


ESTUDOS PALEOGRÁFICOS: GÊNESIS, EVOLUÇÃO E TENDÊNCIAS ATUAIS

Antonio Ackel*

 <http://orcid.org/0000-0002-8283-4417>

Como citar este artigo: ACKEL, A. Estudos paleográficos: gênese, evolução e tendências atuais. *Todas as Letras – Revista de Língua e Literatura*, São Paulo, v. 23, n. 1, p. 1-26, jan./abr. 2021. DOI 10.5935/1980-6914/eLETDO2114367

Submissão: março de 2021. **Aceite:** março de 2021.

Resumo: O presente texto parte de uma reflexão sobre uma concepção do objeto de estudo da Paleografia, busca refletir acerca de seu conceito, considerando-a uma ciência autônoma, aborda fatos históricos para resgatar sua trajetória até chegar ao futuro dos estudos paleográficos, vinculados às Humanidades Digitais. Nessa última parte, são descritos problemas que têm sido enfrentados durante tentativas de automação no reconhecimento de elementos morfografemáticos, bem como uma proposta de resolução. O que será apresentado faz parte de uma pesquisa de maior escopo, ainda em fase inicial e, como exemplo, utilizou-se parte da coleção documental do Fundo da Companhia Índias Ocidentais, guardado no Arquivo Real de Haia.

Palavras-chave: Humanidades digitais. Paleografia digital. Leitura e reconhecimento. HTR. Manuscritos setecentistas.

* Universidade de São Paulo (USP), São Paulo, SP, Brasil; Universiteit Leiden, Leiden, Holanda (bolsista Capes/Print – 2021). E-mail: antonio.ackel@usp.br

INTRODUÇÃO

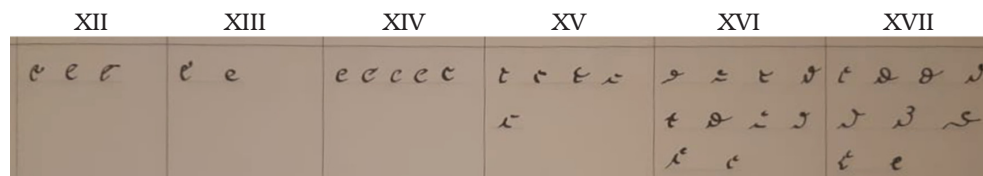
■ **E**ste artigo pretende discutir sobre o fazer paleográfico diante das questões que se colocam na leitura e decifração de manuscritos.

Para tanto, discorre-se inicialmente sobre como é concebido o documento, objeto de estudo da Paleografia, e de que forma este pode ser estudado e analisado, a partir de suas múltiplas características, materiais e gráficas, original e digitalizado. Propõe-se uma organização de elementos de um documento com o objetivo de refletir sobre diferentes perspectivas adotadas na análise documental.

Na sequência, tenta-se compreender os significados do termo *paleografia* e agrupá-los de forma que possam abranger a teoria e os métodos dessa ciência quando colocados em prática. Por natureza, a Paleografia demanda leitura e decifração de elementos gráficos contidos em um manuscrito. Isso significa dizer que se faz necessário também compreender a história e a evolução dos modos de escrita.

A exemplo do que se verá, para uma leitura exata do tipo de escrita que será apresentado, precisou-se recorrer, em função do material utilizado como exemplo, a manuais de paleografia holandesa que indicassem a evolução das formas das letras. Abaixo, é possível notar a diferença de formas que uma letra possui a depender do século em que foi escrita e como tais formas podem dar margem a interpretações equivocadas.

Figura 1 – Evolução por século da letra <e> nos Países Baixos



Fonte: Dekker, Roland e Dechamps (1992, p. 187).

Buscando compreender os passos que levaram aos estudos apresentados neste artigo, propõe-se uma leitura dos fatos históricos que desencadearam a necessidade de leitura e decifração de documentos manuais, tornando a Paleografia um saber pautado em teoria e método até a possibilidade de automação dessas atividades de forma precisa a fim de promover o avanço dessa ciência. Nota-se que o desenvolvimento de técnicas paleográficas representa um movimento circular que sempre retorna ao seu objetivo primeiro e final: compreender um manuscrito.

Por fim, na tentativa de colaborar com hipóteses para os estudos documentais, apresenta-se um projeto, em fase inicial, que pretende implementar um sistema computacional que utiliza um banco de dados para o reconhecimento de letras e palavras manuscritas em língua portuguesa.

O OBJETO DOS ESTUDOS PALEOGRÁFICOS

O documento é uma forma de materializar um pensamento. Mesmo depois da invenção da imprensa¹, o ato de escrever à mão existe e existirá.

Desde a Antiguidade, o termo *documento* pode denotar variados referentes. Para Júlio César, por exemplo, o termo tinha o sentido de *aviso*; para Tito Lívio, *amostra* ou *ensinamento*; para Cícero, era o *modelo*. É certo que não eram usos aleatórios, mas sujeitos às circunstâncias.

O fato é que, assim como nos tempos dos clássicos, na nossa língua, hoje em dia, podemos chamar um documento de muitos nomes: declaração, certificado, contrato, diploma, carta, bilhete, manuscrito, testemunho e por aí vão outros nomes que mudam de acordo com sua função. Como se verá mais adiante, o termo *modelo*, utilizado por Cícero, será retomado nos estudos da Paleografia Digital no tocante à Diplomática.

Ainda que múltiplas designações façam referências a diferentes áreas de atuação, o termo refere-se a uma mesma ideia: um objeto manuscrito que informa algo.

Efêmeros por natureza, eventos, atos, fatos precisam do documento para possibilitar informação sobre eles; é o único que fica. Objeto histórico, o documento permite defender ou não domínios lógicos, históricos e ontológicos. A elucubração de um assunto qualquer faz-se também por meio de documentos e do que neles está escrito.

É fonte de informação, guarda a memória do que o homem quis ou precisou preservar quando grafou. Nem sempre está em bom estado, nem sempre é facilmente acessado, mas, quando estudado, torna-se imprescindível reconhecê-lo em todas as suas características para que se possa adentrar em sua materialidade constitutiva.

Resumindo, o documento é o objeto de estudo da Paleografia. Ao haver contato entre o pesquisador e esse objeto, deve-se desenvolver um processo de recordar acontecimentos que o fizeram ser produzido.

Entre erros e acertos, deve-se compreender que um documento não se apresenta em estado puro, antes é revestido de estrita dependência da sua relação com o seu entorno diacrônico e sincrônico.

A informação sobre o objeto em si mesmo, sobre a autoria do texto, sobre as circunstâncias sociais e históricas que fizeram parte da produção do texto, sobre os modos de circulação e de transmissão, até seu lugar de pouso (onde ele está guardado) e tantas ocorrências que o envolvem são relevantes para a compreensão de suas escrituras.

Identificar no suporte todos os elementos materiais e gráficos que o caracterizam como fonte de informação é essencial para que se compreendam funções que poderíamos chamar de *spino-transcendentes*².

Para questões de facilitação na compreensão do que o documento pode representar na pesquisa, a seguir mostra-se um exemplo de organização prática que pode servir para sua descrição. Divide-se em três níveis de caracterização:

-
- 1 O termo *imprensa* faz referência a máquinas pelas quais o texto e as imagens são transferidos para o papel ou outra mídia por meio de tinta. Embora o papel tenha aparecido pela primeira vez na China, foi na Europa que a impressão tornou-se mecanizada pela primeira vez. A primeira menção a uma impressora tipográfica ocorre em um processo em Estrasburgo, em 1439, revelando a construção de uma impressora feita por Johannes Gutenberg e associados.
 - 2 O neologismo faz referência ao que o professor Segismundo Spina (1977, p. 76-77) chama de *função transcendente*, "que faz com que o texto deixe de ser um fim em si mesmo para transformar-se num instrumento que permite ao filólogo reconstruir a história da cultura de uma sociedade de determinada época". Parece pertinente substituir aqui o termo *filólogo* por *paleógrafo*.

1. *material*, que se refere ao tipo de suporte e de seu estado de conservação, por exemplo: se é papel, tecido; se é a tela do celular³; se há elementos adicionados como manchas, dobras, rasgos, carimbos; qual o instrumento utilizado para escrever – tinta, grafite, ouro;
2. *gráfico*, de onde se extraem elementos que foram produzidos pelo homem com um instrumento, por exemplo, uma assinatura, de que modo o texto foi disposto no papel, se há destaque de fragmentos como datas, vocativos, anotações, qual a forma da letra, se há intervenções de terceiros;
3. *essencial*, que provê a informação sobre o que nele está escrito e, por consequência, sobre a história que circunstanciou sua produção. Em alguma medida, poder-se-ia dizer que, nesse nível, importam mais os documentos sobre o documento estudado do que ele mesmo.

Não se pretende aqui estabelecer regras de classificação ou composição de um documento para seu estudo e descrição. Quer-se apenas contribuir com novos olhares que ampliem as dimensões de estudos sobre um objeto textual, para o bem da ciência paleográfica.

Nesse sentido, é oferecido o exemplo de esquematização acima para que se possa controlar e ordenar características documentais.

DO QUE TRATA A PALEOGRAFIA

Dentre as questões iniciais que se colocam ao especialista sobre essa ciência atualmente, destacam-se, entre outras, sua concepção, seu objeto de estudo, sua delimitação, sua terminologia, seus métodos de aplicação. Longe de tentar responder a qualquer uma delas, propõem-se aqui diferentes perspectivas, de diversos especialistas em Paleografia.

Etimologicamente, o termo *paleografia* procede do grego e tem um significado bastante amplo. Deriva de *παλαιός*, que significa antigo ou primitivo, e *γραφή*, com o sentido de *escrever*. O sufixo *-ia*, também do grego, permite conceber, além de outros significados, o de instituição ou movimento de que se faz parte, assim pode-se traduzir como uma *ciência que estuda escrituras antigas*.

