Ensaio

BIOLEXICOGRAFIA E PARADIGMA LEXICOGRÁFICO DISTRIBUÍDO: ENSAIO SOBRE LÉXICO DIGITAL, INTELIGÊNCIA COLETIVA E SOBERANIA DA LINGUAGEM

Francisco André da Silva[1]

[1] Universidade Estadual do Ceará (UECE) / Programa de Pós-Graduação em Linguística Aplicada (PosLA), Fortaleza - Ceará / Brasil.

Resumo

Este ensaio propõe a biolexicografia distribuída como reflexão teórico-propositiva sobre os limites da dicionarização centralizada diante da fluidez lexical em ambientes digitais. Fundamentado na Linguística Aplicada como ciência social, o texto articula lexicografia, inteligência coletiva, estigmergia e consenso distribuído para pensar novas formas de registro do léxico. Defende-se que o dicionário pode deixar de operar como arquivo estático para constituir-se como sistema vivo, responsivo e atravessado pela agência dos falantes. Conclui-se que a biolexicografia oferece uma via crítica para pensar soberania lexical e resistência ao apagamento linguístico.

Palavras-chave: Biolexicografia. Lexicografia digital. Linguística Aplicada. Inteligência coletiva. Léxico digital.

Biolexicography and the distributed lexicographical paradigm: an essay on digital lexicon, collective intelligence and language sovereignty

This essay proposes distributed biolexicography as a theoretical and propositional reflection on the limits of centralized dictionary-making in the face of lexical fluidity in digital environments. Grounded in Applied Linguistics as a social science, the text articulates lexicography, collective intelligence, stigmergy and distributed consensus to rethink lexical recording. It argues that dictionaries may cease to function as static archives and become living systems shaped by speakers' agency. The essay concludes that biolexicography offers a critical path to lexical sovereignty and resistance to linguistic erasure.

Keywords: Biolexicography. Digital lexicography. Applied Linguistics. Collective intelligence. Digital lexicon.

Biolexicografía y paradigma lexicográfico distribuido: ensayo sobre léxico digital, inteligencia colectiva y soberanía del lenguaje

Este ensayo propone la biolexicografía distribuida como reflexión teórico-propositiva sobre los límites de la diccionarización centralizada frente a la fluidez léxica en ambientes digitales. Fundamentado en la Lingüística Aplicada como ciencia social, el texto articula lexicografía, inteligencia colectiva, estigmergia y consenso distribuido para repensar el registro léxico. Se sostiene que el diccionario puede dejar de operar como archivo estático para constituirse como sistema vivo, atravesado por la agencia de los hablantes. Se concluye que la biolexicografía ofrece una vía crítica para pensar soberanía lexical y resistencia al borramiento lingüístico.

Palabras clave: Biolexicografía. Lexicografía digital. Lingüística Aplicada. Inteligencia colectiva. Léxico digital.

DOI: https://doi.org/ 10.5281/zenodo.23892501
ISSN: 2359-6252
Editor-chefe: Vinicius de Souza Moreira
Editora-adjunta: Letícia Lima Milani Rodrigues
Artigo submetido em 13 de junho de 2026 e aceito para publicação em 11 de setembro de 2026

A velocidade com que novas palavras, gírias, ressignificações e disputas semânticas emergem nas redes digitais impõe um desafio decisivo à lexicografia contemporânea. O dicionário, historicamente concebido como instrumento de estabilização e legitimação do léxico, vê-se confrontado por uma língua que circula em ritmo acelerado, atravessada por memes, regionalismos, tecnoletos, disputas políticas e formas de pertencimento que frequentemente escapam aos tempos lentos da curadoria editorial. Nesse cenário, a questão não é apenas técnica, mas epistemológica e política: quem tem o poder de registrar uma palavra? Quem decide quando um uso se torna legítimo? Quais vozes permanecem fora do arquivo lexical?

Este ensaio propõe a noção de biolexicografia distribuída como reflexão teórico-propositiva sobre os limites dos modelos centralizados de dicionarização diante da fluidez da linguagem digital. Trata-se de pensar o léxico não como inventário morto de unidades estabilizadas, mas como sistema complexo adaptativo, vivo, mutável e atravessado por relações sociais.

