

**PSICOFILOLOGIA:
UMA EVENTUAL FONTE DE SOLUÇÕES
PARA ALGUNS PROBLEMAS SOCIOFILOLÓGICOS?**

Thomas Daniel Finbow (USP)

thomas.finbow@usp.br

RESUMO

Nesta comunicação, proponho que conhecimentos desenvolvidos na psicolinguística e na psicologia experimental (especificamente nas suas vertentes que estudam a leitura), podem contribuir para esclarecer dúvidas de natureza filológica e sociofilológica sobre as relações entre língua, linguagem, usos técnicos e costumes metalinguísticos que caracterizavam a Baixa Idade Média, em particular na Península Ibérica. Inspirada na integração à linguística histórica da sociolinguística variacionista, sustentamos que a abordagem "psicofilológica" proposta contribuirá de maneira significativa para melhor compreendermos a evolução das relações entre o latim e seus descendentes românicos. O desenvolvimento de uma abordagem "sociofilológica" desde a década de 1980 tem trazido esclarecimentos importantíssimos a diferentes questões linguísticas e filológicas relacionadas à interligação das esferas linguística e cultural na linguística histórica românica, em particular, à *vexta quaestio* das origens das línguas neolatinas. Da mesma maneira, defendemos que a psicolinguística e a psicologia experimental da leitura trazem uma nova perspectiva para a análise filológica graças a certos fenômenos fisiológicos envolvidos no movimento ocular e nas capacidades visuais e do processamento mental do ser humano.

Palavras-chave: Psicofilologia. Sociofilologia. Sociolinguística. Psicolinguística.

1. Introdução

Neste artigo, proponho que a psicolinguística e a psicologia experimental da leitura podem contribuir para a resolução de certas dificuldades que surgiram devido à aplicação da sociolinguística à história das línguas românicas.

2. A abordagem sociofilológica e o monolinguismo complexo latino

A inspiração da abordagem interdisciplinar que apresento aqui, tal como a origem do problema que procuro solucionar, é o enfoque "sociofilológico" desenvolvido pelo hispanista britânico Roger Wright (2002) com o intuito de:

[...] integrar a análise filológica tradicional com as descobertas da sociolinguística moderna. Tal abordagem exige um conhecimento tão profundo quanto seja possível adquirir das circunstâncias históricas e intelectuais dos tabeliães, amanuenses e autores que nos forneceram as evidências. Embora a sociofilologia se aplique por definição à história linguística de comunidades alfabetizadas, enquanto o enfoque preferido da linguística histórica é, antes, ora a história de línguas com pouca ou até nenhuma história escrita, ora a pré-história de línguas que posteriormente se tornaram línguas literárias, é de suma importância que as duas disciplinas se complementem sem contradições. Isto é, entretanto, bastante difícil. O sociofilólogo bem-sucedido deve esforçar-se para dominar diversas áreas acadêmicas ao mesmo tempo, além de compreender uma ampla gama de dados. (WRIGHT, 2002, p. vii. Tradução nossa)

Wright questiona certos aspectos das explicações para o surgimento das línguas neolatinas à luz dos descobrimentos da sociolinguística. Primeiro, o filólogo britânico critica duramente a noção de que existia uma concepção metalinguística dentre os latinófonos tardo-antigos e alto-medievais de eles falarem duas *línguas* numa relação de bilinguismo ou de diglossia. Wright assevera que a sociedade latinófona manifestava antes um estado de “monolinguismo complexo” em que qualquer variedade escrita ou falada era concebida como parte da mesma língua, o latim, apesar de as modalidades exibirem diferenças estruturais marcantes. A situação monolíngue perduraria até a reforma carolíngia (séc. VIII-IX), ao norte dos Pirineus e na Catalunha, quando o latim dito “medieval”, com sua pronúncia ortográfica, foi inventado por reformadores eclesiásticos visando a padronizar a liturgia sob a direção do monge anglo-saxão, Alcuíno de York, que introduziu tal modelo da latinidade insular. Esse latim reformado foi introduzido no resto da Península Ibérica na reforma gregoriana (séc. XI-XII), explicando a demora histórica de quase trezentos anos entre o surgimento dos primeiros textos propriamente românicos nos domínios carolíngios (séc. IX-X) e seu aparecimento no resto da România ibérica (séc. XII-XIII).

