

Etnografía digital nos estudos da linguagem

Perspectivas críticas, práticas e desafios metodológicos

Thaís Elizabeth Pereira Batista

Universidade Federal de Goiás (UFG), Fapeg, Brasil

ORCID: 0000-0003-3439-8363

Joana Plaza Pinto

Universidade Federal de Goiás (UFG), CNPq, Brasil

ORCID: 0000-0001-8052-9390

Carolina Fernanda Soares Silva

Universidade Federal de Goiás (UFG), CNPq, Brasil

ORCID: 0000-0002-0743-3664

Ana Luiza Krüger Dias

Universidade Estadual de Campinas (Unicamp), PIPD/CAPEs, Brasil

ORCID: 0000-0002-7538-995X

Resumen

La creciente presencia de la tecnología en nuestras interacciones cotidianas lleva a la necesidad de un análisis crítico del mundo digital en los estudios del lenguaje. Así, la etnografía digital se convierte en un enfoque metodológico interesante, debido a su compromiso con la descripción densa, ofreciendo perspectivas críticas sobre cómo ocurren las interacciones digitales, y sobre cómo se ha diseñado y experimentado el entorno digital. En vista de eso, este artículo¹ discute tres aspectos centrales de este enfoque metodológico: 1) los recursos y condiciones del mundo digital en el siglo XXI; 2) los contextos online y offline implicados en la historicidad del mundo digital; 3) los posibles caminos a través de las etnografías digitales. Las discusiones destacan que, para abordar la complejidad de lenguaje y de *big data*, necesitamos considerar sus vínculos con los procesos sociales y políticos, especialmente con la distribución desigual de las infraestructuras de la globalización, que no sólo determina las posibilidades y limitaciones de nuestras acciones en Internet, sino que en realidad produce el significado de “Internet”. Finalmente, presentamos algunas implicaciones metodológicas y desarrollos de esta propuesta, explorando algunos horizontes éticos para la etnografía digital, como una forma de deconstruir ilusiones comunes sobre este campo.

Palabras clave: Lenguaje, Internet, Metodología, Etnografía digital, Crítica.

Abstract

The increasing presence of technology in our daily interactions leads to the need for critical analysis of the digital world in language studies. Thus, digital ethnography becomes an interesting methodological approach, due to its commitment to thick description, offering critical insights not only about how digital interactions occur, but also about how the digital environment has been designed and experienced. Given this scenario, this article discusses three aspects central to this fieldwork: 1) resources and conditions of digital environments in the 21st century; 2) online and offline contexts implied in the historicity of the digital world; and 3) possible routes throughout digital ethnographies. The discussions highlight that, to deal with the complexity of language and big data, we need to consider their links to social and political processes, especially regarding the unequal distribution of globalization infrastructures, which not only determines the possibilities and limitations of our actions in the internet – but actually produce the meaning of “internet” itself. Finally, we present some methodological implications and developments of this proposal, exploring some ethical horizons for digital ethnography, as a way of deconstructing common illusions about this field.

Keywords: Language, Internet, Methodology, Digital ethnography, Criticism.

1 Introdução

Originada na Antropologia, a etnografia é uma abordagem de longa data. Contudo, algumas áreas do conhecimento, como filologia, psicanálise, direito, entre outras, a experimentaram mais tarde, na medida em que precisaram ajustar as ferramentas para os seus problemas de pesquisa, conforme apontam Blommaert e Jie (2010). Ainda na Antropologia, a etnografia conquistou o status de rigor científico na perspectiva moderna de ciência a partir do trabalho desenvolvido pelo polonês Bronisław Malinowski nas primeiras décadas do século XX. Naquele contexto, era considerado científico um trabalho que tivesse como características a neutralidade e o distanciamento do pesquisador de seu objeto de pesquisa, ainda que esses “objetos” fossem pessoas em interações.

Para Hymes (1996), a etnografia tem o interesse de estudar as pessoas e seus mundos, podendo se integrar a outros modos de fazer pesquisa. Mas isso não é consenso entre quem defende a tradição etnográfica, já que nem todas as vertentes, sobretudo as mais tradicionais, consideram esse tipo de integração interdisciplinar algo positivo.

Para nós, vale a pena “eleger o método e a perspectiva etnográfica como muito eficientes para olhar problemas e questões de linguagem” (Batista 2020: 300). Adicionalmente, consideramos pertinente o olhar da etnografia multissituada (Briggs 2007), que busca gerar dados em diferentes ambientes, inclusive os digitais, que se sobressaem no cenário interacional atual. Consideramos ainda que uma perspectiva teórico-metodológica para abordar questões de linguagem nesses moldes deve se apresentar de forma crítica, permitindo “que a pesquisa seja guiada e moldada a partir de um conhecimento aprofundado do campo” (Batista 2020: 301).

No percurso da produção de conhecimento científico em humanidades, a etnografia adentrou os estudos linguísticos, influenciando áreas como Pragmática, Sociolinguística e Análise do Discurso, tendo como ponto forte as descrições detalhadas dos fenômenos, característica que também é criticada pelo modelo de ciência moderna, centrada no positivismo de seus precursores. Em posição crítica a esse positivismo, nos alinhamos com o seguinte pensamento:

[...] a etnografia retira o foco sobre o espécime humano, o indivíduo, e se interessa por grupos (pequenos ou grandes) de pessoas de alguma forma relacionadas (Blommaert; Jie, 2010; Mattos, 2011). Essa natureza coletiva da abordagem etnográfica é pertinente para o estudo das relações sociais (Coletiva Ciborga 2022: 26).

Mas o que tudo isso tem a ver com o campo de pesquisa digital? Esse compromisso etnográfico com a descrição densa dos fenômenos investigados, no lugar de pressuposições *a priori* feitas em outras abordagens antes mesmo do início do trabalho de campo, torna a etnografia interessante para a pesquisa no campo digital. Isso porque a experiência humana, em si, é caótica em todos os espaços, e a etnografia oferece “uma pluralidade criticamente sustentada para lidar com o aparente caos da linguagem no ambiente digital” (Coletiva Ciborga 2022: 27).

Numa obra que já ultrapassa duas décadas, Christine Hine (2000) defende que a etnografia digital pode ajudar a responder uma série de problemas de pesquisa, como a existência ou não de fronteiras entre online e offline, como a internet tem afetado as relações sociais, além de questões sobre performance e autenticidade. Para a autora, uma abordagem etnográfica crítica é capaz de investigar até mesmo as interações de grupos de pessoas na internet, já que ela não prescinde da interação face a face necessária nas abordagens mais tradicionais.

Dessa forma, buscando produzir conhecimentos críticos sobre a linguagem no ambiente digital, a pesquisa etnográfica tem por objetivos observar, descrever e discutir como a linguagem afeta o ambiente digital e como o ambiente digital afeta a linguagem.

A etnografia digital nos estudos da linguagem tem muito a colaborar nesse cenário, pois deve produzir conhecimento fundamentado não apenas sobre as interações que ocorrem no ambiente digital, mas também sobre o que essas interações têm a ver com as nossas ações pela linguagem moldadas pela forma como o ambiente digital tem sido pensado e desenhado de modos específicos. Como agimos linguisticamente nesses ambientes e com quais possibilidades e limites? (Coletiva Ciborga 2022: 23).