A proposta ancora-se na Linguística Aplicada compreendida como ciência social, conforme Moita Lopes (1996), isto é, como campo comprometido com problemas reais mediados pela linguagem. Nesse sentido, a dicionarização deixa de ser apenas prática técnica de descrição lexical e passa a ser compreendida como espaço de disputa entre autoridade institucional, uso social e soberania dos falantes.

A lexicografia tradicional, ao organizar verbetes sob critérios normativos e editoriais, desempenhou papel importante na sistematização do saber linguístico. Contudo, seu funcionamento centralizado também pode produzir apagamentos. Ao depender de conselhos editoriais, tempos longos de atualização e filtros de legitimidade associados à norma culta, o dicionário tende a reconhecer tardiamente usos que já circulam de modo produtivo nas práticas sociais. Biderman (1984) compreende o dicionário como instrumento de organização do universo lexical de uma comunidade; porém, quando essa organização se submete exclusivamente à autoridade institucional, determinados modos de dizer podem permanecer em condição de invisibilidade.

Termos como *flopar, macetar, promptar*, cancelamento e lugar de fala mostram que a vitalidade lexical frequentemente antecede sua consagração oficial. Antes de entrarem em repertórios formais, essas palavras já organizam afetos, conflitos, identidades, práticas profissionais e disputas ideológicas.

A crítica aqui proposta não pretende negar a importância da lexicografia, mas deslocar seu horizonte. Em vez de tratar o dicionário como tribunal que concede existência legítima ao signo, propõe-se imaginá-lo como infraestrutura responsiva de acompanhamento da vida social da palavra. Essa mudança exige compreender a língua como realidade dinâmica. Bakhtin e Volóchinov (2006) apontam que a palavra é território de disputa ideológica, não sinal neutro e fixo. Do mesmo modo, a perspectiva dos sistemas complexos permite compreender que a língua se auto-organiza por meio das interações sociais, respondendo às pressões históricas, tecnológicas e culturais de cada contexto (Larsen-Freeman, 2012).

É nesse ponto que a biolexicografia se apresenta como possibilidade conceitual. O prefixo *bio* não é usado aqui em sentido biologizante, mas como metáfora epistemológica para indicar uma lexicografia viva, processual e responsiva. Diferentemente do verbete estático, que tende a oferecer uma definição estabilizada, a biolexicografia buscaria registrar trajetórias de sentido, camadas de uso, disputas interpretativas e variações pragmáticas. Um termo como cancelamento, por exemplo, não comporta apenas uma definição. Ele pode funcionar como denúncia pública, mecanismo de responsabilização, prática de linchamento virtual, estratégia de visibilidade ou dispositivo de censura, dependendo da comunidade discursiva e da posição ideológica que o mobiliza. Registrar essa palavra exige, portanto, mais do que fixar um significado; exige mapear sua polifonia.

A inteligência coletiva oferece um caminho para pensar esse deslocamento. Em ambientes digitais, os falantes não apenas usam palavras: eles deixam rastros, replicam sentidos, alteram contextos, disputam interpretações e fazem circular novas formas lexicais. Lévy (1997) compreende a inteligência coletiva como forma de produção distribuída de conhecimento. Transposta para a reflexão lexicográfica, essa noção permite pensar que a validação de um uso não precisa depender exclusivamente de um centro editorial. Ela pode emergir da circulação social do termo, de sua

produtividade pragmática, de sua capacidade de atravessar comunidades e de sua estabilidade relativa em contextos distintos.

A estigmergia, conceito oriundo dos estudos sobre coordenação descentralizada em sistemas coletivos, também contribui para essa reflexão. Em sistemas estigmérgicos, os agentes coordenam suas ações por meio de rastros deixados no ambiente (Grassé, 1959). No caso da linguagem digital, cada postagem, comentário, réplica ou reaproveitamento lexical funciona como marca de circulação. A palavra deixa rastros de uso. Esses rastros, quando recorrentes e socialmente produtivos, indicam que determinado signo passou a exercer função comunicativa reconhecível. Assim, a biolexicografia distribuída poderia operar como modelo de escuta desses rastros, sem reduzir a língua a estatísticas brutas ou a decisões verticais.

Nesse horizonte, o protocolo *Proof of Meaning*, aqui entendido apenas como hipótese conceitual, serviria para nomear um princípio de validação distribuída do sentido. Diferentemente de modelos técnicos baseados em prova computacional ou capital acumulado, a prova de significado estaria vinculada ao uso social, à dispersão discursiva e à estabilidade dialógica de uma unidade lexical. Um termo não seria reconhecido apenas porque aparece muitas vezes, mas porque circula em contextos diversos, produz compreensão entre falantes e revela capacidade de organizar práticas sociais.