Wright atribui a pouca aceitação da ideia de que o latim oral popular ou “protorromance” e o latim escrito eram simplesmente duas variedades da mesma língua principalmente à maneira em que relações grafo-fonêmicas desse período têm sido analisadas. Nisso, Wright concorda em certos pontos com a visão diglósica de Lüdtke, por exemplo (LÜDTKE 1964, p. 14-15). Ambos os filólogos duvidam que duas fonologias latinas tenham coexistido, uma especialmente arcaica e outra evoluída, enfatizando não haver como transmitir uma pronúncia ortográfica por muitas gerações, sem que o ensino da mesma deixasse algum vestígio nas obras dos gramáticos. “*Archaic morphology is accessible to the literate from written works. Archaic phonology is not similarly accessible; it cannot be*

taken from books; otherwise historical phonology would be simple". (WRIGHT, 1982, p. 43)

Por conseguinte,

The tradition of reading Latin aloud as an artificial language, a sound for each letter, in the Romance-speaking world as everywhere else, has the air of being obvious and as though it had been forever present. But someone, somewhere, had to establish that as a standardized norm, if Latin is to maintain its monolingual status, for it could not arise in a native Romance community. There was a kind of continuity through the years between Carolingian and Imperial Latin in the vocabulary and syntax of the educated, for these could always be resurrected from Classical books by antiquarians (FONTAINE 1981:786), but what we now think of as traditional Latin pronunciation had no such direct continuity with that of the Empire (Cp. LÜDTKE 1988, p. 63, on [-m], for example). That is why the invention of the need for what we now call Latinate pronunciation (with the sounds determined by the spelling) is the key issue here. As the historian Hobsbawm (1983, p. 1) pointed out: "traditions which appear or claim to be old are often quite recent in origin and sometimes invented. (WRIGHT, 1994a, p. 27-28)

Consequentemente, Wright propõe que, tanto o conservadorismo cultural, quanto as normas representacionais, perpetuava (e continua a perpetuar) uma aparência uniforme e, portanto, enganadora, da língua latina, em particular, quanto ao aspecto sonoro:

[R]eading aloud could have used Old Leonese phonetics even for such apparently Latinate material as a legal document, in the same way as readers of Modern English documents can read legal language aloud with their own phonetic habits. (WRIGHT, 1982, p. 167)

Para Wright, os filólogos não cogitaram tal possibilidade seriamente porque se acreditava que a ortografia latina mantinha sempre uma relação bastante direta e transparente entre as normas representacionais e a fonologia. Essa conjectura foi fundamentada na noção de uma pronúncia culta que tenha participado minimamente nas mudanças ocorridas no vernáculo. Por outro lado, Wright aponta as complexas correspondências grafo-fonêmicas das ortografias conservadoras do francês e inglês. Apesar de ambas as tradições escritas começarem com normas ortográficas que representavam a fonologia de forma razoavelmente direta, com o passar do tempo, mudanças no sistema sonoro não foram incorporados na grafia e, portanto, as correspondências grafo-fonêmicas se tornavam cada vez mais complexas. Hoje, existem desinências escritas francesas que não correspondem a nenhuma realização oral, por exemplo, *je jou-e, tu jou-es, il jou-e, elles jou-ent* = /ʒu:/ (Cf. WRIGHT, 1982, p. xi), enquanto outras flexões, distinguidas rigorosamente na escrita, são homófonas em muitas

variedades. Por exemplo: *-ais, -ait, -aient, -er, -ez, -é, -ée, -és, -ées, -ai* = [e] ~ [ɛ] (WRIGHT, 1982, p. 51). Na escrita inglesa, inversamente, a mesma sequência gráfica pode ser pronunciada de diversas maneiras. Por exemplo, *-ough*: (1) /au/ *bough* “galho”, *slough* “pântano”, *plough* “arado”; (2) /ʌf/ *tough* “duro”, *enough* “suficiente”, *rough* “áspero”; (3) /əʊ/ *although* “embora”, *dough* “massa”; (4) /ə/ *borough* “município”, *thorough* “rigoroso”.

Se a ortografia latina era como a francesa ou inglesa antes da reforma carolíngia, por conseguinte,

Learning to write involved the techniques for reproducing on parchment outdated inflectional morphology (such as -m, -bus, -abit) and a spelling system not closely suited to the evolved Romance. Seventh-century French [ˈvʲerɔʒə] or [ˈvirɔʒə], for example, might be spelled virgen, virginem, virgine or virgini, in the same way as Modern French [ʒa:t] can be spelled chante, chantes, chantent. [...] The old vocabulary, grammar and spelling could be picked up from books. Old pronunciation, however, cannot be picked up directly from books. (WRIGHT, 1982, p. x-xi)

O surgimento das escritas vernáculas seria um resultado da introdução da pronúncia ortográfica, que visava a padronizar o culto. Antes da reforma, no latim franco: *viridarium* = [ver'ɟer] “pomar”, *directum* = [drejt] “direto”. Depois da reforma, *viridarium* = [viridiarium], *directum* = [direktum] etc. (WRIGHT, 1982, p. 107). Além de os reformadores exigirem que a sintaxe e o léxico fossem mais próximos aos modelos bíblicos e patrísticos (e, portanto, mais distantes das falas quotidianas), tais mudanças de pronúncia teriam ressuscitado à oralidade muitas desinências morfológicas que haviam ficado mudas há gerações, embora fossem escritas. Desse modo, o novo estilo eclesiástico teria se tornado incompreensível para os analfabetos. Por conseguinte, os letrados sentiam a necessidade de desenvolver novas convenções de representação grafo-fonê-mica, muitas das quais eram decalcadas nas relações grafo-fonêmicas novamente diretas do uso eclesiástico, a fim de expressar os sons de itens lexicais que existiam apenas no vernáculo, porque as grafias tradicionais para esses lexe-mas haviam recebido novas pronúncias “reformadas”, ou seja, [ver'ɟer] = *vergier*, [drejt] = *dreit* etc. O intuito consciente desses inovadores ortográficos era de produzir textos com finalidades pastorais compreensíveis pelos analfabetos, já que os sermões prescritos eram redigidos no novo linguajar eclesiástico. Ao registrarem o uso do vernáculo na escrita, no entanto, destacavam-se as diversas diferenças entre as variedades populares e a linguagem culta eclesiástica, conduzindo lentamente à conceptualização de duas línguas diferentes.