A Epigrafia também estuda escrituras antigas, mas somente aquelas que estão em suportes rígidos, como arcos de triunfo, lápides e outros que exigem que a letra seja esculpida. À Paleografia cabe ocupar-se com as escrituras que estão em suportes maleáveis, dobráveis (como papel e pergaminho) e que foram realizadas com instrumentos mais delicados (como pena, estilete, lápis). À Codicologia cabe analisar e descrever os suportes dos textos da Epigrafia e da Paleografia.

Contudo, a expressão *escritas antigas* poderá causar certa estranheza, já que, além de soar um tanto simplista, suscita discussões entre pesquisadores, especialmente sobre temporalidade.

Assim como a Arte ou a Filosofia, a Paleografia possui diferentes escolas. No entanto, todas elas têm por objetivo ler corretamente um documento e com-

3 Problematizações de cunho codicológico ou epigráfico não são descritas neste artigo. No exemplo “tela do celular”, quer-se apenas mostrar a possibilidade de serem aplicadas técnicas paleográficas em documentação também digital. Para uma abordagem inicial, ver Ackel e Módolo (2020).

preender a sua história. Desde que decidiu registrar feitos e acontecimentos, o homem tem escrito a vida ao longo dos tempos.

Ler corretamente um documento significa identificar letras, compreender suas diferentes morfologias, interpretar abreviaturas e símbolos gráficos, reconhecer todos os conteúdos gráficos. Investigar a história de um documento significa concebê-lo para além do seu texto, fazer uma leitura sobre o período, o local e os fatos que consubstanciaram naquela produção textual, situá-lo a partir de seu autor, de sua função e de seu estado de conservação.

Vejam o que dizem alguns especialistas e professores que trabalharam e trabalham com Paleografia.

Millares Carlo (1929) referiu-se ao termo como a ciência que trata do conhecimento e interpretação das escrituras antigas e que estuda suas origens e evolução.

Mallon (1952) disse que a Paleografia deve ocupar-se de todo monumento gráfico de qualquer tipo e natureza, o que inclui inscrições em suportes rígidos como paredes, castelos e outras construções.

Castillo Gómez e Sáez (1994) afirmaram que não se trata simplesmente de uma disciplina que estuda as escrituras antigas, mas de uma ciência renovada que considera o estudo global da história dos usos e práticas da escrita. Portanto, é a história dos signos gráficos e também dos usos e das práticas da escrita.

Cencetti (1994) indicou que o objetivo da Paleografia é, na verdade, não só interpretar precisamente os antigos manuscritos, mas também datá-los, localizá-los e, de modo geral, apontar, a partir de sua materialidade exógena, todos os elementos úteis para o estudo do seu conteúdo e, num plano mais amplo, para a história da cultura em geral.

Nuñez Contreras (1994) definiu-a como a ciência que, com um método próprio, estuda o desenvolvimento do processo gráfico, considerando a escrita uma faculdade própria e exclusiva do homem.

Petrucci (1992) representou-a como o estudo histórico do desenvolvimento da escrita como expressão cultural, ampliando seu escopo ao adicionar às quatro questões fundamentais da Paleografia tradicional, que tratam da produção dos testemunhos escritos – *o que está escrito, quando foi escrito, onde e como foi escrito* –, duas novas: *quem escreveu e por que o fez*.

De acordo com Pérez Herrero e Rivero Suárez (2006), há tipos diversos de Paleografia, como a *Paleografia de leitura* ou *Paleografia prática*, que pretende ensinar a ler corretamente os documentos. Trata-se de um primeiro nível, responsável por situar erros que um documento possa apresentar. Com essas informações, é possível fornecer transcrições fidedignas que se tornam fontes de pesquisa a outros campos do conhecimento.

Erros de leitura podem desencadear críticas e interpretações errôneas. Vejam-se, por exemplo, as duas palavras abaixo, digitalizadas do manuscrito VOC_006, do fundo da Companhia das Índias Ocidentais, pertencente ao Arquivo Real de Haia, que trata de assuntos da colonização dos holandeses no Brasil (1637-1654)⁴.

4 O documento faz parte do fundo arquivístico Vereenigde Oostindische Compagnie, do Arquivo Real de Haia, na Holanda. Foi encontrado no catálogo *O Brasil em arquivos neerlandeses (1634-1654)*, de Marianne Wiesebron (2005), sob a cota VOC_006_HaNa_041, e consultado pessoalmente no arquivo holandês.

Figura 2 – Palavras couzas e Saude


Fonte: Arquivo Real de Haia.

Em uma das sessões de Crítica Textual da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (Anpoll), em 2020, ao ser apresentado o manuscrito em questão, chegou-se a aventar a possibilidade de que o grafema <n>, encontrado nas palavras *sande* (saúde) e *conzas* (cousas), pudesse indicar um fenômeno de mudança nas estruturas silábicas. Num primeiro momento, considerou-se a possibilidade de uma fusão por contato fonético, como uma labialização (TRASK, 1992), em que se considerasse <n> e <d>; <n> e <z>, pelo ponto de articulação. Como alternativa de interpretação, propôs-se também que houvesse relação morfológica entre a palavra *sande* e a forma latina *sanu*.

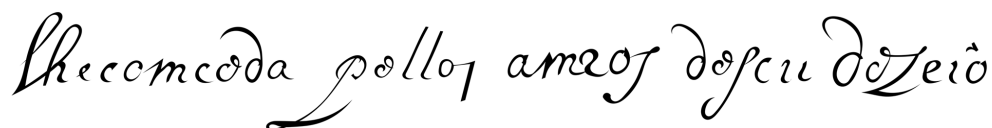
Essas conjecturações, baseadas em teorias da linguística histórica, foram desacreditadas quando se concluiu que, pela comparação com outros documentos coevos ao período a que esse se referia (século XVII), a letra <u> possuía essa forma semelhante à letra <n>.

Ainda de acordo com Pérez Herrero e Rivero Suárez (2006), um segundo nível pode ser representado pela *Paleografia científica*, também chamada *Paleografia crítica* ou *de análise*. Nesse nível, busca-se estudar os signos gráficos para determinar as características da escritura e suas mudanças. Dessa forma, o estudioso poderá saber em que época e lugar um documento foi produzido, se se trata de um texto original ou não e até buscar a interpretação de formas de escritas desconhecidas.

O significado dessa atividade está entre os possíveis limites de compreender a razão dos traços grafados pela mão do autor. Como exemplo, utiliza-se o mesmo manuscrito VOC_006 em que se observa uma forma bastante particular da letra <e>, que pode ser identificada na palavra *Saude* da Figura 2.

Trata-se de uma forma classificada, segundo Marín Martínez (1991), como pré-cortesã, que se refere à variedade gráfica localizada entre a letra de *albalaes* e a cortesã, ou seja, meados do século XIV e primeiro quartel do século XV.

A cursividade da escrita que se observa no manuscrito leva a uma característica morfológica em que a velocidade se sobrepõe à caligrafia, fazendo com que essa letra <e>, de matriz gótica, apresente alógrafos em algumas palavras e que podem levar à confusão com a letra /o/, como nas palavras *the comceda pellos de seu dezeio*, digitalizadas abaixo.

Figura 3 – Frase com alógrafos de <e>


Fonte: Arquivo Real de Haia.

Essa forma de letra <e> é bastante comum e pode ser encontrada em todos os documentos escritos em língua holandesa do mesmo fundo arquivístico. Então, o fato de essa letra ter sido identificada em um manuscrito escrito em língua portuguesa sugere interpretações de cunho histórico sobre uma possível influência da escrita holandesa no aprendizado da escrita no Brasil em uma determinada região e em uma determinada época.

Para que se chegue a qualquer hipótese sobre a história de um documento e de seu autor, essa perspectiva paleográfica requer uma avaliação dos traços elementares que constituem uma determinada tipologia de escrita. Podem ser consideradas estruturas individuais caligráficas e também as formas clássicas de uma escrita. Para tanto, apresenta-se um experimento, na seção 4.1, que trabalha com automação de leitura e decifração de manuscritos.

As características principais a serem estudadas para compreensão de um elemento gráfico contemplam a inclinação da letra, que é a direção e acentuação que os traços seguem a partir da verticalidade de um ângulo de 90°. Como ilustração, na Figura 4, pode-se observar na escrita cursiva do século XVII uma inclinação, bastante acentuada para a direita, das palavras unidas *MeuPrincipe*, retirada do manuscrito cuja autoria é atribuída ao Padre Antonio Vieira⁵.

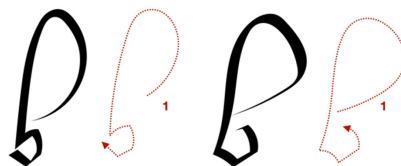
Figura 4 – Inclinação de *MeuPrincipe*



Fonte: Instituto de Estudos Brasileiros.

Outra característica a ser observada em uma análise paleográfica é o *ductus*, que se refere à ordem de sucessão e o sentido de execução dos traços que compõem cada uma das letras. É possível, ao se estudar uma escritura, acompanhar o processo de constituição de um signo gráfico. Por exemplo, na Figura 5, pode-se observar a letra , digitalizada a partir do códice quinhentista, *Le livre de trois vertus*, da autora Christine de Pizan⁶. No alógrafo da esquerda, seu corpo foi grafado de cima para baixo e da direita para a esquerda, e no da direita, pode-se conferir outro movimento, de baixo para cima e da esquerda para a direita. Ao lado direito de cada letra, há um pontilhado que corresponde ao traçado e aos movimentos que o escriba fez. A direção das setas indica o sentido.

Figura 5 – Alógrafos de *b*



Fonte: Biblioteca Nacional da Espanha⁷.

5 O manuscrito faz parte de um códice sob a cota 135-007, da coleção Alberto Lamego. Está disponível apenas para consulta pessoal no Instituto de Estudos Brasileiros da USP.