É imperativo, contudo, demarcar a diferença qualitativa dessa ecologia em relação a iniciativas contemporâneas que já ensaiam o diálogo entre léxico e redes, como a Linguística de *Corpus* Digital e a Dicionarização Aberta (a exemplo do projeto *Wiktionary*). Embora a Linguística de *Corpus* forneça instrumentais robustos para o mapeamento quantitativo e a observação de dados em larga escala, sua natureza é predominantemente descritivo-estatística. Por sua vez, plataformas como o *Wiktionary* democratizam a autoria dos verbetes via *crowdsourcing*, mas frequentemente reproduzem a lógica do consenso enciclopédico, buscando estabilizar um sentido fixo e universal. A biolexicografia distribuída distancia-se de ambos por não almejar o congelamento estatístico ou colaborativo do signo; seu foco recai sobre o rastreamento ontológico da palavra, mapeando o fluxo estigmérgico, a variação pragmática e as tensões ideológicas que impedem o sentido de se fechar em um verbete pacificado.

A proposta é menos uma implementação tecnológica imediata e mais uma provocação teórica: pensar a legitimidade lexical a partir da agência da rede falante. Essa perspectiva permite recolocar o papel do falante na lexicografia. Na tradição centralizada, o falante aparece muitas vezes como fonte de dados a ser observada por especialistas. Na biolexicografia distribuída, ele passa a ser agente de validação, circulação e transformação do léxico. O dicionário, nesse caso, não desaparece; ele muda de natureza. Deixa de ser arquivo de autoridade e torna-se ecologia de sentidos. Sua tarefa não seria congelar a palavra, mas acompanhar sua vida social, preservando a memória das disputas que a constituem.

Essa mudança possui implicações políticas relevantes. Quando termos periféricos, digitais, regionais ou identitários são ignorados pelos repertórios oficiais, não se apaga apenas uma palavra: apaga-se também uma forma de experiência social. A ausência lexical pode funcionar como forma de invisibilização simbólica, aproximando-se daquilo que Bourdieu (1989) compreende como violência simbólica. Por isso, uma biolexicografia distribuída teria o compromisso de registrar não apenas o que já foi autorizado pela norma, mas aquilo que emerge das bordas, dos conflitos, das comunidades de prática e dos usos insurgentes. Nesse sentido, ela se aproxima de uma lexicografia crítica, comprometida com a pluralidade da língua e com a resistência aos apagamentos históricos.

Contudo, essa proposta também exige cautela. A descentralização tecnológica não garante, por si só, democratização. Redes digitais podem reproduzir desigualdades, amplificar discursos dominantes, favorecer bolhas de repetição e excluir sujeitos com menor acesso tecnológico. Como alerta Van Dijck (2013), as plataformas não são infraestruturas neutras, mas ecossistemas guiados por lógicas de datificação e controle algorítmico. O risco de um utopismo tecnológico reside em ignorar que a validação coletiva nas redes pode ser facilmente manipulada por fazendas de cliques (*bot farms*), campanhas de falso engajamento *(astroturfing)* e disseminação coordenada de desinformação. Se o

protocolo *Proof of Meaning* depender unicamente do volume estigmérgico, corre o grave perigo de legitimar vocabulários de ódio ou preconceito apenas pela sua tração métrica.

Essa vulnerabilidade expõe uma tensão inerente entre a mera descrição lexicográfica e a necessidade de moderação. Um dicionário descentralizado não pode ser ingenuamente descritivo; ele exige protocolos de governança comunitária que saibam diferenciar a evolução orgânica do léxico das distorções algorítmicas artificiais. Scott (1998) alerta para os riscos de sistemas que tornam a realidade excessivamente legível para fins de controle; de modo semelhante, uma lexicografia distribuída não pode substituir a velha autoridade editorial por uma nova autoridade algorítmica. O desafio é impedir que a promessa de participação se converta em nova forma de exclusão. A soberania lexical plena só faria sentido se vinculada à acessibilidade, à transparência, à pluralidade e à participação efetiva dos sujeitos.