Assim, a ideia de haver duas línguas diferentes (latim e romance) numa dada comunidade de fala românica, em lugar de dois estilos da mesma língua (latim), desenvolvia-se gradualmente a partir deste momento, estimulada pela falta de compreensão da variedade culta entre os não latinados. Consequentemente, no modelo wrightiano, a separação conceitual entre o latim e o romance surge *depois* e *em decorrência* da invenção das novas modalidades escritas, diferentemente da visão pregada nos modelos binormais diglóticos tradicionais, nas quais a consciência de haver dois níveis linguísticos nitidamente definidos estrutural e praticamente *antecede desde há séculos* o surgimento das primeiras escritas românicas.

Wright argumenta que, embora as correspondências grafo-fonêmicas da ortografia latina tivessem começado razoavelmente diretas e transparentes, após muitos séculos, essas relações teriam atingido um grau de complexidade parecido ao francês e inglês modernos (1982, p. 168). Wright exemplifica seus argumentos sobre as relações grafo-fonêmicas no latim tardio com transcrições hipotéticas da leitura em voz alta de textos escritos antes das reformas eclesiásticas conforme a tradicional norma ortográfica latina não reformada. Por exemplo,

1. In Dej nomine. Ego Splendonius tui Fredesinde In Domino salutem. Ideo placuit

en'die'nwem.ne'io.e.splen'do.ɲo'tie.fre'dzin.de.en'dwe.ɲo.sa'lu.ðe.i.ɲo'plo.go

No nome [de] Deus. Eu, Esplendônio, saúdo-te, Frezindo, no Senhor. Assim, plogo

2. mici atque conuenit nunlljusque cogentis Imperio neque suadentjs artjculo

'mi.ee.kom'bi.ne.nu.ɫjos.ke.ko'dʒjen.tes.em'pe.rjo.ni.swa'djen.tes.ar'te.ɫo.

-me e conveio [-me], e sem a coação, nem a persuasão de ninguém,

3. set probria mici acesi uoluntas ut uinderem tjui Iam dicte Fredesinde terra

se'pro.βri.ja'mi.e.a'ʃe.ze.vo.lun'ta.de.o.ven'dje.re'ti.e.ja'dij.te.fre'dzin.de'tje.ra.

mas conforme a minha livre vontade, que [eu] vendasse-te, já dito Frezindo [um] terreno

4. In uilla Uiasco super Illa senrra domniga lloco predicto Agro rrodundo...

en'vi.la'vjas.ko'so.βre.la'ser.na'ðo.ɲi.ɲa'lwe.ɲo.pre'dij.to'a.ɲro.ro'don.do...

em Vila Viasco, sobre a serna donhiga [no] lugar [já] mencionado: Agro Redondo ...

Transcrição hipotética da leitura

de um contrato de compra-e-venda leonês do ano 946 d.C.

Floriano Cumbreño (1951, vol. II, p. 361-362, apud WRIGHT, 1982, p. 166-167)

O filólogo britânico defende ainda que a leitura em voz alta de um texto escrito de tal latim tardio não reformado seria compreensível para um público analfabeto, embora certas passagens soassem arcaicas e rebuscadas, devido à linguagem pouco corrente na fala contemporânea. O público teria aprendido o significado de tais expressões pelo contato fre-

quente com o estilo escrito no culto e em trâmites burocráticos, por exemplo, testemunhar contratos, testamentos e doações a instituições religiosas etc.

3. *Léxico arcaico e a leitura logográfica*

Uma das críticas das teorias de Wright envolve a questão do léxico arcaico presente nos textos tabelionais alto-medievais. Vários filólogos, como Dworkin (1995), Emiliano (1991, p. 242, 1993, p. 239) e Walsh (1991), disputam que as discrepâncias entre o vocabulário tradicional utilizado nos textos e os lexemas vernaculares correspondentes contemporâneos não criariam “consciência subliminar” (WALSH, 1991, p. 207-278), de haver duas línguas.