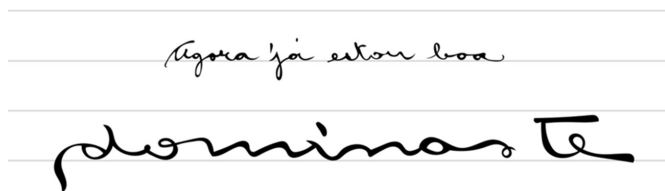
6 O manuscrito está disponível para consulta no portal da Biblioteca Nacional da Espanha (<https://www.bne.es>), sob a cota Ms. 11.515.

7 Disponível em: <https://www.bne.es>

Uma letra é formada por um corpo que pode vir adicionado de uma haste ou de uma cauda. No exemplo acima, temos uma haste e um corpo que compõem a letra . A partir dessas composições entre corpo, haste e cauda, as letras indicam seu módulo, ou seja, a relação de tamanho entre sua largura e altura.

Essa é outra característica morfológica que pode ser analisada. No exemplo abaixo, foram digitalizadas duas formas de escrita diferentes com relação ao seu módulo. Ambas foram retiradas de cartas de pacientes que ficaram internados no Sanatório Pinel, em São Paulo (1929-1944)⁸. Pode-se notar claramente a diferença de módulos entre os dois tipos de escrita.

Figura 6 – Diferença de módulos



Fonte: Arquivo Público do Estado de São Paulo.

Outras características como peso, cursividade, regramento ideal, sinais abreviativos podem ser identificados e analisados em uma escrita. Dado o limite deste artigo, exemplificam-se as anteriores apenas. Todo o processo de análise paleográfica requer rigor teórico-metodológico que permita acompanhar o desenvolvimento do ato de escrever.

Há também um terceiro nível de Paleografia proposto por Tanodi, em 1979, definido como *Paleografia cultural*, centrada na escritura e em seu contexto cultural, sobretudo em aspectos não perceptíveis por meio de outras fontes. Sob essa perspectiva, escrever é um ato sociocultural cujo estudo requer ampliação de horizontes, conhecimento histórico e de todas as implicações que a escrita em sociedades específicas supunha.

Trata-se de desenvolver uma técnica que seja capaz de resolver problemas suscitados pela escrita como prática sociocultural. Como observado, Petrucci (1992), ao introduzir as questões culturais já mencionadas aos estudos paleográficos, propõe teoria e metodologia novas que se articulam ao que hoje se conhece como alfabetização e cultura escrita.

Cherubini e Pratesi (2010) chamam de Paleografia a história da evolução contínua dos signos do alfabeto latino e daqueles que o integram, desde sua formação inicial até as manifestações mais recentes, investigados com metodologia própria nos mecanismos intrínsecos e nas causas externas que concorrem de vez em quando para modificar sua estrutura e aparência.

Muitos outros trabalhos ainda poderiam ser citados aqui, como os de Maffei (1727), Traube (1907), Masai (1956), Marichal (1963), Casamassima (1964), Ganz (1982), Blay (1986), Costamagna (1987), Brown (1993). Quando se procura definir uma ciência com tamanha diversidade de concepções, tenta-se ser o

⁸ A coleção de manuscritos pertence ao Fundo Pinel e pode ser consultada apenas pessoalmente no Arquivo Público do Estado de São Paulo.

mais sucinto e abrangente possível. Para a elaboração deste texto, buscou-se, além das referências citadas, outras descrições acerca de Paleografia.

Mesmo que não se possa resumir a imensa significação, definição, classificação que o termo, por natureza, oferece, propõe-se compreender Paleografia como a ciência que leva a indagar criticamente a origem e a evolução da escrita e de suas formas. Ela ensina a ler e decifrar todos os caracteres gráficos, manuscritos ou impressos, em qualquer tipo de documento e permite compreender o tempo, o espaço e a história por meio de todos os elementos materiais, gráficos e essenciais nele contidos.

UMA BREVE HISTÓRIA DA PALEOGRAFIA

A arte de decifrar documentos nasce da arte das falsificações. Desde muito tempo, o uso de práticas de recriação documental tem levado o homem a questionar a validade de documentos, textos, letras, assinaturas e, conseqüentemente, a estabelecer procedimentos para o seu reconhecimento.

Uma das primeiras identificações de adulteração documental remonta ao século IX, tempo de Carlos, o Calvo, pela obra *Decretos de Pseudo-Isidoro*, que dispunha, entre outras, de coleções sobre legislações de monarcas (*Capitularia Benedicti Levitae*), procedimentos criminais (*Capitula Angilramni*) e cartas com assinaturas dos papas dos primeiros três séculos.

Também o *Capítulo Geral da Ordem*, já de 1220, em que se lê “*falsariis litterarum Domini Papae*” (falsificadores de documentos papais), faz clara referência à prática ilícita.

Por tantas obras recriadas, todas elas revestidas de formas múltiplas e complexas, ao longo de toda a Idade Média, é que se procurou estabelecer métodos e regras para a verificação da autenticidade e integridade de um documento.

Em fins do século XVII, o monge beneditino da Congregação de São Mauro, Mabillon, escreveu um grande volume (*Libri VI*) chamado *De re diplomatica*.

Nessa obra, Mabillon iniciou suas pesquisas desenvolvendo um método científico com base no reconhecimento de formas escritas, descrição e cópia de letras e outros símbolos gráficos contidos em livros antigos, conhecidos como códices, e em outros documentos. Podem ser encontradas escrituras em latim dos séculos XIV e XV, organizadas por data e tipo de escrita.

Seu trabalho era um marco para toda a comunidade acadêmica que se debruçava sobre documentos antigos. Seu propósito era fornecer um método claro e inquestionável para reconhecer a tradição documental de originais e cópias, bem como a autenticidade e confiabilidade de *diplomas*, palavra que se referia a um conjunto de documentos como cartas, atas, tratados, contratos e outros registros gerados durante a Idade Média.

O termo *diplomática* vem de seus estudos e é derivado do latim *diploma*, que significa um escrito dobrado em dois, *duplo*.

Já a palavra *paleografia* só foi empregada pela primeira vez no início do século XVIII, quando outro monge, da mesma congregação, Dom Bernard de Montfaucon, publicou em Paris, em 1708, a obra *Palæographia Græca*. Nela, trabalhou de forma tão detalhada com manuscritos gregos que serviu de referência, no mínimo, pelos dois séculos seguintes. Nesse livro, o monge também reflete sobre *codicologia*.

Desde então, o termo *codicologia*, também estendido a caracteres alfabéticos que não os gregos, passou a ser comumente usado por estudiosos, mesmo que seu escopo e abrangência nem sempre coincidisse com as interpretações dadas por Montfaucon.

Ainda que os monges mauristas houvessem identificado letras como a gótica, a saxônica e outras próprias dos povos germânicos, deve-se considerar que o trabalho de Scipione Maffei, de Verona, publicado em 1727, sobre a divisão dos modos de escrita em maiúscula, minúscula e cursiva, foi o ponto de partida para outros estudos da história da escrita.

De fato, a continuação de suas pesquisas se deu novamente com outros monges mauristas, Dom Charles Toustain e Dom René Tassin, com a publicação *Nouveau traité de diplomatique* (Paris, 1750, especialmente o v. III). A obra consiste em uma coletânea de todas as variedades do latim, além de uma organização sistemática sobre sua evolução.

No entanto, nessa época, a Paleografia ainda era considerada uma atividade auxiliar da Diplomática e baseava-se apenas na correta decifração de escrituras antigas e na datação de manuscritos.

Já no século XIX, Léopold Delisle e Ludwig Traube, da École Nationale des Chartes, e Émile Chatelain, da École Pratique des Hautes Études, na França, foram os responsáveis por aumentar o alcance da Paleografia, dando-lhe autonomia por seus resultados e aliando-a à Filologia e, mais importante, atribuindo-lhe o papel de uma ciência que possibilitou conhecer mais sobre a história do texto e seus processos de transmissão.

Na Alemanha, Wilhelm Wattenbach publicou seus manuais, em 1871, *Anleitung zur lateinischen Paläographie*, e, em 1878, *Das Schriftwesen im Mittelalter*, corroborando os estudos franceses.

Sob influência de um de seus alunos, Chatelain dedica considerável pesquisa ao uso das notas tironianas e publicou, em 1900, *Introduction à la lecture des notes tironniennes*. Nesse volume, apresenta as regras gerais para que um símbolo gráfico seja considerado nota tironiana:

Uma nota tironiana consiste, normalmente, de dois elementos: um sinal radical ou principal e uma terminação ou sinal auxiliar. O sinal que expressa a terminação é menor do que o que representa o radical. Assim, pode-se reconhecer facilmente terminações como os, um, a, ta. O radical pode ser expresso: 1. apenas pela letra inicial; 2. pela sílaba inicial; 3. por várias letras que compõem a palavra; 4. por várias letras da palavra colocadas em uma ordem diferente (CHATELAIN, 1900, p. 22-23, tradução nossa)⁹.

Para dar exemplos de notas tironianas, apresentam-se abaixo duas imagens retiradas do códice quinhentista, *Le livre de trois vertus*, da autora Christine de Pizan¹⁰. A primeira representa uma forma utilizada para indicar a palavra Χριστός, em português, *Cristo*. Inicia-se com o radical formado pelas letras gregas <X>, <ρ> e <o>, e uma terminação que indica a supressão das letras <ι>,

⁹ No original: "Une note tironienne se compose, ordinairement, de deux éléments, un radical ou signe principal et une terminaison ou signe auxiliaire. Le signe exprimant la terminaison est plus petit que celui qui représente le radical. Ainsi dans on reconnaît aisément les désinences os, um, a, ta. Le radical peut être exprimé: 1. par la lettre initiale seule. 2. par la syllabe initiale. 3. par plusieurs lettres entrant dans la composition du mot: 4. par plusieurs lettres du mot placés dans un ordre différent".