Para que essa ecologia de sentidos não sucumba a manipulações ou legitime violências sob o pretexto do volume estigmérgico, a mediação exige um modelo sociotécnico híbrido, superando a falsa dicotomia entre auto-organização irrestrita e curadoria centralizada. Nessa arquitetura, a comunidade atua na linha de frente via moderação por pares, contextualizando e tensionando ativamente os discursos em disputa. O lexicógrafo, por sua vez, desloca-se da clássica função de 'porteiro' *(gatekeeper)* do léxico para a de arquiteto e auditor do ecossistema. Sua tarefa não é aprovar ou vetar palavras por juízo de valor, mas moderar a integridade do sistema, calibrando os protocolos de governança para distinguir o uso social orgânico humano das trações artificiais automatizadas, garantindo assim que o *Proof of Meaning* seja um reflexo de inteligência coletiva autêntica e não de ações orquestradas.

Conclui-se que a biolexicografia distribuída oferece uma proposta teórica fértil para repensar a relação entre dicionário, linguagem digital e poder. Ao deslocar o foco da autoridade central para a circulação social do sentido, ela permite compreender a língua como organismo simbólico em constante transformação. Mais do que substituir a lexicografia tradicional, trata-se de expandir seus limites, incorporando a vitalidade dos usos digitais, a agência dos falantes e a memória das disputas semânticas. Em tempos de mutação acelerada da linguagem, registrar o léxico não pode significar aprisioná-lo. O dicionário do presente precisa aprender a respirar com a língua que pretende descrever[1].

[1] A utilização da ferramenta Gemini (versão Pro) foi empregada como apoio à reorganização textual, adequação linguística e revisão formal do manuscrito. Todo o conteúdo foi revisado criticamente pelo autor, que assume integral responsabilidade pela versão final, pela veracidade das informações, pelas interpretações apresentadas e pela adequada indicação das fontes utilizadas.

REFERÊNCIAS

BAKHTIN, Mikhail; VOLOCHINOV, Valentin. **Marxismo e filosofia da linguagem**: problemas fundamentais do método sociológico na ciência da linguagem**.** Tradução de Michel Lahud e Yara Frateschi Vieira. 12. ed. São Paulo: Hucitec, 2006.

BIDERMAN, Maria Tereza Camargo. A ciência da lexicografia**. Alfa: Revista de Linguística**, São Paulo, v. 28, p. 27-44, 1984. Suplemento.

BOURDIEU, Pierre. **O poder simbólico.** Lisboa: Difel, 1989.

GRASSÉ, Pierre-Paul. La reconstruction du nid et les coordinations interindividuelles chez Bellicositermes natalensis et Cubitermes sp.: la théorie de la stigmergie. **Insectes Sociaux**, v. 6, n. 1, p. 41-80, 1959.

LARSEN-FREEMAN, Diane. Complex, dynamic systems and English language teaching. In: BURNS, Anne; RICHARDS, Jack C. (ed.). **The Cambridge guide to pedagogy and practice in second language teaching.** Cambridge: Cambridge University Press, 2012. P. 165-172.

LÉVY, Pierre. **A inteligência coletiva**: por uma antropologia do ciberespaço. Tradução de Luiz Paulo Rouanet. São Paulo: Loyola, 1997.

MOITA LOPES, Luiz Paulo da. **Oficina de Linguística Aplicada:** a natureza social e educativa da linguagem. Campinas: Mercado de Letras, 1996.

SCOTT, James C. **Seeing like a state:** how certain schemes to improve the human condition have failed. New Haven: Yale University Press, 1998.

VAN DIJCK, José. **The Culture of Connectivity:** A Critical History of Social Media. Oxford: Oxford University Press, 2013.

Sobre o autor

Francisco André da Silva
francisco.anddre@gmail.com

Doutorando e mestre em Linguística Aplicada pela Universidade Estadual do Ceará (UECE), além de licenciado em Letras – Inglês pela mesma instituição. É professor de Língua Inglesa na educação básica e desenvolve pesquisas nas áreas de Linguística Aplicada, ensino de língua inglesa, estudos do léxico, lexicografia, discurso, tecnologias digitais e inteligência artificial generativa. Participa de grupos de pesquisa voltados aos estudos da linguagem e à formação docente, com interesse em práticas linguísticas situadas, produção de sentidos e relações entre linguagem, sociedade e poder.