Ahora, si la lectura en voz alta con fonética vernácula es necesaria para la inteligibilidad de un texto, no es sin embargo suficiente. Si un anglófono puede leer y comprender un texto legal en inglés moderno eso no depende sólo de la fonética empleada. Del mismo modo, un documento notarial inglés del siglo X u XI no es comprensible para un anglófono moderno que no haya estudiado el inglés antiguo, aunque sea leído con fonética moderna (o hasta con morfología moderna). [...] El problema es que la oralización de un texto en latín no reformado por medio de simples reglas de correspondencia grafo-fonémica conduce a resultados ininteligibles desde el punto de vista de la lengua hablada. (EMILIANO, 1993, p. 239)

A admissão não qualificada das *correspondências* de Wright, ou seja, a oralização dos textos a partir de simples regras de conversão grafo-fonêmica levaria, como já tive ocasião de fazer notar (EMILIANO, 1993 e 1994), à produção de seqüências sem significado contemporâneo virtualmente incompreensíveis para os ouvintes romanófonos iletrados. (EMILIANO, 2003, p. 23)

Walsh (1991) protesta que ler *Ingrediamur inquit domum*, uma passagem das glosas emilianenses típica das leituras litúrgicas que os analfabetos experimentavam regularmente (GARCIA LARREGUETA, 1982, p. 100, *apud* WALSH, 1991, p. 206), com uma pronúncia contemporânea, ou seja, algo como *[iŋ.gre'ðja.mo'riŋ.kið'do.mo], não tornaria o texto compreensível para alguém cuja expressão espontânea do mesmo sentido era *[en'tre.mos'di.fo.e.ne.la'ka.za] (Cf. espanhol *entremos, dijo, en la casa*; português *entremos, disse, na casa*).

Apesar de criticar Wright no que diz respeito à linguagem arcaica, Emiliano é favorável ao monolingüismo tardo-latino e à explicação dos fatores que provocaram a separação conceitual. Desse modo, o filólogo

português defende que o latim tardio tabelional envolvia dois tipos de representação escrita (EMILIANO, 1991). A maioria das formas tradicionais representariam a fonologia evoluída mediante uma ortografia conservadora, tal como em inglês ou francês modernos. Por exemplo: *fecerit* = [fi'zer] e *mandauerit* = [man'dar]. Tais formas ortográficas manteriam vínculos tanto com a fonologia quanto com a semântica do item lexical que representavam, porque seu descendente direto ainda vigorava no vernáculo. Entretanto, Emiliano também afirma que certas outras formas escritas tradicionais representariam vocábulos contemporâneos apenas semanticamente. Por exemplo: *percusserit* = [fe'rir], *occiderit* = [ma'tar], *uoluerit* = [ki'zer]. Isto é, as sequências de letras *p.e.r.c.u.s.s.-*, *o.c.c.i.d.-*, *u.o.l.u.-* representavam {fer-}, {mat-} e {kiz-}, respectivamente, e *-erit* expressava {-'ir}, {-'ar}, {-'er}, ou seja, “terceira pessoa do singular do futuro do subjuntivo”. Nessa segunda classe, as formas escritas, eram logogramas que representavam lexemas vernaculares por “correspondências grafo-lexêmicas”¹⁵.

A teoria logográfica de Emiliano talvez provoque estranhamento. Contudo, semelhantes práticas não são incomuns entre os sistemas de escrita no mundo. Caracteres chineses, por exemplo, não explicitam a composição fonológica dos vocábulos que representam. Assim, para Emiliano, nos logogramas, as letras se tornariam meros componentes gráficos, desprovidos de valor sonoro no momento de ler.

Emiliano (1991, 1993, 1994, 1999, 2003) e (1991a, 1991b, 1992, 1995, 1998) propõe uma teoria de logografia “forte” em que a escrita e a leitura são logográficas porque algumas formas escritas são caracteres verdadeiramente plerêmicos. Formas ortográficas tradicionais, cujos descendentes persistem no vernáculo contemporâneo, funcionam mediante a representação grafo-fonêmica por caracteres cenêmicos. O aparente léxico arcaico é, na realidade, nada mais que a tradicional grafia conservadora. Por outro lado, vários outros filólogos, como Penny (1998, 2003), Pensado (1991, 1998), Varvaro (1991) e Wright (1994, 1996) recorreram a uma hipótese logográfica “fraca” para explicar como um sistema de escrita tradicional e as prescrições de um estilo arcaizante poderiam representar um vernáculo evoluído, embora sem admitir estratégias de conversão grafo-

¹⁵ Emiliano baseia sua hipótese logográfica nas alternâncias sincrônicas que encontrou entre as rubricas e os parágrafos do mesmo texto. Os documentos investigados pertencem a uma família de forros/costumes da Estremadura leonesa escritos nos séculos XII e XIII, os forais de Riba-de-Coa. (EMILIANO, 1991, p. 235; CINTRA, 1959)

lexêmica.

Esse segundo grupo defende seu uso do termo “logografia”, apontando que a unidade de interpretação básica em seu modelo é a palavra como unidade visual e gráfica. Parte-se do fato de leitores experientes modernos não reconhecerem as palavras pelo processamento mecânico e linear de símbolos visuais discretos. Antes, percebem-nas imediata e *globalmente*, como um todo visual que associam a certo significado. Por conseguinte, formas ortográficas são apreendidas e reconhecidas com pouquíssima análise sistemática das correspondências grafo-fonêmicas abaixo do nível da palavra por meio de acesso léxico direto, sem que uma etapa de conversão fonográfica intermedeie no acesso léxico.