¹⁰ Utilizou-se esse material como exemplo por apresentar variados exemplos de notas tironianas em todo o texto. Está disponível para consulta em: <https://bne.es>.

<σ>, <τ> e <ς>, representada pelo traço sobre as letras. A segunda nota tironiana apresenta um radical formado pelas letras <D> e <s> e uma terminação que representa as letras <e> e <u>, também representada por um traço, significando a palavra *Deus*.

Figura 7 – Notas tironianas



Fonte: Biblioteca Nacional da Espanha (<https://www.bne.es>).

Pode-se dizer que Ludwig Traube, professor de Filologia Medieval Latina na Universidade de Munique, incluiu definitivamente a Paleografia na história da cultura europeia, destacando sua importância como componente da pesquisa filológica. Em 1907, apresentou *Nomina sacra: Versuch einer Geschichte der christlichen Kürzung*, um estudo histórico e sistemático sobre nomes e títulos que, nos manuscritos gregos das escrituras sagradas, apareciam geralmente abreviados.

Sob supervisão de Traube, Elias Lowe publicou, em 1908, a partir de seu doutorado, pela Princeton University, *Codices Latini antiquiores* (CLA), em que é encontrada uma descrição paleográfica bastante detalhada de todos os manuscritos da literatura latina, datados até 800 d.C.

Ainda na passagem dos anos 1800 para 1900, na Itália, grandes pesquisadores se destacaram por sua contribuição à ciência paleográfica, dentre eles, Luigi Schiaparelli, com suas obras sobre a escrita latina, como *La scrittura latina nell'età romana*, de 1921, e *Note paleografiche (1910-1932) Raccolte a cura di Giorgio Cencetti*, de 1969.

Durante o início dos anos 1930, Bernard Bischoff era estudante de doutorado de Paul Lehmann, sucessor de Traube. Também nesse período, serviu de assistente a Lowe, na pesquisa dos CLA e, em 1979, publicou *Paläographie des römischen Altertums und des abendländischen Mittelalters*, um manual atualizado de suas publicações anteriores e que abrange escrituras do século I até o século XIV.

Vicenzo Federici foi professor de Paleografia e Diplomática na Università di Roma La Sapienza, onde teve como aluno o arquivista e filólogo Giorgio Cencetti. Em 1934, publicou a obra *La scrittura delle cancellerie italiane dal secolo XII al XVII: fac-simili per le scuole di paleografia degli archivi di stato*.

Professor de diversas universidades italianas, como Università Lateranense e Università di Macerata, Giulio Batelli foi um dos enviados do Vaticano para salvar o patrimônio arquivístico, bibliográfico e artístico em risco, em particular o da Abadia de Montecassino, durante a Segunda Guerra Mundial. Dentre suas obras, destaca-se o manual *Lezioni di paleografia, a cura della Pontificia Scuola Vaticana di Paleografia e Diplomatica*.

Giorgio Cencetti, tendo sido reitor da *Scuola speciale per archivisti e bibliotecari* e membro ministerial do *Consiglio superiore delle accademie e biblioteche*,

abordou a paleografia latina nas obras *Lineamenti di storia della scrittura latina e Vecchi e nuovi orientamenti nello studio di paleografia*, nas quais propôs os conceitos de “canonização” e “densificação caligráfica”, o que mudou a compreensão do desenvolvimento da escrita latina da Antiguidade e do início da Idade Média.

Após a Segunda Guerra Mundial, a pesquisa paleográfica foi profundamente condicionada pela publicação de *Paléographie romaine*, em 1952, de Jean Mallon. Progenitor e expoente mais radical da chamada *Nouvelle École Française*, representada não só por ele, mas também por Robert Marichal e Charles Perrat, visava a revolucionar a visão de futuro da escrita latina. O produto da nova interpretação do fenômeno gráfico organizou novos conceitos da evolução da escrita, principalmente com o estudo do *ductus*, além de propor uma genealogia de vários tipos de escrita.

Compreendendo algumas limitações da Paleografia ensinada até então, Armando Petrucci estendeu seu campo de investigação a todos os produtos escritos, sem limites cronológicos; coerentemente com sua orientação marxista, privilegiou o estudo da função e difusão social da escrita, da escrita das classes populares, dos vários níveis de letramento presentes em uma dada sociedade e em uma época.

Seus estudos renovaram o método paleográfico e orientaram novas pesquisas para além da erudição circunscrita em bibliotecas e arquivos de materiais medievais. Petrucci viu que às questões fundamentais da paleografia – *o que, quando, onde e como* – faltavam duas novas perguntas: *quem e por que*. A ampliação dessa curiosidade científica resultou em uma mudança substancial da investigação e dos resultados paleográficos.

Hoje, estudos paleográficos são realizados em grande número nas universidades. Muitas vezes associada à História, a Paleografia cede seu lugar de ciência primeira e mostra-se auxiliar a outras disciplinas.

Entretanto, aqui, pretende-se ressaltar a importância de conceber a ciência paleográfica como autônoma, que se coloca acima das pretensões acadêmicas quando se volta para sua função social, capacitando a geração de conhecimentos, a partir de teoria e metodologia próprias, para que possam ser aplicadas às práticas de análise documental.

Como exemplos de aplicações metodológicas, que também levam a pensar na própria ciência paleográfica como o objeto de estudo, apresentam-se, a seguir, pressupostos que podem ser adotados na investigação e na didática desse conhecimento.

Adaptando-se a tríplice distinção da paleografia – *de leitura, análise e cultural* –, pensa-se, em primeiro lugar, sobre os critérios que podem ser adotados para o procedimento da leitura de um manuscrito:

1. Há que se fazer uma leitura calma, atenta a todos os signos gráficos que estão no papel, mesmo que não sejam reconhecidos num primeiro momento.
2. Observar a característica de escrita do punho, compreender que a mesma letra pode mudar sua forma, como no caso dos alógrafos ou a depender de seu contexto (se no começo, meio ou fim de uma palavra).
3. Não se deve ignorar nenhum elemento gráfico entremeado na escrita. Num primeiro momento, pode parecer apenas um rabisco aleatório, mas depois poderá significar uma vírgula, um diacrítico ou uma abreviatura, como ilustrado anteriormente.

4. Criar e preencher um formulário que contenha todas as formas de letras, abreviaturas e outros símbolos do alfabeto utilizado na escrita que se estuda, para que possa servir de referência quando houver dúvidas. O método mais fácil e recorrente usado para a composição desse alfabeto é o recorte de imagens pelo computador e a colagem em um processador de textos, para a ordenação dos elementos alfabéticos, como mostrado na Figura 8.

Figura 8 – Construção do alfabeto do Ms. 355-004¹¹



Fonte: Instituto de Estudos Brasileiros.

Para melhor compreensão do conteúdo do texto, faz-se necessário o uso de diferentes tipos de dicionários, por exemplo, de abreviaturas, como os de Borges Nunes (1981) e Flexor (2020), de vocabulários coetâneos ao manuscrito, nesse caso, Bluteau (1789), e de vocabulários contemporâneos, como Houaiss (2001) e Ferreira (2014). Facilita também a criação de um glossário de palavras descobertas.

Uma prática bastante efetiva para a interpretação da leitura é transcrevê-la. Não caberia, neste momento, discorrer sobre critérios e tipologias de transcrições, mas apontar para a necessidade de estabelecer e adaptar normas para uma transcrição fidedigna que possa ser lida e comparada com o texto manuscrito.

No caso de se desenvolver uma análise paleográfica, deve-se ter em mente a etapa anterior, a leitura, que oferece uma rigorosa observação dos elementos

¹¹ O manuscrito faz parte do códice 355-004, da coleção Alberto Lamego. Está disponível apenas para consulta pessoal no Instituto de Estudos Brasileiros da Universidade de São Paulo. Trata-se de uma declaração assinada por Dom Macedo Carvalhal, em Lisboa, aos 23 de abril de 1727, informando sobre a partida da Fragata Nossa Senhora do Rozario, de Lisboa para o Rio de Janeiro, capitaneada por Jozeph Soares de Andrade.

gráficos constitutivos, de forma que é possível estabelecer uma relação entre escrita e suporte. Mais ainda, o pesquisador, a partir da identificação de diferentes formas gráficas, busca compreender o processo evolutivo gráfico.

Ao serem identificados tais elementos, o paleógrafo poderá também relacionar datações cronológica e tópica (às vezes aproximadas) daquela escrita (CARDONA, 1981). No estudo dos diferentes tipos de escrita e das evidências gráficas individuais, o pesquisador usa determinados elementos de análise e uma série precisa de termos relacionados tanto ao processo de composição dos sinais gráficos quanto à identificação de categorias de tipos de escrita.

Como exemplo de metodologia aplicada à análise paleográfica, observam-se os critérios e os termos utilizados por Mallon (1952) e Bischoff (1979) e utiliza-se a proposta criada por Ackel e Módolo (2020), que se refere a cópias manuais com instrumentos digitais. Essa última tem se mostrado eficiente na medida em que permite analisar elementos baseados na composição de uma escrita, como forma, módulo, *ductus*, ângulo, peso e também definir categorizações gerais da escrita, por exemplo, se maiúscula ou minúscula, se cursiva ou monolítica etc. Tem também oferecido meios facilitadores na aplicação da atividade de cópia, como aumento de visualização do elemento gráfico, para observar detalhes como velocidade e peso, repetição contínua dos movimentos de composição do signo no mesmo suporte ou reprodução da mesma cor da tinta observada no modelo.

Os resultados mostram-se promissores com relação à facilitação de leitura e análise paleográfica, à compreensão do processo de construção da escrita e às práticas de reconstrução de diversos tipos de escrita. Isso se dá, especialmente, pelo fato de permitir visualização mais detalhada de uma letra específica por meio de aumento e isolamento de uma unidade gráfica ou visualizações mais detalhadas do módulo da letra e separação de sobreposições.

