registro humano, e sobre como o sabemos”; minha tradução.”

[4] “As ferramentas atualmente a nosso dispor, por primitivas que poderão parecer no futuro, são já adequadas para criar uma esfera dinâmica para a vida intelectual que é tão diferente daquilo que a precede como a cultura oral difere do mundo da escrita”; minha tradução.

[5] “Treebanks (corpora anotados) são coleções de textos com extensas categorias de anotação morfológica, sintática e similares, e constituem instrumentos familiares na pesquisa de corpus e de linguística computacional. Ao construir corpora anotados para línguas stóricas como o grego e o latim, encontramos um novo espaço intelectual que combinava elementos da linguística computacional e de corpus e da antiga disciplina da filologia”; minha tradução.

[6] Esse primeiro dado, note-se, é pouco confiável, pois, apesar da investigação pioneira do pesquisador, é preciso lembrar que não há propriamente uma entidade oficial que centralize as informações sobre o que há “na internet”, e portanto os números, inferidos indiretamente, correm o risco de estarem subestimados.

[7] Penso que estamos chegando a um momento no qual a forma da atenção que damos às fontes primárias está se deslocando da digitação para a análise, de artefatos para agregados, e da representação para a abstração; minha tradução.

[8] Em Paixão de Sousa (2010), mostrei como a correlação “internet-biblioteca” é falha, discutindo os problemas do acervo “percirculante” e da angústia do catalogador, e defendendo que aos humanistas nos resta dominar as novas formas de ordenação do texto, ou deixar que outros a dominem – e nessa segunda hipótese, abrir mão do coração do nosso trabalho e, desconfio, da nossa relevância social.

[9] “A tecnologia digital não é novidade nos estudos clássicos: há hoje professores plenos que sempre pesquisaram repositórios de textos gregos e latinos, compuseram suas ideias em forma eletrônica, encontraram fontes secundárias online, e que exploraram de modo oportunista quaisquer ferramentas digitais que servissem seus propósitos. Entretanto, a inércia das práticas pregressas fez preservarem-se intactas as formas de evoluírem para explorar as vantagens e minimizar as desvantagens da cultura impressa: criamos documentos que mimetizam

escravizadamente seus predecessores impressos; enviamos esses documentos para os mesmos tipos de periódicos e editoras; nossas obras de referência e nossas edições já começam a se tornar obsoletas no momento em que são publicadas e daí para frente ficam estagnadas; mesmo quando novas, nossas publicações são estáticas e não conseguem se adaptar às necessidades de seus usuários variados; enquanto um público crescente, global, poderia agora encontrar os resultados do nosso trabalho, nós encapsulamos nossas ideias em linguagem especializada e por trás de barreiras de assinaturas que fazem perpetuar, para o século vinte e um, o minúsculo público leitor do vinte.”; minha tradução.

[10] “A questão que se nos depara é muito mais profunda que o desafio de produzir mais, ou, de preferência, melhores artigos e monografias. Precisamos, de modo mais geral, nos perguntar que tipo de espaço queremos produzir para explorar o registro linguístico da humanidade – quer estejamos contemplando a Odisséia, os arquivos administrativos da Suméria, ou investigando o pensamento matemático por meio de fontes gregas ou árabes. Mais importante, talvez, do que a questão sobre o que podemos fazer, pode ser a oportunidade de redefinirmos quem pode fazer o quê – de abrir a vida intelectual como nunca antes, e de criar um solo fértil onde a humanidade poderá cultivar a vida da mente com maior vigor e alegria.”; minha tradução.

[11] “... essa visão de um tipo de leitura e produção acadêmica pública e interativa que democraticamente coloca o leitor num camarote e ao mesmo tempo empodera o estudioso a produzir e assinar mais edições especializadas, fazendo boa parte do trabalho de descoberta para nós, é uma perspectiva muito atraente. Não devíamos temer os bárbaros passando pelos portões da cidadela acadêmica. Eles normalmente têm coisas menos trabalhosas a fazer, de qualquer maneira; e mesmo quando decidem interferir (como no caso de passagens controversas da bíblia a cujos termos eles apresentem objeções), ou mesmo se se sentirem encorajados a produzir suas próprias edições, minha resposta é: Que façam!”; minha tradução.

[12] “L’objet texte ici désigné nest pas un objet purement sémiotique, ce nest pas seulement un ensemble de codes ; cest un objet fabriqué, artisanalement et/ou industriellement (dans le régime des médias informatisés par un mixte des deux), doté de propriétés matérielles, inscrit dans un certain type déchange” [Pédaque 2004, ¶10] (meus grifos).