No desenvolvimento da etnografia digital como campo de pesquisa, foram observadas algumas formas de empreender pesquisa etnográfica no ambiente digital. De acordo com Frago, Recuero e Amaral (2011), diferentes abordagens etnográficas nesse contexto foram nomeadas como etnografia digital, etnografia virtual, netnografia, ciberantropologia e webnografia.

Com base nessas autoras, sintetizamos essas abordagens, ressaltando suas principais características. A netnografia foi criada nos anos 1990 no campo da comunicação e do marketing, onde também é usado o termo webnografia. A ciberantropologia também foi pensada na década de 1990, com o objetivo de pesquisar seres humanos e seus meios de conexão, com base nos estudos de Donna Haraway (1991) e sua discussão sobre o ciborgue e a reinvenção da natureza. Já a etnografia virtual surgiu com Hine (2000), com o objetivo de adaptar a etnografia ao ambiente virtual. Por fim, a etnografia digital, como estamos chamando aqui, se coloca como uma ampliação do conceito de etnografia virtual e propõe a criação de materiais e narrativas que integrem o nexos online-offline para que possam ser usados nas pesquisas acadêmicas, mas que também cheguem a um público fora da academia (Coletiva Ciborga 2022).

Reconhecemos a necessidade de pensar em abordagens de pesquisa que se mostrem pertinentes aos contextos nos quais interagimos. Com o aumento cada vez maior das tecnologias em nossa vida, as fronteiras entre online e offline podem se mostrar muito fluidas, problemáticas ou mesmo inexistentes, de modo que pesquisas sobre linguagem precisam buscar analisar criticamente os desafios do digital.

Com essa perspectiva, colocamos algumas questões: como lidar criticamente com a metodologia qualitativa etnográfica diante de megadados digitais e do nexos online-offline? É possível fazer “estudo de caso” rigoroso num ambiente de conexões de terabytes de dados e dezenas de milhares de pessoas em escalas diferentes de interação?

Ante tais questionamentos, neste artigo, temos o objetivo principal de focar em três aspectos que têm sido centrais na perspectiva crítica da etnografia digital como metodologia nos estudos da linguagem: a identificação e descrição de recursos, condições e configurações que informam o digital no século XXI; os vários contextos (online e offline) implicados na historicidade do digital; os tipos de trânsitos e percursos possíveis nesses ambientes. Retomaremos ao final uma síntese das implicações e desdobramentos metodológicos desses aspectos, articulados aos horizontes éticos para a etnografia digital. Acreditamos que um olhar crítico e ético para os estudos etnográficos da linguagem possa render bons frutos nessa seara de pesquisa.

2 Recursos, condições e configurações do ambiente digital

Blommaert (2013) aponta a década de 1990 como um marco importante para o que ficou conhecido como globalização geocultural, já que a criação da *Web 2.0* permitiu a disseminação de informações de maneira intensa e volumosa. Desde então, o aumento do acesso às redes de internet tem distribuído as interações online de forma muito mais intensa, modificando os nossos modos de vida na contemporaneidade. Ao mesmo tempo, conforme Wang et al (2014), as conexões possibilitadas por infraestruturas de globalização são distribuídas de forma desigual ao redor do globo, escancarando hierarquias sociais e raciais historicamente sedimentadas e promovendo uma complexa organização nos padrões comunicativos, já que

[...] percebemos que ‘no contexto da globalização, recursos linguísticos mudam valor, função, direito de propriedade/domínios, e assim por diante, porque eles podem ser inseridos em padrões de mobilidade’ (Blommaert, 2010, p. 32). (Batista 2022: 21)

Em tempos de redes sociais e popularização de *smartphones*, é possível observar como a distribuição de recursos é organizada de forma a moldar e ser moldada por padrões comunicativos, em uma simbiose com o modelo capitalista global vigente. Na sociedade capitalista, as interações no nexos online-offline dizem respeito às ações que acontecem no ambiente online também estarem diretamente relacionadas às ações offline (Coletiva Ciborga 2022). Esse nexos assume contornos específicos, alinhados com os objetivos centrais desse modelo econômico, manutenção de lucros a qualquer custo.

Com a pandemia mundial de Covid-19 em 2020, que levou ao isolamento social por um período, a interação mediada por tecnologias escalou vertiginosamente, levando ao ápice de um processo que estava em curso e expondo ainda mais as diferenças sociais, como discutido em Thaís Batista (2022).

Tarnoff (2022) discute como a internet, que surgiu sob um discurso de democratização da informação e do conhecimento, tem sido objeto de privatização que atende à demanda do capitalismo global por maximização dos lucros. Isso está levando a internet do século XXI a ser

governada pelo capital, pelo lucro, e não pelas pessoas, de modo que cabe questionar: como as pessoas moldam e são moldadas nesse espaço digital que está totalmente conectado e atrelado ao seu cotidiano, que agora acontece no nexo online-offline?

Para pensar em questões como essa, podemos resgatar um exemplo ocorrido em 2021, quando houve o mais longo apagão de várias plataformas da empresa de tecnologia Meta, de propriedade de Mark Zuckerberg, detentora de várias redes sociais famosas, como Facebook, Instagram e WhatsApp. A pane nos sistemas de um único conglomerado tecnológico dirigido por um grupo muito reduzido de empresários dos Estados Unidos por pouco mais de 5 horas causou um alvoroço de escala mundial. Isso porque os aplicativos mais populares ficaram fora do ar, o que leva ao questionamento sobre o que significa internet hoje em dia para as pessoas: “para muita gente, ‘a internet’ é um conjunto de aplicativos de uma única empresa do Vale do Silício” (Coletiva Ciborga 2022: 21).

Mas se engana quem acha que isso não tem razão de ser ou que ocorre simplesmente por escolhas autônomas das pessoas por determinados aplicativos considerados mais “interessantes” ou “confiáveis”. De fato, os recursos e condições são manejados alheios aos interesses das pessoas:

Em 2015, o Facebook, empresa que controla o aplicativo de seu nome e outros dois onipresentes nas nossas vidas (WhatsApp e Instagram), fez um acordo com o governo brasileiro para levar conexão de banda larga para populações de baixa renda, com a promessa de promover o maior acesso a serviços públicos por meio da rede, como parte de uma iniciativa liderada pelo Facebook com o objetivo de ampliar o acesso à rede (Coletiva Ciborga 2022: 21-22).

Mas o que essa promessa de ampliação de acesso gera é o monopólio de uma empresa que torna as pessoas dependentes dela e restritas a poucas plataformas. Além de monopolizarem a comunicação – de modo que pessoas e pequenas empresas passam a ter esses aplicativos como únicas ferramentas de trabalho, e quando elas ficam fora do ar, a comunicação fica inviável mesmo que a infraestrutura da internet esteja funcionando, também ocorre um monopólio da veiculação da informação, pois muitas pessoas se informam por essas redes, confiando mais nelas do que em outros veículos de comunicação consolidados ao longo do século XX, como a mídia impressa, de rádio e TV. É nessa esteira que um número pequeno de executivos e investidores passam a deter o poder de realizar escolhas que atendam às chamadas *demandas de mercado*, em nome de toda a população.