Ainda que as técnicas de Mallon (1952) e Bischoff (1979) sejam discutidas, a delimitação da proposta de Ackel e Módolo (2020) é específica, no tocante à prática de leitura e cópia de manuscritos: 1. trata-se de selecionar o material para análise – a imagem digital a ser copiada –, localizar o fragmento com que se quer trabalhar; 2. a primeira identificação a ser feita é quais instrumento e tinta foram utilizados no documento digitalizado para que o resultado do traçado digital assumira efeitos e cores semelhantes ao que se vê no fac-símile; 3. uma vez identificado o tipo de instrumento utilizado, procede-se, pois, com o ato da cópia: retraça-se a letra, buscando, com precisão, refazer o caminho percorrido pelo instrumento utilizado no manuscrito.

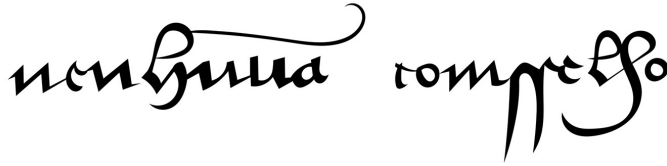
Copiar um movimento escritural novo requer assimilação de padrões gráficos por etapas. A sucessão de movimentos que um autor realizou em seu manuscrito pode não ser familiar àquele que o copia, dessa maneira, requer entendimento lógico e teórico sobre o tipo da escrita (pressão exercida sobre o suporte com o instrumento, ângulo do instrumento¹², *ductus* etc.).

Como exemplo dos resultados gráficos obtidos por essa proposta metodológica, mostra-se a Figura 9, em que se pode conjecturar que o ângulo da pena que o escriba usou foi entre 35° e 45°, para escrever as palavras *nenhūua* e

12 O ângulo da pena é medido pela relação entre a linha da escrita e a inclinação do instrumento utilizado. A partir de uma identificação do ângulo utilizado pelo escriba, pode-se avaliar o peso da escrita, ou seja, a natureza grossa ou fina dos traços constitutivos da letra.

conselho existentes no códice quincentista, *Le livre de trois vertus*, da autora Christine de Pizan¹³.

Figura 9 – Palavras *nenhūua* e *conselho*



Fonte: Biblioteca Nacional da Espanha¹⁴.

No que se refere à Paleografia cultural, é difícil estabelecer uma metodologia única para se conhecer um documento a partir das circunstâncias que motivaram sua produção, circulação, transmissão.

No entanto, há aspectos elementares que o pesquisador deve considerar no exame da gênese documental, por exemplo, o fato gráfico associado à criatividade humana, ou seja, a escrita não pode ser analisada como uma ocorrência isolada. Isso significa dizer que a Paleografia cultural é estudada por uma perspectiva da história social e da história da escrita.

Fundamenta-se a teoria em Petrucci (1992), com suas abordagens e problematizações sobre alfabetização e história da cultura escrita. Faz parte de sua discussão a correta identificação dos protagonistas da história: os autores que escreveram para a sociedade.

Daí propostas recentes que propõem responder a algumas perguntas sobre essas discussões têm buscado a confluência entre a história social e outras histórias. Por exemplo, Barros (1991, 1992) sobre os horizontes da história das mentalidades, Chartier (1993) e Burke (1993) no campo da história cultural e social.

Nesse sentido, a Paleografia cultural é entendida em virtude dessa confluência de abordagens, que, de outros ângulos e perspectivas, fazem dela uma ciência que abrange o humano, sua história e sua escrita.

Não se trata de uma justaposição de métodos e objetivos, mas de uma intersecção de modos de fazer história que surgiram do exame documental. Como dito, Petrucci (1992, p. XV, tradução nossa) inaugura um novo campo paleográfico ao elaborar outras questões que até então eram novas no meio acadêmico:

A própria formulação destas questões [de “quem” e “porquê”] subverte, em certo sentido, o método tradicional da paleografia, não só porque considera a escrita no contexto da sociedade que a produziu [...], mas sobretudo porque, em vez de partir das formas gráficas para depois ligá-las a outras manifestações da sociedade contemporânea, pretende partir propriamente do estudo do sentido que uma dada sociedade formada necessariamente por escritores e não escritores atribui à escrita e do conhecimento do número e da qualidade dos escritores nessa sociedade, em particular, ligar o estudo de todas as formas gráficas produzidas por aquela sociedade na sua variedade e no seu conjunto aos resultados

¹³ O manuscrito está disponível para consulta do portal da Biblioteca Nacional da Espanha, <https://www.bne.es>, sob a cota Ms. 11.515.

¹⁴ Disponível em: <https://www.bne.es>

*dessas pesquisas e explicar as atitudes gerais dessa sociedade com a lógica desta ligação e comparação, cultura e vice-versa, as particularidades das formas gráficas adotadas, suas mudanças, as influências estilísticas que sofreram ou exerceram*¹⁵.

Dessa forma, poder-se-ia conceber que todo o fazer paleográfico perfaz essas três etapas para consubstanciar-se: o primeiro contato com seu objeto de estudo será por meio de uma leitura. A leitura fornecerá subsídios necessários à investigação, análise, interpretação e compreensão documental. Essas resultados em conhecimento crítico sobre os fenômenos gráficos realizados pelo homem, ao longo da vida, na sociedade.

A CAMINHO DA PALEOGRAFIA DIGITAL

A versatilidade e a capacidade de combinar imagem e significado com técnicas analógicas e digitais têm contribuído sobremaneira para os estudos paleográficos.

De documentos históricos a contemporâneos, as técnicas de leitura e escrita com instrumentos e suportes digitais têm estabelecido pilares para o futuro da informação e documentação. Tal fato sugere reflexões prospectivas e antecipatórias, na medida em que, não em vão, o rico patrimônio histórico documental a que se tem acesso apenas pessoalmente em bibliotecas e arquivos ainda exige soluções que integrem inovações no acesso, tanto de natureza quantitativa quanto qualitativa.

Aliar a aplicação tecnológica à leitura e análise documental é função da Paleografia digital. Aqui, a expressão associa-se à habilidade no uso da tecnologia computacional à habilidade caligráfica (ACKEL; MÓDOLO, 2020) e ao conhecimento teórico-metodológico de análise paleográfica.

Isso considerando o escopo das humanidades digitais¹⁶, formadas pela contribuição transdisciplinar de diferentes profissionais, desde linguistas computacionais a desenvolvedores de ferramentas de visualização de dados.

Para se pensar em um primeiro conceito sobre essa associação, deve-se pensar no texto como imagem e significado, quer dizer, a Paleografia digital entende a escrita manual por meio de duas variáveis fundamentais: o texto como forma (parâmetros morfológicos) e o texto como significado (parâmetros semânticos), que vão desde o processamento de letras como conjuntos harmônicos de linhas que formam imagens identificáveis até a inclusão de vários metadados associados ao texto linguístico expresso em determinado *corpus*.

15 No original: "*L'impostazione stessa di queste domande [del 'chi' e del 'perché'] rovescia, in un certo senso, il metodo tradizionale della paleografia, non soltanto perché considera la scrittura nell'ambito della società che l'ha prodotta [...], ma soprattutto perché, invece di partire dalle forme grafiche per collegare poi queste ultime ad altre manifestazioni della società coeva, intende partire propriamente dallo studio del significato che una determinata società formata necessariamente di scriventi e non scriventi attribuiva alla scrittura, e dalla conoscenza del numero e della qualità degli scriventi in quella determinata società; per collegare poi ai risultati di queste ricerche lo studio di tutte le forme grafiche prodotte da quella società nella loro varietà e nel loro complesso, e spiegare con la logica di tale collegamento e di tale confronto gli atteggiamenti generali di quella società stessa rispetto alla scrittura e alla cultura, e, viceversa, le particolarità delle forme grafiche adottate, i loro mutamenti, le influenze stilistiche da esse subite o esercitate*".

16 "Humanidades digitais" é uma área de atividade acadêmica transdisciplinar que reúne computação ou tecnologias digitais e disciplinas das humanidades. Inclui o uso sistemático de recursos digitais nas Ciências Humanas, bem como a análise de sua aplicação (COHEN et al., 2008).

Dentre as múltiplas sistematizações possíveis, a partir dessas duas variáveis, podem-se destacar: imagens de fontes e sua evolução morfológica ao longo do tempo; geradores automáticos de padrões de cada ciclo escriturário; mapas de distribuição geográfica de variantes de letras; ferramentas que identificam modificações nas mãos do escriba.

Dar forma à aparência gráfica manuscrita de um lado (conceber o texto como imagem) e processar os dados associados ao texto como porção linguística (texto como significado) do outro mostra certa complementação para a formação da Paleografia digital. É fato que esse ramo inova o ensino e o tratamento do texto manuscrito.

Estudiosos da computação em Ciências Humanas estão mais do que familiarizados com as técnicas e tecnologias usadas para marcar dados textuais: a marcação de texto eletrônico e o desenvolvimento de ferramentas e tecnologias associadas continuam sendo o principal impulso da disciplina.

Esquemas de codificação, a partir do nível do caractere, que têm a intenção de representar todos os tipos de textos literários e linguísticos para pesquisa e ensino *on-line*, fornecem excelentes padrões para manipular digitalmente o texto e permitir pesquisas acadêmicas.

No entanto, faz falta para pesquisadores na área da Paleografia, que se debruçam sobre textos manuscritos, conseguir compreender um nível anterior ao do caractere. O foco está nos elementos básicos que compõem o caractere, como os já citados anteriormente.

A construção de bancos de dados paleográficos que capturam informações dessa natureza também é imprescindível para alimentar sistemas de inteligência artificial voltados para o auxílio à leitura de textos em documentos danificados e deteriorados.

Embora tenha havido algum interesse no desenvolvimento de esquemas de codificação para capturar esses dados textuais, poucos trabalhos até o momento desenvolvem técnicas sistemáticas que podem ser usadas para marcar dados – textos e imagens – nesse nível para uso nas Ciências Humanas.