4. Sociofilologia, logografia e psicofilologia

A sociofilologia levanta a possibilidade de haver existido logografia na alta idade média e tal hipótese não pode ser descartada facilmente, porque sabemos que tais práticas, de fato, existem em outras sociedades. Na teoria, não é difícil imaginarmos um sistema de escrita de origem alfabética em que alguns conjuntos de grafemas básicos funcionariam como logogramas (SAMPSON, 1985, p. 203). Se a leitura do latim tardio fosse realmente logográfica, pelo menos quanto aos itens lexicais exigidos pela tradição literária, sem descendentes diretos na linguagem coloquial contemporânea, não haveria problemas de compreensão entre os analfabetos.

Defronta-nos agora o problema de determinar se a leitura podia ser logográfica. Procurei solução numa fonte de dados empíricos confiáveis: a fisiologia do olho e da vista e (de uma maneira mais tentativa) as descobertas da psicologia experimental da leitura. Portanto, optei por denominar essa abordagem “psicofilologia”, pois, a psicolinguística e a psicologia esclarecem questões filológicas, de forma parecida à em que a sociofilologia wrightiana tem expandido o conhecimento filológico, com a aplicação das descobertas da sociolinguística.

Baseamo-nos inicialmente no trabalho de Saenger sobre o surgimento da “leitura referencial”, típica da maioria das escritas alfabéticas modernas (a rapidíssima perscrutação mental e silenciosa) (SAENGER, 1982, 1989, 1990a, 1990b, 1991, 1994, 1997a, 1997b). Segundo o bibliotecário pesquisador americano, a evidência experimental demonstra que essa leitura depende da capacidade de perceber palavras escritas global-

mente e captar seu significado diretamente, da mesma maneira como caracteres chineses ou japoneses, ou seja, como logogramas. Isto é, na leitura referencial, o acesso lexical é direto: as seqüências de símbolos não são analisadas primeiro para extrair as unidades fonológicas que elas representam, para depois identificar os itens lexicais pelas cadeias fonológicas que os compõem (acesso lexical intermediado).

Segundo Saenger, o estímulo visual primário nessa leitura global advém principalmente dos *contornos* das palavras escritas. Para que leitores possam discriminar rapidamente e sem ambiguidade a posição de palavras e suas características visuais, é preciso que espaços separem cada palavra escrita e que esses espaços sejam de 1,5 a 2,0 vezes maiores do que as demais unidades de espaço presentes, por exemplo, no traçar das letras e nos espaços que separem as letras dentro de uma seqüência gráfica. Essa proporção é relevante porque, espaços menores de 1,5:1 não são perceptíveis fora da visão foveal e, conseqüentemente, não possibilitam nem a localização rápida de palavras escritas, nem a percepção dos seus contornos.

A razão 1,5-2,0:1 é determinada pela estrutura física do olho humano: a relação entre a densidade de células cone e a acuidade visual (vide a seção seguinte). É por isso que a abordagem psicofilológica é confiável. Podemos ter certeza de que o olho não evoluiu durante os últimos 3-4.000 anos em que se escreve, e ainda mais confiança disso quanto aos 1.500 anos desde a produção de textos em latim tardio.

5. A fisiologia e psicologia da leitura referencial moderna

A luz refletida por um objeto passa pela lente da pupila e recai na retina, onde há dois tipos de células: cones e bastonetes. As células cone se concentram no centro da retina e sua frequência diminui rapidamente, conforme se distancie dessa área. Portanto, a acuidade visual é máxima na fóvea (2-3° ao redor do ponto de fixação). Só percebemos todos os detalhes de uma letra na vista foveal. Na *parafóvea* (a 3-6° de excentricidade), apenas distinguem-se aspectos visuais mais gerais (o comprimento, o contorno) porque a falta de precisão visual faz com que a interferência lateral mútua torne nebulosas as características das letras. Na vista *periférica* (a 7°+), somente os grandes espaços intervocabulares (quando acima de 1,5:1) se destacam como interrupções nas linhas de símbolos gráficos indistintos.

A saliência visual dos espaços intervocabulares no formato convencional (1,5-2,0:1) na vista não foveal permite o reconhecimento global de palavras escritas, até antes da vista foveal recair nelas. Poder localizar espaços intervocabulares é fundamental para planejar os movimentos oculares (“pulos” (saques) de aproximadamente 30 milissegundos e períodos de descanso (fixações) que perduram 200 a 300 milissegundos). Podendo perceber espaços intervocabulares na vista periférica, posicionamos a fôvea sobre a parte mais informativa (aproximadamente no meio da primeira metade de uma palavra de tamanho médio). O resto, que é menos informativo, é processado na vista parafoveal, junto com as silhuetas das próximas palavras.