Esse movimento produz vários efeitos nas práticas e usos da linguagem, inclusive a desinformação e a propagação de *fake news*, que se popularizaram nas redes sociais e tiveram papel importante na ascensão da nova direita ultraconservadora que tem tomado o poder em vários países (Maly 2024). Por isso, existe uma luta transnacional pela regulação das empresas detentoras desses monopólios capazes de manipular a vida política por meio dos ambientes digitais e extraindo dados de maneira irrestrita (Thibout 2022).

Nesse contexto, é possível vislumbrar as potencialidades da etnografia digital para os estudos da linguagem. Embora os primeiros estudos de etnografia digital não tenham necessariamente sido relacionados com linguagem (do mesmo modo que já foram realizados vários estudos sobre linguagem no ambiente digital sem adotar a abordagem etnográfica), a articulação de ambos permite uma combinação estratégica, pois une o aprofundamento da pesquisa etnográfica com a sistemática de análise da linguística (Varis y Hou 2020).

Além disso, conforme aponta Santos (2016), as novas tecnologias se efetivam com o intermédio de práticas linguísticas, estimulando recursos e escalas de linguagem. Pela natureza diversa desses dados, é importante compreender seu acúmulo e como isso se relaciona com a linguagem. Assim, com a passagem de grande parte do fluxo midiático para as mídias digitais, “práticas linguísticas estão sendo mediadas por hardwares e softwares computacionais que, através de seus próprios designs, influenciam modos e escalas da linguagem e ação humanas” (Santos 2016: 08).

A ampliação do fluxo de dados gerado principalmente pela *Web 2.0*, mas também por outras plataformas de tecnologia, resultou em um aumento desmedido da massa de dados digitais, o que passou a ser chamado de *big data*, um desafio para os modelos disciplinares tradicionais que procuram reduzir seus objetos de pesquisa (Santos 2016).

Mas mesmo em modelos teóricos e metodológicos interdisciplinares e que buscam olhar os problemas de pesquisa de forma mais ampla, lidar com a complexidade da linguagem articulada com a complexidade do *big data* pode ser desafiador (conforme discutimos na seção 4).

Diante disso, a próxima seção aborda especificidades dos vários contextos (online e offline) implicados na historicidade do campo na etnografia digital. Esse debate é fundamental para refletir criticamente não só sobre o que as pessoas estão fazendo na internet, mas também sobre o que a internet tem feito para e com as pessoas, como esse ambiente molda e é moldado nas interações.

3 Contextos, transformações e a historicidade do campo digital

3.1 Interfaces digitais e mediações de contextos

Ao contrário do que se possa pensar apressadamente, as pessoas humanas usuárias não são agentes nas mídias digitais, mas *ambiente* para a *agência* de sistemas não humanos: não somos livres e soberanas nas redes e entender isso é primordial para não cairmos na lógica de influência que ampara a hodierna economia digital, bem como suas reverberações políticas e epistêmicas:

[...] o comportamento aparentemente livre dos usuários se desenrola numa interface montada para produzir, nos ‘bastidores’, *outros* tipos de efeito, de ordem propriamente sistêmica, dos quais os usuários são, por *default*, totalmente alienados (Cesarino 2022: 107).

Desse modo, “os usuários constituem o ambiente a partir do qual os algoritmos aperfeiçoam suas habilidades” (Cesarino 2022: 106), ao passo que a *agência* de usuárias e usuários se atém ao plano local dos conteúdos postados e visualizados. Dessa forma, as mídias digitais não nos oferecem um acesso não mediado do real, mas sim uma visão fragmentada e personalizada: realidades produzidas não somente simbolicamente, mas também materialmente (infraestrutura técnica), lhe dando forma (*afford*). Não existem metaenquadramentos estáveis, confiáveis e a linguagem no campo é sobrecarregada pelo potencial enunciado falseado, editado e filtrado (Cesarino 2022).

Uma pesquisa realizada em contexto digital precisa da compreensão das mudanças constantes que surgem no campo e das limitações pragmáticas das programações. Como proceder no caso de uma pesquisa que acontece na rede social Instagram e a(s) página(s) e/ou postagens em análise “desaparecem” (são deletadas) *aparentemente* sem motivo? Por isso a importância de que tudo considerado relevante para os objetivos da pesquisadora seja registrado (com a captura de tela, por exemplo) a fim de evitar a “surpresa” da “perda” de dados. Lidar com o campo digital

é lidar com imprevistos, e como vamos nos adaptar a eles é a questão-chave. Nesse sentido, a pesquisa etnográfica é, em seu cerne, um processo de construção de conhecimento guiado pelas descobertas e experiências acumuladas em campo (Varis 2014), ou seja, o próprio campo guiará a pesquisadora a seguir, ou não, com suas escolhas prévias. Assim,

[...] as formas de geração e interpretação dos dados também envolvem os movimentos do campo - o que inclui considerar os eventuais (diríamos certos!) imprevistos que irão ocorrer ao longo do processo de pesquisa. Caso esse campo envolva ainda a participação de outras pessoas, é importante que se estabeleça uma ética de decisão coletiva sobre os processos da etnografia, reconhecendo as tensões que atravessam a pesquisa e as constantes negociações que precisamos fazer diante delas (Coletiva Ciborga 2022: 54).

Outro ponto a ser considerado é o do nexo online-offline, que atravessa aspectos dos dois ambientes e gera complexidade e multiplicidade aos eventos (Blommaert y colaboradores 2019). É possível, durante a exploração do campo digital, que seja feita a escolha focal de uma rede social ou um portal, por exemplo, e, a partir daí, seja delineada a circulação dos textos nesse contexto. Mas também (e por que não, juntamente) existe a possibilidade de acompanhar as transformações dos textos nessa plataforma, mas que caminha e se modifica por outras, considerando particularidades dos *affordances* em cada ambiente (Coletiva Ciborga 2022).

O termo *affordance* tem sido usado, por um lado, para tratar de “um certo grau de objetividade do mundo material que limita nossos atos de linguagem” e, por outro, “[d]as circunstâncias materiais objetivas do mundo [que] não são determinísticas e estáticas, porque os contornos materialmente regulados são somente uma possibilidade que pode ou não ser realizada” (Coletiva Ciborga 2022: 44). Portanto, *affordances* “não são propriedades fixas das plataformas, mas potencialidades que *emergem* entre a arquitetura de mídia e o comportamento dos usuários” (Cesarino 2022: 100).

As interações online emergem, nesse sentido, em transformações linguísticas acopladas nos *affordances* das plataformas, cuja estrutura depende de um enquadre próprio para hierarquias predefinidas das participações. Enquadre pode ser definido como “aquele recurso que coloca limites ao nosso ‘campo visual’ interpretativo, referenciado em valorizações culturais [...] nas bases ideológicas do acesso ao significado” (Pinto 2019: 226). Os *affordances* limitam os enquadres porque definem o “campo visual” da linguagem na rede social – em que quantidade podemos postar, quais os tipos de textos, arquivos ou links podemos incluir, quem pode ou não “ver” o que postamos, e assim por diante – assim como transformam os enquadres ao longo dos movimentos dos textos, arquivos ou links nos diferentes ambientes – como no caso de uma postagem que pode ser recortada para extrair trecho a ser recontextualizado em outra rede de interações, assentando um novo enquadre (Coletiva Ciborga 2022).