Como exemplo e sem se deter exaustivamente nos termos e conceitos técnicos computacionais, descreve-se a seguir o método utilizado para se pensar em estudos futuros.

CLASSIFICAÇÃO DE MANUSCRITOS E HTR

Handwriting text recognition (HTR) é o processo de transformar letras ou frases manuscritas em um formato digital. Aqui, será apresentado um modo de reconhecimento *off-line*¹⁷. Para qualquer uma dessas atividades de reconhecimento, é necessário o envolvimento de pesquisadores das áreas da Ciência da Computação. O resultado do trabalho colabora com pesquisas acadêmicas que se valem de automação para serem mais proficuas, por exemplo, atestar verificação de assinaturas, reconhecimento de punhos, datações tópica e cronológica.

Os estudos paleográficos em HTR que aqui serão mencionados foram baseados em Scheidl (2018) e fazem parte de uma pesquisa de maior escopo que abrange manuscritos de diferentes naturezas. Para a elaboração deste texto, foi

¹⁷ Esse modo de processamento utiliza fac-símiles, manuscritos digitalizados e o *off-line* é executado enquanto o texto está sendo escrito (por exemplo, por meio da sensibilidade de pressão de um dispositivo).

utilizada, como ilustração dos estudos de caso, parte da coleção documental relativa à colonização holandesa no Brasil (1635-1654), guardada no Arquivo Real de Haia.

O propósito da pesquisa é encontrar soluções para os problemas que têm surgido ao longo de seu labor e abordar o reconhecimento de letras, *ductus*, eliminação de interferências como sobreposição, corrosão e manchas no papel, detecção de interferência de terceiros no documento. O propósito da seção é compartilhar os experimentos que têm sido desenvolvidos até o momento.

Atualmente, há uma grande variedade de projetos de transcrição e indexação de manuscritos históricos. Como exemplo do que vem sendo realizado no campo da Paleografia digital, podem ser citados: 1. o pioneiro *Digital Resource and Database of Palaeography, Manuscripts and Diplomatic* (DigiPal), do King's College de Londres, que oferece imagens digitais de caligrafia medieval com descrições e caracterizações detalhadas da escrita, além de incorporar diferentes formas de explorar e manipular a informação, como imagens marcadas, mapas dinâmicos e linhas do tempo e processamento de imagens¹⁸; e 2. o projeto "Cuneiform Digital Palaeography", da Universidade de Birmingham e do British Museum, custeado pela bolsa Leverhulme Trust, que procura estabelecer uma análise paleográfica bastante detalhada da escrita cuneiforme. O eixo central desse projeto é a construção de um banco de dados *on-line* que oferece imagens digitais de signos cuneiformes individuais retirados diretamente das fontes originais¹⁹.

A Paleografia digital tem levado pesquisadores a cada vez mais assumir um repensar metodológico e inovador na compreensão, tratamento, pesquisa e ensino de manuscritos.

Isso se dá porque à medida que o mundo tem se apoiado na utilização de tecnologias digitais e as novas gerações consideram os computadores, aparelhos móveis e a *internet* ferramentas indispensáveis para a execução de suas tarefas, o futuro das formas tradicionais de cultura, conhecimento e erudição parece estar em risco. O próprio *status* do texto, herdeiro de uma longa tradição de livros manuscritos e impressos, está evoluindo com a escrita multimídia, com tecnologias em constante desenvolvimento e novas expectativas do leitor. Vemos a criação dinâmica de novos espaços e mídias para o conhecimento gradualmente substituir a autoridade de objetos culturais seculares.

Muitos pesquisadores das Ciências Humanas e Sociais já trabalham há algum tempo com ferramentas digitais e estão ativamente empenhados em demonstrar como as novas funcionalidades possibilitadas por essas ferramentas enriquecem seu trabalho crítico, por exemplo, Buzzetti (2002), Ciula e Marras (2009), Paixão de Sousa (2014), Apollon, Bélisle e Régnier (2014). No entanto, as tecnologias digitais recentes para codificação e marcação de textos ainda não são familiares para a maioria dos pesquisadores que trabalham no campo da Paleografia.

18 Pode ser acessado em: <https://kclpure.kcl.ac.uk>

19 Disponível em: <https://research.birmingham.ac.uk/>

Esforços combinados e produção de ferramentas digitais significativas para o estudo de objetos de análise da Paleografia têm levantado questões sobre o futuro da prática e a habilidade dos pesquisadores em integrar tais ferramentas às suas pesquisas.

O estudo de leitura que é proposto aqui, como forma de ampliar o repertório de metodologias aplicadas à Paleografia digital, aborda o reconhecimento de manuscritos, fazendo uso de uma rede neural artificial (*artificial neural network* – ANN), por meio de aprendizagem profunda (*deep learning* – DL). Isso significa dizer que são utilizados sistemas de computador para imitar o processamento de dados do cérebro humano (com neurônios artificiais).

Por exemplo, da mesma forma que os olhos do paleógrafo procuram pelo elemento gráfico relevante à sua pesquisa em uma grande quantidade de manuscritos, o computador também executa todos esses passos, mas em segundos.

De forma técnica, o sistema utiliza-se de algoritmos para processar dados. A aprendizagem profunda, subárea da aprendizagem de máquina (*learning machine* – LM), usa camadas de neurônios matemáticos para compreender a fala humana (por exemplo, perguntar ou pedir algo para o celular) e reconhecer objetos visualmente (por exemplo, reconhecer características morfografemáticas, como se verá a seguir).

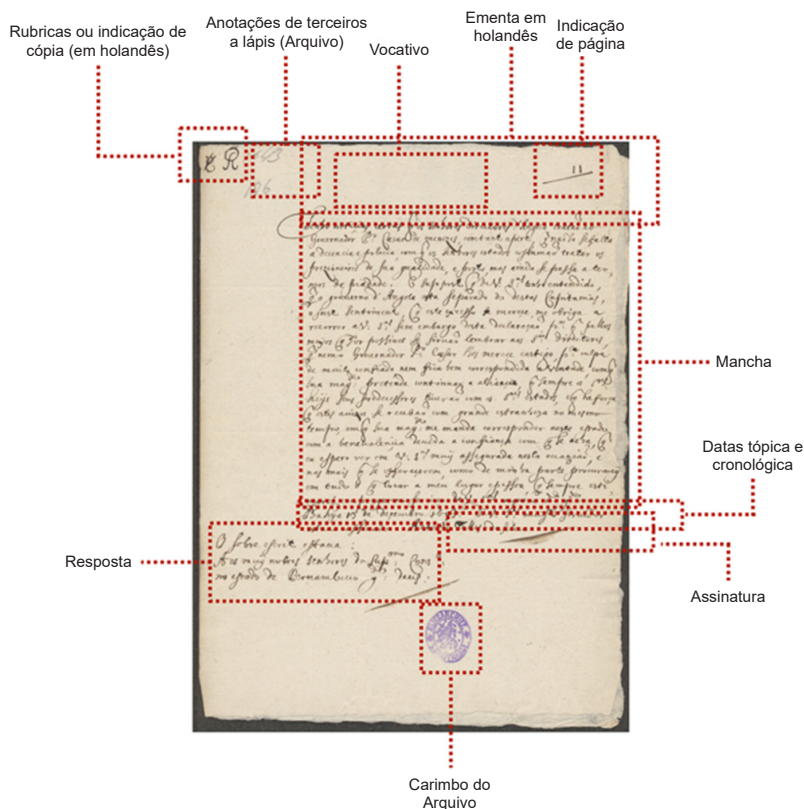
A informação é passada por meio de cada camada, com a saída da camada anterior fornecendo entrada para a próxima camada, e assim sucessivamente. Cada camada é tipicamente um algoritmo simples, ou seja, uma fórmula matemática que representa uma função de ativação.

A Figura 11 mostra a utilização de uma rede neural convolucional (*convolutional neural network* – CNN), que é uma arquitetura de redes especialmente desenhada para classificar imagens por meio de um cálculo de camadas. A CNN cria um algoritmo que é enviado para a rede neural recorrente (*recurrent neural network* – RNN). Essa rede reconhece os padrões em sequência de dados (texto ou imagem) e envia para uma operação de classificação temporal conexionista (*connectionist temporal classification* – CTC). Essa operação é o resultado da transcrição do elemento gráfico (BUNKE, 2003).

O sistema utilizado aqui compreende um modelo que possa apresentar todas as áreas da imagem em que os elementos gráficos estão localizados no suporte (com base na teoria e metodologia da Diplomática).

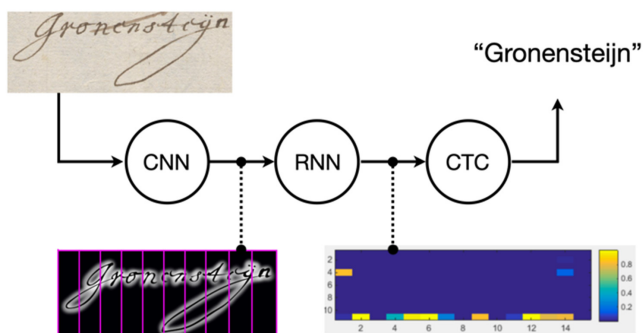
Como ilustração do primeiro modelo, a Figura 10 mostra a associação de cada uma das áreas marcadas no manuscrito para que o computador pudesse dizer se existe um padrão de escrita quanto à disposição textual. O documento refere-se a uma carta de Antonio Telles da Silva com justificativas sobre as condutas de Pedro Cesar de Menezes, escrita na Bahia, em 13 de dezembro de 1643²⁰.

²⁰ O documento faz parte do fundo arquivístico Vereenigde Oostindische Compagnie, do Arquivo Real de Haia, na Holanda. Foi encontrado no catálogo *O Brasil em arquivos neerlandeses (1634-1654)*, de Marianne Wiesebron (2005), sob a cota VOC_01a_HaNa_044 e consultado pessoalmente no arquivo holandês.