Works Cited

- Artières and Berg 2010** Artières, Philippe; Bert, Jean-François. “La Bibliothèque Foucaldienne”. [internet]. Centre Michel Foucault, École normale supérieure de Lyon; 2010. Disponível em <http://lbf-ehess.ens-lyon.fr/pages/index.html>.
- Belon 1555** Belon, Pierre. *Histoire de la nature des oiseaux*. Paris, 1555. Imagens destacadas, Bibliothèque Interuniversitaire de Santé: <http://www2.biusante.parisdescartes.fr/img/?refphot=09330>, <http://www2.biusante.parisdescartes.fr/img/?refphot=09331>.
- Berry 2011** Berry, David M. “The computational turn. Culture Machine:” thinking about the Digital Humanities. 2011, 12. Disponível em <http://culturemacne.net/index.php/cm/article/view/440/470>.
- Betts 2009** Betts, Russell. *Colloquium on Digital Humanities and Computer Science*, Chicago, 2009.
- Chaudiron and Iahadjadene 2008** Chaudiron, S., Iahadjadene, M, Maredj, A. “La fragmentation et l’unité documentaire em question”. In: Actes ... CONGRÈS DE LA SFSIC, 16. Compiègne, 2008.
- Clement et al. 2008** Clement, T., Steger, S., Unsworth, J., & Uszkalo, K. (2008) “How Not to Read a Million Books”, accessed 21 June 2010. Disponível em <http://www3.isrl.illinois.edu/~unsworth/hownot2read.html#sdendnote4sym>.
- Crane et al. 2008** Crane, Gregory; Bamman, David; Babeu, Alison. *ePhilology: when the books talk to their readers*. Blackwell Companion to Digital Literary Studies, R. Siemens; S. Schreibman (eds). Oxford: Blackwell, 2008. <http://www.digitalhumanities.org/companion>.
- Dacos 2013** Dacos, Marin. “La stratégie du Sauna finlandais. Les frontières de Digital Humanities. Essai de Géographie politique d’une communauté scientifique”, 2013. <http://bn.hypotheses.org/11138>.
- Dewey 2015** Dewey, Caitlin. “If you could print out the whole Internet, how many pages would it be?” The Washington Post, 2105. Disponível em https://www.washingtonpost.com/news/the-intersect/wp/2015/05/18/if-you-could-print-out-the-whole-internet-how-many-pages-would-it-be/?utm_term=.1cb813d38f45.
- Fiormonte 2012** Fiormonte, Domenico. “Towards a Cultural Critique of the Digital Humanities”. *Historical Social Research / Historische Sozialforschung*, vol. 37, no. 3 (141), 2012, pp. 59–76. JSTOR, www.jstor.org/stable/41636597.
- Fiormonte 2014** Fiormonte, D. “A quando i Brics della conoscenza?” *Infolet: Cultura e critica dei media digitali*. Número 6, 2014.i Disponível em <https://infolet.it/?issue=numero-6>.