Assim, podemos nos defrontar com transformações como na etnografia digital de Carolina Silva (2022), que se deparou com a circulação do cartaz de Rosie, a Rebitadora², em múltiplos contextos digitais: diversas edições com textos feministas e antifeministas em diferentes plataformas, como a imagem do cartaz grafitado em um muro e recontextualizado em uma foto do muro em uma página no Facebook; a imagem recontextualizada em uma xícara num banheiro compondo o arquivo de transformações da figura em ambientes digitais. Por isso, não podemos ignorar o hibridismo (Blommaert y colaboradores 2019) de ambos os espaços enquanto estamos pesquisando e também em nossas próprias vidas integradas às nossas investigações digitais.

Mas como podemos entender melhor essas transformações dos textos e enquadres nos nexos online-offline e *affordances* do contexto digital? Não se trata de uma pergunta de fácil resposta, já que há variados debates em torno da própria noção de *contexto* e, no que diz respeito ao

online, isso se complexifica ainda mais. Apesar de a desordem do campo digital ser aparentemente mais caótica que a das interações face a face, os eventos sociais acontecem em camadas contextuais, em eventos considerados como microcontextuais (a exemplo dos eventos casuais, como um vídeo em um perfil a ser investigado no TikTok ser deletado ou de a pesquisadora perder acesso à internet no momento exato em que ocorre uma transmissão ao vivo de seu interesse de pesquisa) e outros macrocontextuais (a queda de uma plataforma, a restrição de acesso por razões de censura ou regulação, entre outros eventos históricos, políticos, sociais e culturais) (Blommaert y Jie 2010). Nesse sentido, aqui, os contextos são sedimentados e policêntricos, o que costuma fazer com que as pesquisas acabem se encarregando mais do produto textual ou discursivo, que pode ser ilusório, do que dos processos textuais e discursivos, o que nos força a procurar um “fundo contextual” para além do que é possível ver na tela (boyd y Crawford 2012; Varis 2014).

3.2 Contradições entre colapso e expansão dos contextos

Alice Marwick e danah boyd (2010), ao discutirem o colapso contextual no ambiente digital, salientam que esses contextos são multiplicados e, além disso, as audiências são indeterminadas e sobrepostas em um mesmo evento de interação – em que a disposição de tempo e espaço se torna mais intrincada. Esse colapso é mais evidente nas redes sociais do tipo Facebook ou Twitter, onde uma postagem direcionada para uma audiência de colegas de trabalho potencialmente pode ser lida por pessoas da família.

Complicando a noção de colapso de contexto, Blommaert, Smits e Yacoubi (2018) propõem a ideia de expansão de contexto, que explica como ações sociais são desenvolvidas colaborativamente através de recursos disponíveis e de expectativas normativas em relação a formas específicas de organização social. Conforme os autores, as pessoas possuem uma capacidade expansível e dinâmica em suas interações nos contextos digitais, de modo que atuam nas contextualizações como um processo de negociação ativo e reflexivo. Nesse sentido, o contexto não colapsa; é a noção de contexto estável que deixa de ser proveitosa.

Para Cesarino (2022), a questão é menos sobre o que as pessoas podem ou não fazer na interface com a mídia e mais sobre uma propriedade transversal à infraestrutura que reflete não apenas em ordenações de experiências individuais, mas, principalmente, na organização do sistema sociotécnico como um todo. Ou seja, o colapso de contextos indica como as novas mídias desconstróem a metacomunicação em sociedade.

Mesmo que fazer pesquisas em ambientes digitais seja um trabalho árduo (Blommaert 2020), apagar etapas e dados “inoportunos” contradiz a realidade de contextos em colapso e em expansão. É necessário inovação e flexibilidade, de maneira que se reconheça e explicita questões e desafios implicados na investigação online, especialmente para observar e descrever o nexo online-offline e os manejos de contextos e *affordances* (Barton y Lee 2017).

Refletindo sobre contextos dinâmicos e não-lineares e sua relação com processos sociais e políticos, Cesarino (2022) argumenta que agentes cibernéticos atuam regularmente para ajuste do comportamento no ambiente, em interações aparentemente sem consequências ou ao acaso. Ainda assim, é possível obter alguma espécie de ordem, tendo em vista que a conduta desses agentes cibernéticos não é contingente: “a atual infraestrutura das novas mídias possui um *viés político* [...] favorável à direita iliberal, aos conspiracionismos e às demais forças antiestruturais em seu entorno” (Cesarino 2022: 87).

Essa reflexão se articula com outro aspecto importante aqui: a ideia de que a internet é incorporada, corporificada e cotidiana (Hine 2015). Isso porque está entrelaçada a variadas formas de contexto e enquadres de sentido (incorporada); integrada nas vidas e nas formas de produção de sentido em que é difícil separá-la dos corpos (corporificada); e, por assim ser, se faz presente na vida das pessoas todos os dias, o tempo todo, o que pode gerar uma falta de atenção crítica por poder aparentar ser uma atividade não marcada (cotidiana). Logo, é capaz de *parecer ser* somente uma infraestrutura neutra e disponível em que podemos realizar as mais diversas ações de linguagem, mas

[p]ode ser importante examinar atentamente o trabalho que permitimos que as diversas aplicações da Internet em que confiamos façam por nós, e as decisões que implicitamente permitimos que essas infraestruturas tomem em nosso nome (Hine 2015: 46).

A própria virtualidade da internet se sustenta num aparato material, bem como sua interação com pessoas usuárias humanas corporificadas, produzindo, assim, efeitos reais nos corpos em forma de afetos, loops cibernéticos que nossa cognição satisfaz com as máquinas desenhadas para isso (Cesarino 2022). McLuhan (1964) já alertava que as tecnologias da informação se transformam em uma espécie de “extensão da mente” da falante. A “realidade” oferecida pelas redes, nesse sentido, está integrada à escala da cognição individual, forjada por meio de fragmentos de conteúdos gerenciados em grande parte por algoritmos e um pouco por outros usuários (Cesarino 2022).

Dessa maneira, há um treinamento cognitivo externalizado pelas plataformas atuantes enquanto infraestruturas, relativas por exemplo, à necessidade de estar conectada ou conectado constantemente para não “perder” algo. Podemos exemplificar esse manejo e cognição externalizada com o caso do *affordance* “rolagem infinita do *feed*”. Criada em 2006 pelo Facebook, esse tipo de rolagem privilegia, via algoritmização, postagens recentes e relação entre comportamentos contíguos. Seguindo essa tendência, o TikTok desenhou a “rolagem automática” dos vídeos, o que dispensa quem usa o aplicativo de tocar na tela enquanto assiste aos conteúdos. Então, o propósito desse tipo de *affordance* é, progressivamente, “reduzir a fricção para que o usuário permaneça mais tempo conectado” (Cesarino 2022:104). Trata-se, portanto, de um *affordance* desenhado para atender a uma economia da atenção.

As infraestruturas dizem respeito não somente ao aspecto técnico, mas também ao fenômeno humano, impactando também os processos de subjetivação. Para que isso aconteça, os agentes adicionam, crescentemente, aparatos técnicos, biológicos, linguísticos e culturais que inter põem as compreensões humanas em relação ao mundo e, para que exista esse “acoplamento” ao ambiente, a cognição precisa concordar com os algoritmos em algum âmbito (Cesarino 2022).