Figura 10 – Determinação de elementos gráficos distribuídos no manuscrito

Fonte: Elaborada pelo autor.

Abaixo, a palavra holandesa *Gronsteijn*, retirada do manuscrito VOC_006, em que a letra <e>, mencionada em seção anterior, foi reconhecida com facilidade pelo sistema e também as letras <ij>, que poderiam ter sido confundidas com <y>²¹.

Figura 11 – Ilustração do sistema HTR

Fonte: Elaborada pelo autor.

²¹ Ainda que o banco de dados se refira a manuscritos em língua portuguesa, a inclusão de formas estrangeiras tem se mostrado necessária no desenvolvimento desse estudo paleográfico.

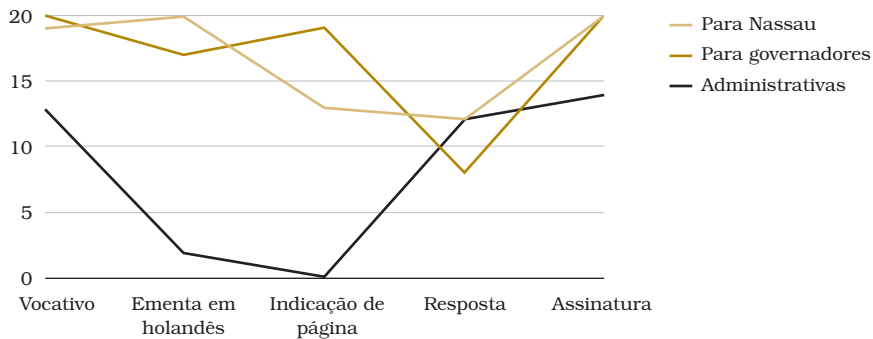
O reconhecimento de elementos gráficos manuscritos, como o caso acima, usa o método de aprendizagem de máquina (ML), que requer dados de entrada marcados para indicar o modelo. É a associação entre imagem e significado que o paleógrafo informa ao computador.

O sistema utilizado aqui compreende um modelo que indica o reconhecimento de símbolos gráficos nas porções textuais.

Para uma consulta de qualquer dado sobre essa coleção, formada por 6.700 documentos, aproximadamente 23.000 fólios, a máquina poderia oferecer muitas respostas com os mais variados filtros, indicando, inclusive, documentos que apresentam padrões diferentes.

Abaixo, mostra-se o resultado de uma pesquisa de ocorrência de elementos gráficos em 60 documentos, divididos igualmente nos seguintes conjuntos de dados: 20 cartas para Maurício de Nassau, 20 cartas para os governadores do estado de Pernambuco e 20 cartas enviadas a outras pessoas²².

Figura 12 – Estatística de conjuntos de dados



Conjunto de dados	Vocativo	Ementa em holandês	Indicação de página	Resposta	Assinatura
Para Nassau	19	20	13	12	20
Para governadores	20	17	19	8	20
Administrativas	13	2	0	12	14

Fonte: Elaborada pelo autor.

A escrita característica dessa coleção documental, como já se observou, tem se mostrado desafiadora quando se trata de reconhecer suas letras, especialmente no caso da cursividade (*cursive writing recognition* – CWR).

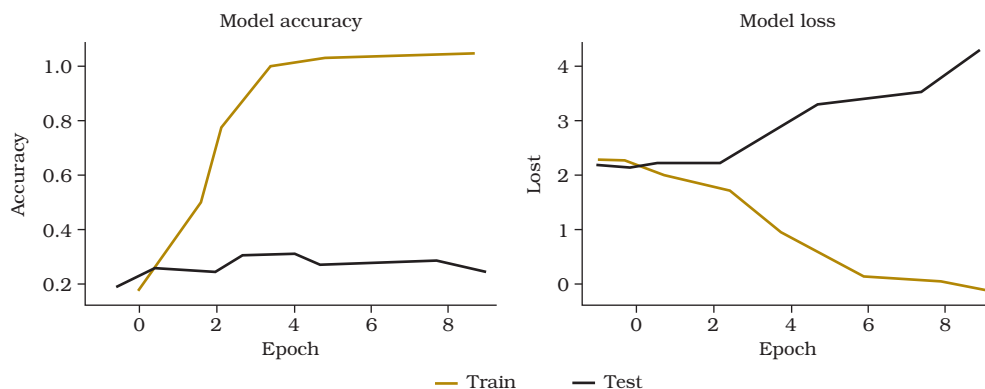
²² Toda a coleção de documentos foi identificada e classificada no catálogo de Wiesebron (2008). O fundo arquivístico Vereenigde Oostindische Compagnie, que compreende essa coleção, está sob custódia do Arquivo Real de Haia, na Holanda. Durante o ano de 2019, por meio da bolsa Erasmus+ICM, foi realizado um estágio de pós-graduação na Universidade de Groningen, pela qual se obteve acesso pessoal a esse fundo. Todos os documentos que estão em língua portuguesa nesse fundo, no total de 89 fólios, foram fac-similados e apresentados como objeto de estudo para um projeto de doutorado, aprovado pelo programa de Filologia e Língua Portuguesa da Universidade de São Paulo em 2020.

A exemplo das letras <e> e <u>, já citadas, que confundiram o reconhecimento automático das letras <o> e <n>, respectivamente, considerou-se apresentar aqui a aplicação do modelo CNN²³ para uma possível solução de parte do problema²⁴.

Nos manuscritos selecionados foram encontradas 122 letras <e>, no entanto, para que todas fossem reconhecidas pela máquina, foi necessário treiná-la, ou seja, cada vez que ela não lia uma forma diferente na letra <e>, era preciso adicionar o elemento gráfico novo como parte da letra.

Os dois gráficos da Figura 13 mostram, pela linha amarela, a diferença sensível de resultados depois que a máquina foi treinada para compreender as várias formas do grafema em questão.

Figura 13 – Resultados do experimento CNN



Fonte: Elaborada pelo autor.

A partir dos exemplos apresentados, nota-se que as possibilidades de exploração tecnológica em manuscritos vão muito além da transcrição de textos de uma forma automatizada e amparada na inteligência artificial.

CONSIDERAÇÕES FINAIS

Se desde o século IX, como foi apontado na breve recuperação de fatos históricos da ciência paleográfica, tinha-se como objetivo a leitura dos manuscritos para investigação e interpretação documental, hoje a era digital, com todas as novidades tecnológicas que pode oferecer, mantém essa atividade, facilitando e acelerando sobremaneira as pesquisas nessa área.

Pensando-se na possibilidade de contribuir com as atividades paleográficas aqui expostas, propõe-se validar novos métodos e projetar horizontes que se abrem para outros campos de pesquisa. Acredita-se que a Paleografia deve buscar sua construção científica por meio da complementaridade com ciências de outras áreas.

²³ A pesquisa visa a um estudo mais aprofundado na classificação do texto manuscrito e sua tradução para o formato digital. Dessa maneira, considera-se também utilizar o *recurrent neural network* (RNN) em HTR.

²⁴ Nessa etapa inicial da pesquisa, além da informação dos elementos gráficos e suas disposições no documento, formulários para coleta de amostras de caligrafia por palavras-chave estão sendo criados. Neste texto, foi selecionado um pequeno conjunto de dados sobre a morfologia da letra <e>.

Não à toa, foram incluídos os termos cultural e digital neste artigo para, antes de tudo, propor reflexões críticas no campo paleográfico que ofereçam visão de compartilhamento de informações e que sejam humanistas, a ponto de não ver mais uma obra documental de forma racional, isolada de seu contexto de produção social e histórico e isolada fisicamente em bibliotecas e arquivos de difícil acesso.

Em meio à documentação a ser estudada, o pesquisador pode aprofundar seu conhecimento sobre o objeto de estudo também por meio do reconhecimento dos protagonistas ou de comunidades, utilizando-se de métodos de interpretação e de narrativa que possam dar conta, no texto escrito, do entrecruzamento do social e do cultural.

A associação entre essas disciplinas é apresentada neste artigo com a intenção de contribuir com abordagens metodológicas e projetar um futuro paleográfico que se utilize de modernas técnicas em suas pesquisas, tendo sempre como objetivo a leitura correta de um documento.

Nesse sentido, a Paleografia, sem renunciar ao seu passado, reorienta práticas teórico-metodológicas, tornando-se um campo aberto a saberes de outras áreas.

A partir dos exemplos de práticas digitais que aqui foram descritas, planeja-se melhorar os modos de reconhecimento textual da máquina. Na realidade, acredita-se na leitura automática como complementação da leitura humana, ou seja, é vista como uma ferramenta que pode ser chave na integração de múltiplas fontes, todas elas relacionadas ao apoio e suporte da ciência paleográfica.

Com a emergência da necessidade de aplicativos e ferramentas tecnológicas desenvolvidas para trabalhos paleográficos, faz-se imperativo pensar nos modos de ampliação no tratamento de textos. Antes de uma descrição ou delimitação sobre essa nova disciplina paleográfica, fica aqui um chamado para especialistas de áreas correlatas construir saberes que possam ser compartilhados.

PALEOGRAPHY STUDIES: GENESIS, EVOLUTION, AND CURRENT TRENDS

Abstract: In this text, we discuss Paleography, considering several theoretical perspectives. It starts with a reflection about its object of study; we also consider historical facts to discuss future Digital Paleography studies. In the last section, we describe problems we are facing regarding reading and recognizing handwriting letters and words. Thus, our aim is to present possible positions to be adopted and activities to be developed within the scope of Digital Humanities. Manuscripts from The Hague Royal Archive's West Indian Company collection will be used as an example.

Keywords: Digital Humanities. Digital Paleography. Reading and recognizing. HTR. Seventh-century manuscripts.