A informação lexical extraída numa fixação ocular é integrada ao contexto dinâmico formado do que já foi visto/lido e é armazenada na memória de curto prazo. Acrescentam-se informações extraídas da parafovea sobre o eventual contexto vindouro para formar previsões, restringidas pela coerência, sobre o que puder aparecer no texto ainda a ser lido. Assim podemos desambiguar as palavras quase até à individualidade, antes mesmo de a vista foveal chegar nelas. Consequentemente, mal é necessário fixá-las com a vista foveal para confirmar a hipótese já calculada a partir desta “previsão parafoveal”. Simultaneamente, a vista periférica localiza os espaços intervocabulares distantes e programa “pistas de pouso” mais propícias para futuros saques. Escolhe-se tipicamente um ponto logo à direita de um espaço que é seguido por uma palavra mais comprida, por conter provavelmente informação lexical.

6. A separação canônica e a leitura de *scriptio continua*

Segundo Saenger, essa extraordinária eficiência, tanto da extração de informação lexical (a base de qualquer leitura de qualquer escrita), quanto de movimento ocular, se deve à *separação canônica* (espaços de tamanho significativamente maiores do que os demais são inseridos entre cada vocábulo). Tal convenção faz com que os espaços separadores se destaquem visualmente ao máximo, até a vista periférica. No entanto, a escrita alfabética latina não era sempre assim. As primeiras escritas romanas separavam as palavras com pontos. Tais símbolos não permitem a percepção dos contornos das palavras porque a quantidade de espaço usada para traçá-los não é distintiva relativamente aos demais caracteres. Durante o séc. II d.C., os romanos passaram a escrever em *scriptio continua*, como

os gregos. Essa convenção não indica as divisas lexicais. Às vezes, pequenos espaços de tamanhos diferentes, eventualmente complementados por pontos de diferentes tipos, serviam para pontuar unidades de sentido.

Os trabalhos de Saenger demonstram que a separação canônica somente se generalizaria na Europa Ocidental a partir do século X-XI, inicialmente sob a influência de manuscritos irlandeses, anglo-saxônicos e ibéricos (esses derivados das normas de separação do árabe). Entretanto, não se sentiam os benefícios da convenção de uma maneira mais geral antes do século XIV-XV, quando a leitura se tornava mais difundida entre os leigos europeus.

Na Antiguidade ou na Alta Idade Média, ninguém podia abrir um rolo ou códice desconhecido e começar a ler fluentemente, como se pode hoje. Era necessário estudar o texto de perto, seguindo um procedimento chamado as *officia grammaticae* que progredia por etapas: (1) *praelectio*, a identificação das palavras no texto a partir de recodificação grafo-fonêmica em que se juntam letras e sílabas. Especialmente no contexto escolar, esta fase envolvia a marcação das palavras (*discretio*) com símbolos gráficos, como o *hifen* e o *diástole*; (2) *lectio*, a identificação das unidades de sentido: frases (*commata*), cláusulas (*cola*), sentenças (*sententia/ periodus*) e parágrafos (*capitula*); e, finalmente, (3) *pronuntiatio*, a leitura em voz alta.

A necessidade de seguir as *officia grammaticae* decorria do fato de que ler textos em *scriptio continua* ou em escrita “arejada” (com espaços inseridos irregularmente entre sílabas e morfemas) é muito mais penoso do que a leitura referencial moderna. Sem espaçamento regular e icônico, não existem “pistas de pouso” para os saques, que passam a transitar apenas 2 a 3 caracteres, reduzindo drasticamente a quantidade de informação extraída por fixação: apenas algumas letras ou, no máximo, algumas sílabas. Tampouco há como perceber os contornos das palavras porque sequer há “palavras” como unidades visuais. A pouca acuidade parafoveal faz com que a sobreposição mútua encubra as demais características. Portanto, o leitor não pode aproveitar a previsão parafoveal.

O leitor de um texto sem separação canônica avança lenta, sistemática e linearmente, procurando sequências de letras inadmissíveis conforme as normas ortográficas, as quais indicariam uma eventual divisa vocabular. Após realizar esse *discretio*, o leitor não mantém uma sequência de palavras ou até frases na memória (como os leitores modernos fariam),

mas antes uma cadeia de letras da qual é preciso extrair primeiro os fonemas representados, para depois juntá-los em seqüências de sílabas, as quais são estudadas por sua vez para discernir as palavras que constituem (*praelectio*). Consequentemente, o leitor precisa reconhecer as correspondências fonográficas antes de poder acessar o nível lexical. Ainda que o leitor consiga identificar as palavras corretamente, sua capacidade de compreender seu significado numa frase coerente está bastante prejudicada pelo acúmulo de pressões dos processos cognitivos sobre sua atenção. Tipicamente, portanto, é necessário pausar na decodificação de *scriptio continua* de modo a processar o conteúdo de uma passagem (*lectio*).

7. Conclusões

Conclui-se, a partir da análise acima, que não se pode afirmar que a leitura logográfica era praticada na România alto-medieval, sem sabermos qual o tipo de separação de palavras naqueles tempos. Se não houvesse espaços visualmente icônicos entre cada palavra escrita, não haveria grafias que pudessem funcionar facilmente como logogramas, porque as palavras escritas estariam essencialmente invisíveis fora da vista foveal. A localização e interpretação de itens lexicais, sejam como forem as relações grafo-fonêmicas, procederia forçosamente de maneira composicional, conforme as *officia grammaticae*.