Cesarino (2022) apresenta um bom exemplo a partir da plataforma de buscas Google: ao buscarmos uma informação, ela também está à nossa procura, isso porque os algoritmos estão captando nossos dados comportamentais para procedimentos de clusterização – técnica de mineração de dados por meio do agrupamento de dados de pessoas usuárias variadas, a partir da identificação de padrões. As funções dos *affordances* também podem ser vistas a partir da “curtida” vislumbrada como uma forma de gerar afetividade com outra pessoa, mostrar interesses, criar redes, mas a plataforma que desenha essa propiciação, como o Instagram, tem o objetivo de *input* para clusterizar pessoas usuárias umas com as outras, a fim de produzir microdirecionamentos de anúncios, por exemplo.

A alienação técnica é a base para essa opacidade das redes e é organizada em três escalas: a ação tática camuflada nas plataformas; a assimetria entre plataformas e usuárias e usuários; e as orientações de confirmação próprias da cognição humana. Retomaremos ao final uma síntese das implicações e desdobramentos metodológicos desses aspectos, articulando-os aos horizontes éticos para a etnografia digital.

A seguir, nos aprofundaremos nos desafios existentes em torno da seleção, registro e sistematização do material de pesquisa, ou seja, nos tipos de trânsitos e percursos possíveis da pesquisa nesses ambientes.

4 Trânsitos e percursos possíveis da pesquisa linguística no ambiente digital

4.1 Os dados de linguagem na infraestrutura de dados digitais

As infraestruturas digitais não são triviais. Os algoritmos tomam decisões textuais junto conosco, produzindo câmaras de eco e bolhas, e processos de clusterização. Algoritmos homofílicos entregam um ambiente digital que converge ainda mais com o que já produzimos. Ainda que a maior parte das pessoas usuárias trate esse ambiente como uma caixa preta naturalizada como superfície para os próprios afetos e ações linguísticas, com ilusão de decisão unilateral e significado controlado, esses mesmos sujeitos recebem notícias cotidianamente sobre processos de *deep fake* e alterações textuais e semióticas construídas por IAs e outras programações.

Diante dessas infraestruturas digitais complexas e opacas, enfrentamos desafios para lidar com a seleção de material, registro e sistematização nos ambientes digitais.

Desde o final da década de 1990, os modelos de análise de dados digitais detectaram alta velocidade de aumento exponencial de uma enorme variedade de dados, com projeções de dobrar o volume de dados a cada dois anos. O potencial do processamento deste enorme volume de dados é reconhecido por empresas e governos desde o início dos anos 2010 e impactou profundamente a forma de funcionamento do ambiente digital, porque programas são cada vez mais orientados para a criação de condições de uso dos dados – e menos pela criação de lógicas diferenciadas de processamento (Chen, Mao y Liu 2014: 1). Essa centralidade dos dados – ou do conjunto conhecido como *big data* (megadados) – vem reorganizando os próprios recursos online e condicionando nossas participações.

Como o *big data* contém valores enormes, dominá-lo significa dominar os recursos. Através da análise da cadeia de valor de *big data*, pode-se perceber que seu valor vem dos próprios dados, tecnologias e ideias, e o núcleo são os recursos de dados. A reorganização e integração de diferentes conjuntos de dados pode criar mais valores associados à sua mineração e manipulação. O manejo do *big data* lida com 5Vs: *Volume*, ou seja, uma enorme quantidade de dados; *Variedade*, ou seja, diferentes tipos de dados de diferentes fontes; *Velocidade*, ou seja, a rapidez de produção dos dados; *Valor*, ou seja, a extração de dados significativos; e *Veracidade*, ou seja, o grau de confiabilidade dos dados. A definição desses Vs é construída a partir do ponto de vista de quem analisa os dados.

Essa densidade precisa ser considerada no desenho da pesquisa, exploração do campo empírico, seleção, sistematização e armazenamento de materiais empíricos. É inevitável a combinação de procedimentos que levem em consideração a inseparabilidade online-offline, estabelecendo a relevância desse nexo para o objeto delimitado.

Como toda pesquisa etnográfica, a etnografia digital precisa levar em consideração os trânsitos e percursos na escolha do tema, na revisão bibliográfica, na exploração e contextualização do campo.

O ponto de vista de quem analisa os dados define o foco a partir do ambiente digital em direção aos textos a serem analisados, ou dos textos em direção aos ambientes digitais. Assim, as perguntas de pesquisa são historicamente sustentadas e dizem respeito a exequibilidade: que recursos materiais e financeiros estão disponíveis para realizar a pesquisa? Quais as condições de armazenamento – extensão dos arquivos, metadados, backups? Qual o tempo de dedicação e prazos? Em que condições funcionam os *affordances* dos diferentes espaços digitais percorridos e como seu funcionamento afeta a pesquisa – o que a pesquisadora pode ver, como está acessível, quais os sistemas de busca e visualização dos ambientes estudados?

Essas perguntas são indicadoras de práticas etnográficas centrais: descrição densa, conexão entre contextualização e categorias de análise, discussão das categorias em relação ao problema de pesquisa. A descrição densa ajuda a reconhecer as características dos ambientes estudados – quais os *affordances*, quais as condições para o nexos online-offline, como esses *affordances* têm sido usados pelas pessoas envolvidas, para realizar que ações linguísticas, e que agentes sociais, normas e valores são evocados nas ações. Cada parte da descrição é conectável com material empírico dos ambientes, como vídeos publicados, endereços de páginas eletrônicas e muitos e muitos prints de ações linguísticas. Os ambientes digitais nos passam a impressão de um depósito eterno de dados por causa dos 5Vs, mas o volume é sinônimo de perda em meio à variação decorrente das mudanças de algoritmos e das querelas de propriedade – a quem pertence o texto ou imagem postada publicamente numa rede social? Quem pode buscar por um texto específico? Quem detém as ferramentas de busca? Que portas estão abertas, quais estão fechadas, no decorrer do percurso da pesquisa?

Essas questões nos ajudam a enfrentar os desafios do ambiente digital em constante mudança e disputa. Essas mudanças e disputas não podem ser resolvidas com roteiros padronizados sobre a etnografia digital, mas sim com a desconstrução aberta de algumas ilusões comuns e com a consideração de algumas características das práticas linguísticas digitais durante o processo de seleção de material, registro e sistematização.

4.2 Ilusões comuns sobre as práticas linguísticas em ambientes digitais

Apresentamos essas ilusões, exemplificando-as com casos de pesquisas que nos ajudaram a desconstruí-las.

A primeira e mais desafiadora ilusão que precisamos desconstruir é a ilusão de transparência. Originada na Modernidade e infecciosa nos estudos da linguagem até hoje (Moita Lopes y Fabrício 2019), essa ilusão nos impede de encarar que a opacidade é a regra do ambiente digital. Isso significa que o acesso a caminhos nas redes, plataformas, sites, em suma na Internet, não é autoevidente e transparente. A pesquisa de Diego Silva (2018) indicia esse momento de opacidade como desafio metodológico para lidar com o sistema de busca de uma plataforma de legendagem voluntária feita por fãs (*fansubbing*). O sistema de busca não entregava mais do que a primeira página de resultados em todas as máquinas e buscadores testados – ou seja, o funcionamento ou o resultado disfuncional do sistema evidencia que nem tudo que era inserido na plataforma estaria disponível para o pesquisador, portanto, os sentidos construídos em torno do *fansubbing* foram descritos nos limites da sua opacidade.