REFERÊNCIAS

ACKEL, A.; MÓDOLO, M. Leitura e escrita de manuscritos: pena e suporte digitais na prática paleográfica. *Filologia e Linguística Portuguesa*, São Paulo, v. 22, n. 2, p. 233-245, jul./dez. 2020. Disponível em: <https://www.revistas.usp.br/flp/article/view/176142/169499>. Acesso em: 1º mar. 2021.

- APOLLON, D.; BÉLISLE, C.; RÉGNIER, P. *Digital critical editions*. Illinois: University of Illinois, 2014. Disponível em: <https://www.digitalcriticaleditions.com>. Acesso em: 10 jan. 2021.
- BARROS, C. La “Nouvelle Histoire” y sus criticos. *Manuscripts*, v. 9, p. 83-111, 1991.
- BARROS, C. Historia de las mentalidades, Historia social. *Temas Medievales*, Buenos Aires, v. 2, p. 205-230, 1992.
- BATELLI, G. *Lezioni di paleografia, a cura della Pontificia Scuola Vaticana di Paleografia e Diplomatica*. Città del Vaticano, Arquivo segreto vaticano, 1936.
- BISCHOFF, B. *Paläographie des römischen Altertums und des abendländischen Mittelalters*. Berlim: E. Schmidt, 1979.
- BLAY, G. *Las llamadas ciencias auxiliares de la historia: errónea interpretación?* Zaragoza: Institución Fernando el Católico, 1986.
- BLUTEAU, R. *Diccionario da lingua portugueza*. Lisboa: Simão Thaddeo Ferreira, 1789.
- BORGES NUNES, E. *Abreviaturas paleograficas portuguesas*. Lisboa: F. L., 1981.
- BROWN, J. *A palaeographer's view: the selected writings*. Londres: H. Miller, 1993.
- BUNKE, H. Recognition of cursive Roman handwriting: past, present and future. In: INTERNATIONAL CONFERENCE ON DOCUMENT ANALYSIS AND RECOGNITION, 7., 2003, Edinburgh. *Proceedings [...]*. Edinburgh, 2003. p. 448-459. DOI 10.1109/ICDAR.2003.1227707
- BURKE, P. La nueva historia socio-cultural. *Historia Social*, v. 17, p. 105-114, 1993.
- BUZZETTI, D. Digital representation and the text model. *New Literary History*, v. 33, p. 61-78, 2002.
- CARDONA, G. R. *Antropologia della scrittura*. Torino: Loescher Editore, 1981.
- CASAMASSIMA, E. Per una storia delle dottrine paleografiche dall'Umanesimo a Jean Mabillon, I. *Studi medievali*, 3e s., v. 5, p. 525-578, 1964.
- CASTILLO GÓMEZ, A.; SAÉZ, C. *La conquista del alfabeto: escritura y clases populares*. Gijón: Ediciones Trea, 1994.
- CENCETTI, G. *Lineamenti di storia della scrittura latina, dalle lezioni di paleografia* (Bologna, a.a. 1953-54). Bologna: Pàtron, 1994.
- CHARTIER, R. De la historia social de la cultura a la historia cultural de lo social. *Historia Social*, v. 17, p. 97-103, 1993.
- CHATELAIN, É. *Introduction à la lecture des notes tironiennes*. Paris: Chez l'Auteur, 1900. Disponível em: <https://archive.org/details/introductionala00chatgoog/page/n20/mode/2up>. Acesso em: 10 jan. 2021.
- CHERUBINI, P.; PRATESI, A. *Paleografia latina: l'avventura grafica del mondo occidentale*. Città del Vaticano: Scuola vaticana di paleografia, diplomatica e archivística, 2010.
- CIULA, A.; MARRAS, C. Circling around texts and language: towards 'pragmatic modelling' in digital humanities. *Digital Humanities Quarterly*, v. 10, n. 3, 2009. Disponível em: <http://www.digitalhumanities.org/dhq/vol/10/3/000258/000258.html>. Acesso em: 10 jan. 2021.

- COHEN, D. J. *et al.* Interchange: the promise of digital history. *The Journal of American History*, v. 95, n. 2, p. 452-491, 2008. Disponível em: <https://www.jstor.org/stable/25095630?seq=1>. Acesso em: 10 jan. 2021.
- COSTAMAGNA, G. *Perché scriviamo cosí: invito alla paleografia latina*. Roma: Viella, 1987. Disponível em: <http://archivi.beniculturali.it/Biblioteca/IndexCostamagna.html>. Acesso em: 10 jan. 2021.
- DEKKER, C.; ROLAND, B.; DECHAMPS, S. *Album Palaeographicum XVII Provinciarum*. Amsterdã: Brepols, 1992.
- FEDERICI, V. *La scrittura delle cancellerie italiane dal secolo XII al XVII: facsimili per le scuole di paleografia degli archivi di Stato*. Roma: Sansaini, 1934.
- FERREIRA, A. B. H. *Dicionário Aurélio. Século XXI*. Rio de Janeiro: Nova Fronteira, 2014.
- FLEXOR, M. H. O. *Abreviaturas: manuscritos luso-brasileiros dos séculos XVI ao XX*. Guaratinguetá: CRV, 2020.
- GANZ, D. Traube on Schrifttypen. *Scriptorium*, v. 36, p. 293-303, 1982.
- HOUAISS, A. *Dicionário Houaiss da língua portuguesa*. Rio de Janeiro: Objetiva, 2001.
- LOWE, E. A. *Codices Latini antiquiores: a palaeographical guide to Latin manuscripts prior to the ninth century*. Oxford: Oxford University Press, 1908.
- MABILLON, J. D. *De re diplomatica libri VI*. Paris, 1681. XIX-664. Disponível em: <http://141.20.85.26/mabillon/>. Acesso em: 1º mar. 2021.
- MAFFEI, S. *Istoria diplomatica, che serve da' introduzione all'arte*. Mantova: A. Tumermani, 1727.
- MALLON, J. *Paléographie romaine*. Madri: Consejo Superior de Investigaciones Científicas, 1952.
- MARICHAL, R. L'écriture latine et la civilisation occidentale du Ier au XVIe siècle. In: MARICHAL, R. *L'écriture et la psychologie des peuples*. Paris: Armand Colin, 1963. p. 199-247.
- MARÍN MARTÍNEZ, T. *Paleografía y diplomática*. Madri: Universidad Nacional de Educación a Distancia, 1991.
- MASAI, F. La paléographie gréco-latine, ses tâches, ses méthodes. *Scriptorium*, v. 10, p. 281-302, 1956.
- MILLARES CARLO, A. *Tratado de paleografía española*. Madri: Editorial Labor, 1929.
- MONTFAUCON, B. D. *Palæographia Græca sive de ortu et progressu literarum græcarum, et de variis omnium sæculorum Scriptionis Græcæ generibus: itemque de Abbreviationibus & de Notis variarum Artium ac Disciplinarum*. Paris: L. Guérin, J. Boudot e C. Robustel, 1708. Disponível em: <http://gallica.bnf.fr/ark:/12148/bpt6k1041709q/f7.image>. Acesso em: 1º mar. 2021.
- NUÑES CONTRERAS, L. *Manual de paleografía: fundamentos e historia de la escritura latina hasta el siglo VIII*. Madri: Cátedra, 1994.
- PAIXÃO DE SOUSA, M. C. O Corpus Tycho Brahe: contribuições para as humanidades digitais no Brasil. *Filologia e Linguística Portuguesa*, v. 16, p. 53-93, 2014. Disponível em: <http://www.revistas.usp.br/flp/article/view/88404>. Acesso em: 10 jan. 2021.

- PÉREZ HERRERO, E.; RIVERO SUÁREZ, B. *Escritura manuscrita y letra procesal: Canarias en el siglo XVI*. Las Palmas de Gran Canaria: Anroart, 2006.
- PETRUCCI, A. *Breve storia della scrittura latina*. Roma: Bagatto, 1992.
- SCHEIDL, H. *Handwritten text recognition in historical documents*. 2018. 286 f. Dissertação (Mestrado em Engenharia da Computação Visual) – Technische Universität Wien, Viena, 2018. Disponível em: <https://repositum.tuwien.at/handle/20.500.12708/5409>. Acesso em: 1º mar. 2021.
- SCHIAPARELLI, L. *La scrittura latina nell'età romana (note paleografiche): avvio allo studio della scrittura latina nel Medio Evo*. Cuomo: Tipografia editrice Ostinelli, 1921.
- SCHIAPARELLI, L. *Note paleografiche (1910-1932)*. Turin: Bottega d'Erasmus, 1969.
- SPINA, S. *Introdução à edótica: crítica textual*. São Paulo: Ars Poetica: Edusp, 1977.
- TANODI, A. *Manual de archivologia hispanoamericana*. Córdoba: Centro Interamericano de Desarrollo de Archivos, 1979.
- TASSIN, R.; TOUSTAIN, C. F. *Nouveau traité de diplomatie*. Paris: Guillaume Deprez e Pierre-Guillaume Cavelier, 1750-1765. Disponível em: <https://books.google.fr/books?id=QG3PAAAAMAAJ&dq>. Acesso em: 1º mar. 2021.
- TRASK, L. *Historical linguistics*. Londres: Arnold, 1992.
- TRAUBE, L. *Nomina sacra: Versuch einer Geschichte der christlichen Kürzung*. Munique: Franz Boll, 1907.
- WATTENBACH, W. *Anleitung zur lateinischen Paläographie*. Leipzig: Graz, 1871.
- WATTENBACH, W. *Das Schriftwesen im Mittelalter*. Leipzig: S. Hirzel, 1878. Disponível em: http://webserver.erwin-rauner.de/crophius/watt_images_plus.asp?quelle=watt. Acesso em: 1º mar. 2021.
- WIESEBRON, M. *O Brasil em arquivos neerlandeses (1634-1654)*. Documenten in het Koninklijk Huisarchief en in het archeif van de Staten-General. Documentos no arquivo da Casa Real e no Arquivo dos Estados Gerais. Volume 3. Série Mauritaniana. Leiden: CNWS, 2008.