- Foucault 2000** Foucault, Michel. *[As palavras e as coisas: uma arqueologia das ciências humanas]*. Tradução de Salma Tannus Muchail para Les mots et les choses: une archéologie des sciences humaines, 1966. São Paulo: Martins Fontes; 2000.
- Gradmann and Meister 2008** Gradmann, Stefan; Meister, Jan Christoph. "Digital document and interpretation: re-thinking "text" and scholarship in electronic settings". *Poiesis Prax* (2008) 5:139–153. DOI 10.1007/s10202-007-0042-y. <http://link.springer.com/content/pdf/10.1007%2Fs10202-007-0042-y.pdf>.
- Kermode 2016** Kermode F. *Forms of attention*. The University of Chicago Press, Chicago; 2016 [1985].
- Khabsa 2014** Khabsa M, Giles CL (2014) "The Number of Scholarly Documents on the Public Web". *PLoS ONE* 9(5): e93949. <https://doi.org/10.1371/journal.pone.0093949>.
- Kirschenbaum 2010** Kirschenbaum, Matthew G. "What is Digital Humanities and what is it doing in your English department?", *ADE Bulletin*, 150, 2010. https://humanidadesdigitais.files.wordpress.com/2011/09/kirschenbaum_whatisdigitalhumanities.pdf.
- Kunder 2018** Kunder, Maurice. "The size of the World Wide Web". Disponível em <http://www.worldwidewebsite.com>. Acesso em: 01 mar. 2018.
- Marche 2013** Marche, Stephen. "Literature is not Data: Against Digital Humanities", 2013. <http://lareviewofbooks.org/essay/literature-is-not-data-against-digital-humanities>.
- Meister 2012** Meister, Jan Christoph. "DH is us or on the unbearable lightness of a shared methodology". In: *Historical Social Research* 37 (2012), 3, pp. 77-85. URN: <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-378413>.
- Paixão de Sousa 2010** Paixão de Sousa, Maria Clara. "O Leitor na biblioteca digital". II Seminário Mindlin: O Futuro das bibliotecas, São Paulo, 2010. Disponível em <http://oleitornabibliotecadigital.wordpress.com>.
- Paixão de Sousa 2013** Paixão de Sousa, Maria Clara. "Texto digital: Uma perspectiva material". *Revista ANPOLL* (Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística). Volume 1, Número 35, 2013. ISSN: 1982-7830. Disponível em <https://revistadaanpoll.emnuvens.com.br/revista/article/view/643712>.
- Paixão de Sousa 2014** Paixão de Sousa, Maria Clara. "O Corpus Tycho Brahe: contribuições para as humanidades digitais no Brasil". *Filologia e Linguística Portuguesa*, v. 16, p. 53-93, 2014. <http://www.revistas.usp.br/flp/article/view/88404>.
- Paixão de Sousa 2015** Paixão de Sousa, M. C. "As Humanidades Digitais Globais?" Ciclo de Conferências: Congresso Humanidades Digitais em Portugal (Universidade Nova de Lisboa, 8/10/2015), CIDEHUS (Universidade de Évora, 6/10/2015), Programa Materialidades da Literatura (Universidade de Coimbra, 12/10/2015). Disponível em <http://humanidadesdigitais.org/hd2015/>.
- Peckel 2012** Peckel, Joris. "The State of Digitisation". [Internet]. Open Knowledge Foundation / DM2E, 2012. Disponível em <https://openglam.org/2012/08/03/the-state-of-digitisation>; acesso em 12/12/2012.
- Porta 1655** Porta G. *La physionomie humaine de lean Baptiste Porta neapolitain, divisée en quatre liures : enrichie de quantité de figures tirées au naturel, ou par les signes extérieurs du corps, on voit si clairement la complexion, les mœurs, & les desseins des hommes, qu'on semble penetrer iusques au plus profond de leurs ames*. A Rouen: Chez lean & David Berthelin; 1655. Disponível em https://arcve.org/details/gri_33125008642148.
- Priani Saisó et al. 2014** Priani Saisó, Ernesto, et. al. "Las humanidades digitales en español y portugués. Un estudio de caso: DíaHD/DiaHD". *ANUARIO AMERICANISTA EUROPEO*, 2221-3872, N° 12, 2014, Sección Tema Central p. 5-18 5. <http://www.red-redial.net/revista/anuario-americanista-europeo/article/view/267>.
- Pédauque 2004** Pédauque, Roger T. "Document et texte: Permanence et transformations". In: *HAL-SHS- Sciences de l'Homme et de la Société*. Version du 15-06-2004. Disponível em http://halshs.arcv.es-ouvertes.fr/sic_00001003. Acesso em: 03 de jul. 2013.
- Pédauque 2006** Pédauque, Roger T. *Le document à la lumière du numérique*. Caen, France: C & F, 2006.
- Pédauque 2007** Pédauque, Roger T. *La redocumentarisation du monde*. Toulouse: Cepaduès, 2007.
- Pêcheux 1982** Pêcheux, Michel. "[Ler o arquivo hoje]". Tradução de Maria das Graças Lopes Morin do Amaral para Lire l'archive aujourd'hui, 1982. In E. Orlandi et al. (Orgs.), *Gestos de leitura: da história no discurso*. Tradução: Bethânia S. C. Mariani [et. al]. Campinas: Editora da Unicamp, 1994, p.55-66 (Coleção Repertórios).
- Schreibman et al. 2004** Schreibman, Susan; Siemens, Ray; Unsworth, John (eds.). *A Companion to Digital Humanities*. Oxford: Blackwell, 2004. <http://www.digitalhumanities.org/companion>.

Unsworth 2001 Unsworth, John. "Knowledge Representation in Humanities Computing". Lecture I in the *eHumanities* NEH Lecture Series on Technology & the Humanities, Washington, DC, April 3, 2001. Disponível em: <http://www.people.virginia.edu/~jmu2m/FOA/>.

Unsworth 2006 Unsworth, John. "Forms of Attention: Digital Humanities Beyond Representation." Paper delivered at "The Face of Text: Computer-Assisted Text Analysis in the Humanities," III Conference of the Canadian Symposium on Text Analysis (CaSTA), McMaster University, November 19-21, 2006. <http://people.brandeis.edu/~unsworth/FOA/>.