Decorrente dessa ilusão de transparência, constrói-se a ilusão de que se centrar no “conteúdo” dos textos digitais resolveria o problema. No entanto, o “conteúdo” a que um perfil de pesquisadora tem acesso está adensado pelo “viés de confirmação”, como podemos ler na pesquisa de Letícia Cesarino (2022) sobre grupos antidemocráticos nas redes sociais. A pesquisadora identificou o alinhamento constante de feedback positivo dos algoritmos em termos de preferências e repetição *do mesmo*, retroalimentado pelo aparato da plataforma que possibilita consistência do “conteúdo” visto. Assim, o “conteúdo” dos textos e enunciados nos ambientes digitais precisa ser cuidadosamente analisado como parte do “viés de confirmação” a que a própria pesquisadora está exposta pelos modos de funcionamento dos espaços percorridos.

Centrar-se no “conteúdo” faz emergir uma terceira ilusão: a ilusão de que o desconhecimento estruturado/estrutural dos mecanismos de funcionamento dos ambientes digitais não afeta nossos desenhos de pesquisa. Amanda Vallada (2021), ao pesquisar as configurações do novo biologismo (Cameron 2009) no YouTube, se aproveitou do funcionamento do algoritmo de indicação de vídeos dessa plataforma para encontrar mais vídeos. No entanto, o algoritmo não utilizava os mesmos parâmetros de “conteúdo” que faziam parte do desenho de pesquisa – o novo biologismo – mas sim parâmetros desconhecidos sobre audiências clusterizadas no YouTube, tendo indicado cada vez mais vídeos de pregação evangélica para Vallada (2021).

Por fim, uma última ilusão em que apostamos é que os dados online estão disponíveis à pesquisa, como se nosso caminho entre os textos fosse controlado por agentes humanos, quando, de fato, os textos, sua distribuição e alcance “não são feitos para isso”. Isso ficou evidente na pesquisa de Karoline Silva (2020), cujo objetivo era explorar a circulação de notícias sobre a associação feita pelo buscador Google entre o texto “mulher negra dando aula” e imagens pornográficas. Esse caminho levou a pesquisadora a uma variedade de textos correlacionados em diferentes blogs de feminismo negro, blogs de tecnologia da informação que traziam explicações sobre como o buscador Google funciona, sites com textos expositivos sobre a repercussão negativa desse processamento enviesado do buscador, entre outros, tudo relacionado ao mesmo texto “mulher negra dando aula”. A busca inicial por entender a associação entre o texto “mulher negra dando aula” e imagens pornográficas mostrou não as bases textuais dessa associação, mas a interação entre agentes humanos (programadores) e não humanos (algoritmos) em face das repercussões negativas das notícias e a alteração que essa interação provocou nos resultados oferecidos pelo buscador.

4.3 Características das práticas linguísticas digitais no processo de seleção de material, registro e sistematização

Atentas a essas ilusões, defendemos especial zelo a seis características das práticas linguísticas digitais – algumas já mencionadas anteriormente e aqui sintetizadas.

1. Nexos online-offline dos dados: os ambientes digitais não funcionam apartados das interações face-a-face, portanto é mister seguir os textos nos trânsitos e percursos entre o online e o offline.
2. Sobreposições e fragmentos de textos através de diferentes plataformas digitais, mídia tradicionais e interações face-a-face: um debate numa rede social pode gerar um processo jurídico face-a-face offline que pode derrubar vídeos ou textos em uma plataforma de conteúdo. Esses movimentos dos textos digitais levam à transcendência espacial e a outra temporalidade na produção de sentidos que precisam ser consideradas nas nossas análises.

3. Metadados associados aos dados: além dos textos superficialmente visíveis nos ambientes digitais (voz ou escrita), seus sentidos são articulados com metadados associados pelos sistemas de *tag* e perfilação. Uma postagem é associada a outra nem sempre pelo que ela diz (“conteúdo”), mas pela marcação de seus metadados (uma hashtag, uma geolocalização, um perfil comercial etc.). Os metadados nem sempre são visíveis na superfície dos ambientes, por isso exigem uma leitura crítica da contextualização do material empírico analisado para evitar conclusões apressadas.
4. Clusterização de dados e enviesamento do caminho de quem pesquisa: associada à ilusão de que nossas pesquisas não são afetadas pelo funcionamento das plataformas, essa característica expõe o drama positivista dos dados neutros e nos convoca a assumir o conhecimento situado que produzimos em toda e qualquer pesquisa (Haraway 1995).
5. Articulação da linguagem com os dispositivos em uso: o acesso gratuito ao WhatsApp e Facebook em celulares pré-pagos em muitos países da América Latina e a anonimização de janela de navegação são algumas das práticas disponíveis apenas em certos dispositivos (celular, tablet, notebook ou computador de mesa), afetando que práticas de linguagem circulem mais que outras, como o aumento de mensagens encaminhadas em celulares.
6. Ética e níveis de privacidade: os metadados e dados visíveis em materiais públicos em ambiente digitais sempre nos colocam diante do dilema ético do que deixar visível em nossas análises: que parte dos processos semióticos devemos considerar “privada” e que parte devemos considerar de “interesse público” para a pesquisa? Essas perguntas expõem o horizonte ético da própria prática linguística, uma antiga problemática dos estudos da linguagem (Rajagopalan 2003).

Essa última característica da linguagem, a ética como prática linguística, nos impele a uma síntese das implicações e desdobramentos metodológicos desse debate para uma perspectiva crítica da etnografia digital nos estudos da linguagem.

5 Implicações e desdobramentos metodológicos: horizontes éticos para pesquisas linguísticas em ambientes digitais

Nas seções anteriores, discutimos as relações entre a metodologia etnográfica e o campo digital, destacando as complexidades de se fazer pesquisas linguísticas neste ambiente, especialmente quanto à maneira de funcionamento dos recursos e plataformas disponíveis para tanto e como isso afeta as ações linguísticas nonexo online-offline. Entre os principais desafios que identificamos enquanto pesquisadoras da linguagem em ambiente digital, estão as conexões de dados em escalas gigantescas, as formatações específicas dos espaços da internet que orientam o seu modo de operar e as ilusões de transparência sustentadas pela imaginação científica acerca do funcionamento dos ambientes digitais.

Para finalizar, exploramos os desafios éticos implicados nessas relações, incluindo as condições materiais e ideológicas sob as quais os dados são engendrados e operados, as desigualdades de acesso ao digital e as decisões eticamente orientadas que podemos construir enquanto pesquisadoras deste campo. Para nós, a etnografia digital deve produzir conhecimento fundamentado não apenas sobre o “conteúdo” das interações que ocorrem no ambiente digital, mas também sobre como o ambiente digital tem sido pensado e desenhado de forma a produzir tais interações.