No entanto, mesmo que uma forma ortográfica fosse reconhecida de forma composicional, isso não impossibilitaria que a mesma funcionasse como um logograma. Sempre existe a remota possibilidade de reconhecer *percusserit*, de maneira composicional no *praelectio* (*p.e.r+c.u.s.s+e.r.i.t*) e, a seguir, associá-la com a palavra “ferir”, no *lectio*. No entanto, sem separação canônica, tal procedimento não seria realizado por meio de captação visual imediata e global.

Resta descobrir, porém, a data em que a separação de palavras por espaços atingiu a proporção de 1,5-2,0 vezes as demais unidades de espaço. Esse foi o tópico da investigação que eu realizei para o doutoramento e que espero tratar em maior detalhe em outro artigo futuramente. Resumidamente, após examinar 104 textos da Península Ibérica escritos entre o século VIII e o século XIII, concluí que os textos do *corpus* produzidos antes do século XI tipicamente não apresentam convenções de separação de palavras que permitiriam sua identificação na parafóvea e, portanto, apenas após isso a “logografia” por captação global poderia começar a funcionar. Antes desse período, as restrições físicas e mentais impostas

pelas condições visuais impróprias que vigoravam nos textos teriam complicado o funcionamento desse tipo de leitura além do ponto de ele se tornar impraticável.

Deste modo, espero ter demonstrado como informações e modelos teóricos oriundos de outras disciplinas científicas podem se configurar novamente para enriquecer nosso conhecimento de alguns fenômenos ligados a questões filológicas.

REFERÊNCIAS BIBLIOGRÁFICAS

BLAKE, R. J. Syntactic Aspects of Latinized Texts of the Early Middle Ages. In: WRIGHT, R. (Org.). *Latin and the Romance Languages in the Early Middle Ages*, 1991a, p. 219-232.

_____. Squeezing the Spanish Turnip Dry: Latinized Documents from the Early Middle Ages. In: CRAVENS, T. D.; HARRIS-NORTHALL, R. (Orgs.). *Linguistic Studies in Medieval Spanish*. Madison, WI: Hispanic Seminary for Medieval Studies, 1991b, p. 1-14.

_____. Aspectos sintácticos del español antiguo: la prosa latinizada del *Cartulario de San Millán de la Cogolla*. In: VIGUERA, M. Ariza; AGUILAR, R. Cano; MENDOZA, J. M.; NARBONA, A. (Orgs.). *Actas del II Congreso de Historia de la Lengua Española* (Sevilla, 1990). Madrid: Pabellón de España, 1992, vol. 2, p. 291-305.

_____. El latín notarial de un escriba bilingüe o “bígrafo” del XIII. In: PÉREZ GONZÁLEZ, M. (Org.), *Actas del I Congreso Nacional de Latín Medieval*, 1995, vol. 3, p. 463-468.

_____. Las glosas de San Millán y de Silos en su contexto sociolingüístico. In: TURZA, C. García *et al.* (Eds.). *Actas del IV Congreso Internacional de la Historia de la Lengua Española*, 1998, p. 925-932.

CINTRA, Luis Felipe Lindley. *A linguagem dos foros de Castelo Rodrigo*. Lisboa: Publicações do Centro de Estudos Filológicos, 1959.

DWORKIN, Stephen N. Latín tardío y romance temprano: implicaciones léxicas de una hipótesis controvertida. In: PÉREZ GONZÁLEZ, M. (Org.). *Actas del I Congreso Nacional de Latín Medieval*, 1995, vol. 3, p. 489-494.

EMILIANO, António H. de Albuquerque. Latin or Romance? Graphemic Variation and Scripto-Linguistic Change in Medieval Spain. In:

WRIGHT, R. (Ed.). *Latin and the Romance Languages in the Early Middle Ages*, 1991, p. 233-247.

_____. Latín y romance y las glosas de San Millán y de Silos: apuntes para un planteamiento grafémico. In: PENNY, R.; DEYERMOND, A.; HITCHCOCK, R. (Orgs.). *Actas del I Congreso Anglo-Hispano*. Madrid: Junta de Andalucía/Castalia. 1993, vol. 1, n. 3 (Linguística), p. 235-244.

_____. Tradicionalidad y exigencias de realismo en la lengua notarial hispánica (hasta el siglo XIII). In: PÉREZ GONZÁLEZ, M. (Org.). *Actas del I Congreso Nacional de Latín Medieval*, 1995, vol. 3, p. 511-518.

_____. O mais antigo documento latino-português (882 A.D.) - edição e estudo grafémico. *Verba*, n. 26, p. 7-42, 1999.

_____. *Latim e romance em documentação notarial da segunda metade do século XI*. Análise scripto-lingüística de documentos notariais do Liber Fidei de Braga de 1050 a 1110. 2 vol. Coimbra: Fundação Calouste Gulbenkian/Fundação para a Ciência e Tecnologia, 2003.