Como já salientamos, as pesquisas etnográficas em ambientes digitais precisam delimitar seu objeto e a geração dos dados de modo a considerar a abundância de informações de diversas áreas disponíveis às pessoas usuárias das plataformas digitais, as quais são atualizadas constantemente, ou seja, considerar os 5Vs do *big data*. Afinal de contas, atualmente há 5.44 bilhões de pessoas usando a internet no mundo todo, o equivalente a 67,1% da população mundial, com uma tendência de crescimento, especialmente nos países de economia em desenvolvimento³:

Em 2023, mais de 2/3 da população do planeta se utilizavam das principais mídias sociais. Na condição de usuários eles disponibilizam seus dados que, ao serem extraídos, se transformam em commodities (metadados) e são armazenados em Big Datas e a seguir processados, a partir da orientação dos algoritmos. O uso ampliado da digitalização para além das mídias sociais, dentro do que hoje se chama de Economia de Dados, permite estimar que o volume de dados produzidos no mundo deve passar dos 33 zetabytes que estava em 2018, para 175 zetabytes em 2025, ampliando para 291 zetabytes em 2027. Algo inimaginável que demonstra o domínio da tecnologia digital como um setor transversal que como as finanças atravessa todos os demais⁴.

Isso significa que o uso das tecnologias digitais vem modificando rapidamente as formas como nós “experimentamos os espaços públicos e privados, permitindo incorporar as comunicações pela internet a novos domínios de interação social” (Hine 2016: 11), incluindo as próprias práticas acadêmicas, que, além de passarem a considerar o ambiente digital como um novo e importante campo de pesquisa, vêm utilizando cada vez mais algoritmos para realizar análises em larga escala de grandes quantidades de dados⁵.

Chris Anderson, em artigo de 2008, tece elogios para como essa prática altera positivamente o que conhecemos como método científico, que era até então construído a partir de hipóteses cuja testagem confirmava ou descartava uma teoria sobre determinada particularidade do mundo. Com o uso de tecnologias para manejo e sistematização de *big data*, passa a ser possível que analisemos grandes quantidades de dados sem necessariamente lançarmos hipóteses sobre seu funcionamento:

Quem sabe por que as pessoas fazem o que fazem? A questão é que elas fazem, e podemos acompanhar e medir isso com uma fidelidade sem precedentes. Com dados suficientes, os números falam por si (Anderson 2008: n.p.).

Nesse raciocínio, quanto maior o volume de dados, melhor será a qualidade das análises produzidas a partir deles; e se apenas supercomputadores e seus complexos softwares são capazes de processá-los, apenas as pesquisas que se valham deles serão realmente relevantes.

Para dannah boyd e Kate Crawford (2011), automatizar a pesquisa muda a definição de conhecimento, naquilo que as autoras chamam de “virada computacional” no pensamento e na pesquisa. É comum nos depararmos atualmente com discursos que concebem o algoritmo – um conjunto de regras previamente definidas executáveis por programas de computador – como a panaceia capaz de encontrar a “solução” para os “problemas” do mundo, que incluem, entre outros, a automação da indústria, a tradução e geração de textos e imagens, o desenvolvimento de assistentes virtuais e, é claro, a criação e alimentação de um mercado de consumo maciço. Entre os transbordamentos desses discursos na/para a ciência, está a imaginação de que o algoritmo fornece uma espécie de linha direta com o conhecimento bruto (boyd y Crawford 2011), graças à sua capacidade de reconhecer padrões em grandes volumes de dados. Isso significa dizer que o *big data* não alterou apenas o método científico, mas a própria concepção do que configura um objeto de conhecimento, na medida em que suas ferramentas não apenas auxiliam as cientistas na manipulação de um volume massivo de dados, mas efetivamente dão forma à realidade que elas medem.

O motivo pelo qual as pessoas fazem coisas, escrevem coisas ou fazem coisas é apagado pelo enorme volume de repetições numéricas e grandes padrões. Este não é um espaço para reflexão ou para as formas mais antigas de arte intelectual. Como escreve David Berry (2011, p. 8), o Big Data fornece ‘quantidades desestabilizadoras de conhecimento e informação que carecem da força reguladora da filosofia’. Em vez da filosofia – que Kant via como a base racional para todas as instituições – ‘a ciência computacional poderia então ser entendida como uma ontoteologia, criando um novo espírito ontológico’ enquanto uma nova constelação histórica de inteligibilidade (boyd y Crawford 2011: 4).

Tradicionalmente, o desenvolvimento de modelos teóricos explicativos da realidade faria com que dados que eram então incertos nos garantissem *valores de verdade* (Santos 2016). É curioso notar, neste movimento de mudança das ferramentas de análise de dados, que, ao invés de constituir uma ferramenta crítica capaz de mostrar a cumplicidade dos modelos científicos enquanto regimes ontológicos de verdade referencial produtores de exclusão – deslocando, portanto, as próprias ideias de verdade e universalidade enquanto um projeto corporificado no sujeito branco, masculino e eurocentrado – o uso de modelos matemáticos capazes de processar grandes massas de dados passa a operar enquanto um novo modelo de geração de verdades supostamente transparentes.

Nessa substituição do modelo científico pelo modelo algorítmico para a construção de explicações racionais e universais sobre o mundo, estamos diante do seguinte cenário: há um grande volume de dados impossível de ser processado pelo ser humano e apenas modelos matemáticos sofisticados são capazes de oferecer a saída para produzir análises relevantes no contexto do *big data*. Essa solução discursiva funciona enquanto estratégia muito eficaz para excluir boa parte das pessoas do debate público sobre o desenvolvimento das tecnologias algorítmicas em razão de sua “alta complexidade”, camuflando, ao mesmo tempo, seu papel na própria produção dessa imensa quantidade de dados, num ciclo que se retroalimenta.

Cathy O’Neil (2016) argumenta que os modelos algorítmicos que atualmente sustentam não apenas as nossas formas de interação na internet, mas controlam os movimentos de mercados e as políticas públicas, definem sua própria realidade e a usam para justificar seus resultados, num modelo autogestado e altamente destrutivo, que corporifica a própria noção de verdade no capitalismo tardio, cabendo a poucos privilegiados a sua interpretação: “tal como os deuses, estes modelos matemáticos são opacos e o seu funcionamento é invisível para todos, exceto para os sacerdotes mais importantes no seu domínio: matemáticos e cientistas da computação” (O’Neil 2016: 10). E quando esses sacerdotes “deixam as máquinas falarem por eles”, seus vereditos se tornam incontestáveis; afinal, os modelos, justamente por serem matemáticos, são os únicos capazes de produzir análises rápidas, precisas e supostamente imparciais sobre a realidade.

Contudo, é importante lembrar o óbvio: os números não falam por si sós. Os modelos algorítmicos nada mais são do que “opiniões embutidas de matemática” (O’Neil 2016:10) e, como tais, têm suas restrições e limitações, refletindo os valores daqueles que os desenham, desde a definição de seus objetivos, modos de funcionamento, público-alvo e, mais importante, sua capacidade de geração de lucro para as companhias detentoras de tecnologia. Seu uso não é desinteressado, tanto que muitas corporações de big tech buscam proteger de forma rígida os modelos algorítmicos e códigos de funcionamento de Inteligência Artificial enquanto propriedade intelectual.