FLORIANO CUMBREÑO, António Cristino. *Diplomatica española del período astur*. 2 vol. Oviedo: Imprenta "La Cruz"; Universidad de Oviedo, 1951.

GARCÍA LARRAGUETA, S. *Las glosas emilianenses*: edición y estudio. Logroño: Instituto de Estudios Riojanos, 1984.

LÜDTKE, Heinrich. Die Entstehung romanischer Schriftsprachen. *Vox Romanica*, n. 23, p. 3-21, 1964.

PENNY, Ralph. La grafía de los textos notariales castellanos de la alta edad media: ¿sistema logográfico o fonológico?. In: BLECUA, J. M. et al. (Eds.). *Estudios de grafemática en el dominio hispánico*, 1998, p. 211-223.

_____. Ambigüedad grafemática: correspondencia entre fonemas y grafemas en los textos peninsulares anteriores al S. XIII. In: PERDIGUERO VILLARREAL, H. (Org.). *Lengua romance en textos latinos de la edad media*, 2003, p. 221-228.

PENSADO, Carmen. How was Leonese Vulgar Latin Read?. In: WRIGHT, R. (Org.). *Latin and the Romance Languages in the Early Middle Ages*, 1991, p. 177-189.

_____. Sobre los límites de la mala ortografía en romance. ¿Por qué el inglés *fish* no se escribe *ghoti* después de todo? In: BLECUA, J. M. et al. (Orgs.). *Estudios de grafemática en el dominio hispánico*, 1998, p. 225-

241.

SAMPSON, Geoffrey. *Writing Systems: A Linguistic Introduction*. London: Hutchinson & Co.; Stanford, CA: Stanford University Press, 1985.

SAENGER, Paul Henry. Silent reading: its impact on late medieval script and society. *Viator*, n. 13, p. 367-414, 1982.

_____. Books of hours and the reading habits of the later Middle Ages. In: CHARTIER, R. (Org.). *The Culture of Print: Power and Uses of print in Early Modern Europe*. Princeton, NJ: Polity Press, 1989, p. 143-173.

_____. La naissance de la coupure et de la séparation des mots. In: MARTIN, H.-J.; VEZIN, J. (Orgs.). *Mise en page et mise en texte du livre manuscrit*. Paris: Éditions du Cercle, 1990a, p. 447-449.

_____. Coupure et separation des mots sur le Continent au Moyen Âge. In: MARTIN, H.-J.; VEZIN, J. (Orgs.). *Mise en page et mise en texte du livre manuscrit*. Paris: Éditions du Cercle, 1990b, p. 451-455.

_____. The separation of words and the psychology of reading. In: OLSON, D. R.; TORRENCE, N. (Orgs.). *Literacy and Orality*. Cambridge: Cambridge University Press, 1991, p. 198-214.

_____. Word separation and its implications for manuscript production. In: RÜCK, P. (Org.). *Die Rationalisierung der Buchstellung im Mittelalter und in der frühen Neuzeit. Ergebnisse eines buchgeschichtlichen Seminars* (Wolfenbüttel, 1990). Marburg an der Lahn: Institut für Hilfswissenschaften, 1994, p. 41-50.

_____. *Space Between Words. The Origins of Silent Reading*. Stanford, CA: Stanford University Press, 1997a.

_____. Separated script at Rheims and Fleury at the time of Gerbert and Abbo. In: BARBIER, F.; PARENT-CHARON, A.; DESROUSSILES, F. Dupuisgrenet; JOLLY, C.; VARRY, D. (Eds.). *Le livre et l'historien. Études offertes à l'honneur du professeur Henri-Jean Martin*. Geneva: Droz, 1997b, p. 3-23.

VARVARO, Alberto. Latin and Romance: Fragmentation or Restructuring? In: WRIGHT, R. (Ed.). *Latin and the Romance Languages in the Early Middle Ages*, 1991, p. 44-51.

WALSH, Thomas J. Spelling Lapses in Early Medieval Latin Documents and the Reconstruction of Primitive Romance Phonology. In: WRIGHT, R. (ed.). *Latin and the Romance Languages in the Early Middle Ages*,

1991, p. 205-218.

WRIGHT, Roger. *Late Latin and Early Romance in Spain and Carolingian France*. Liverpool: F. Cairns, 1982.

_____. Logographic Script and Assumptions of Literacy in Tenth-Century Spain. In: PARRY, M. M.; DAVIES, W. V.; TEMPLE, R. A. M. (Orgs.). *The Changing Voices of Europe. Papers in Honour of Prof. Glanville Price*. Cardiff: Cardiff University Press, 1994a, p. 125-135.

_____. *Early Ibero-romance: Twenty-One Studies on Language and Texts from the Iberian Peninsula between the Roman Empire and the Thirteenth Century*. Newark, DEL: Juan de la Cuesta, 1994b.

_____. *A Sociophilological Study of Late Latin*. Turnhout: Brepols, 2002.