Reconhecer a estreita relação entre os modelos algorítmicos que regem o funcionamento das plataformas digitais e os interesses mercadológicos embutidos no seu desenvolvimento e uso é fundamental para que possamos efetuar os deslocamentos necessários nas pressuposições sobre

objetividade e autoexplicação em contextos de *big data*. A pesquisa etnográfica digital pode se tornar, então, uma abertura de reflexões críticas sobre esse modelo, a partir da incorporação das perspectivas metodológicas discutidas neste artigo. Tais reflexões estão em estreita sintonia com o compromisso ético para pesquisadoras empiristas como nós, e envolvem, entre outras práticas, a produção coletiva de análises e demais materiais teóricos, a discussão sobre as assimetrias experimentadas no campo, a proteção da identidade das pessoas participantes das nossas pesquisas etnográficas e o cuidado com os dados gerados. Nesse sentido, mencionamos as diretrizes e recomendações da *Association of Internet Researchers* (AoIR), que lançou em 2002 um relatório de boas práticas acadêmicas para pesquisas digitais, atualizado em versões de 2012 e 2019. Em suma, tais recomendações se concentram numa atenção à pesquisa digital enquanto um processo contextualmente orientado, numa costura inextricável entre ética e metodologia:

Esta abordagem orientada para o processo e o contexto requer ainda o ponto desenvolvido por Annette Markham: *ética é método – método é ética* (Markham, 2006). Isto é, nossa escolha de métodos em relação a questões de pesquisa e design evocam questões éticas específicas – mas estas, por sua vez, moldam (ou deveriam moldar) nossas escolhas metodológicas. Além disso, especialmente porque somos forçados ao longo do próprio projeto de pesquisa a revisar o design de pesquisa original e as escolhas metodológicas – somos igualmente confrontados com a necessidade de revisitar nossas suposições e designs éticos iniciais. Uma virtude fundamental deste ponto é que ele ajuda a combater uma presunção comum de ‘ética’ como um item de checklist que configura essencialmente um obstáculo à pesquisa. Pelo contrário – como nossa experiência subsequente demonstrou – assumir uma atenção contínua à ética como inextricavelmente entrelaçada com o método geralmente leva a uma pesquisa melhor (franzke y colaboradores 2019: 4).

Portanto, as decisões éticas tomadas no contexto de pesquisas etnográficas da linguagem em ambientes digitais configuram, pois, um verdadeiro horizonte que deve nos levar a questionar quais são, afinal, as possibilidades e as limitações das nossas formas de agir na internet, para que possamos nos mover na direção de pesquisas menos ingênuas e mais comprometidas com as transformações que queremos ver no mundo.

Neste artigo, buscamos contribuir para uma abordagem crítica da linguagem em ambientes digitais, articulando a etnografia digital a processos sociais, políticos e infraestruturas tecnológicas que moldam nossas interações linguísticas nesses ambientes. Arelado a isso, considerar a linguagem como um eixo central nas etnografias digitais possibilita pensar criticamente também as desigualdades, articulando os impactos de sistemas técnicos e humanos na produção de significados. Nesse sentido, defendemos que a linguagem, em seu caráter performativo, é afetada pelos ambientes digitais que, com isso, fazem emergir características específicas das práticas linguísticas digitais. A etnografia digital, então, não apenas fornece ferramentas de pesquisa linguística, mas também oferta o potencial reflexivo sobre sua própria posição enquanto prática ética situada.

Referências bibliográficas

Anderson, Chris. 2008. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*, 23 jun. Disponível em: <https://www.wired.com/2008/06/pb-theory/>. Acesso em: 10 mai. 2024.

- Barton, David; Lee, Carmen. 2017. Methodologies for researching multilingual online texts and practices. En M. Martin-Jones y D. Martin. *Researching multilingualism: Critical and ethnographic perspectives*. New York: Routledge, pp. 141-154.
- Batista, Thaís E. P. 2020. Lentos viajantes: o olho errante ocidental na visão de ciência única. En L. A. da Silva, P. R. Gimenez, R. M. Canedo y R. Damasceno-Morais, eds. *Pesquisas 2020: linguística em foco*. Goiânia: Cegraf UFG, pp. 296-315.
- Batista, Thaís E. P. 2022. *As muitas linguagens em guerra com o português: A língua-afeto entre mulheres negras quilombolas em luta com as ideologias linguísticas no contexto universitário*. Tese (Doutorado em Letras e Linguística) – Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- Blommaert, Jan. 2020. Political discourse in post-digital societies. *Trabalhos em Linguística Aplicada* 59 (1): 390-403.
- Blommaert, Jan. 2013. *Ethnography, superdiversity and linguistic landscapes: Chronicles of complexity*. Bristol: Multilingual Matters.
- Blommaert, Jan; Jie, Dong. 2010. *Ethnographic fieldwork. A beginner's guide*. Bristol: Multilingual Matters.
- Blommaert, Jan; Smits, Laura; Yacoubi, Noura. 2018. Context and its complications. *Tilburg Papers in Cultural Studies* 208: 1-19.
- Blommaert, Jan; Szabla, Malgorzata; Maly, Ico; Procházka, Ondřej; Ying, Lu; Li, Kunming. 2019. Online with Garfinkel: Essays on Social Action in the Online-Offline Nexus. *Tilburg Papers in Culture Studies* 229: 1-106.
- boyd, danah; Crawford, Kate. 2011. *Six Provocations for Big Data*. En *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*. Oxford, 21 set, 2011. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431. Acesso em: 10 mai. 2024
- boyd, danah; Crawford, Kate. 2012. Critical questions for big data. *Information. Communication and Society* 15 (5): 662-679.
- Briggs, Charles L. 2007. Anthropology, Interviewing, and Communicability in Contemporary Society. *Current Anthropology* 48 (4): 551-580.
- Cameron, Deborah. 2010. Sex/gender, language and the new biologism. *Applied Linguistics* 31 (2): 173-192. doi:10.1093/applin/amp022.
- Cesarino, Leticia. 2022. *O mundo do avesso: verdade e política na era digital*. São Paulo: Ubu Editora.
- Chen, Min; Mao, Shiwen; Liu, Yunhao. 2014. Big Data: A Survey. *Mobile Networks and Applications* 19: 171–209. doi:10.1007/s11036-013-0489-0.

- Coletiva Ciborga. 2022. *Etnografia Digital: Um guia para iniciantes nos estudos da linguagem em ambientes digitais*. Goiânia: Centro Editorial e Gráfico da UFG.
- Fragoso, Suely; Recuero, Raquel; Amaral, Adriana. 2011. *Método de pesquisa para internet*. Porto Alegre: Sulina.
- Franzke, Aline Shakti; Bechmann, Anja; Zimmer, Michael; Ess, Charles; y the Association of Internet Researchers. 2020. *Internet Research: Ethical Guidelines 3.0*. Disponível em: <https://aoir.org/reports/ethics3.pdf>. Acesso em: 02 jun. 2024.
- Haraway, Donna J. 1991. *Simians, Cyborgs, and Women. The Reinvention of Nature*. New York: Routledge.
- Haraway, Donna J. 1995. Saberes localizados: a questão da ciência para o feminismo e o privilégio da perspectiva parcial. *Cadernos Pagu* 5: 7–41.
- Hine, Christine. 2000. *Virtual Ethnography*. Londres: Sage.
- Hine, Christine. 2016. Estratégias para etnografia da internet em estudos de mídia. Em B. Campanella y C. Barros (Org.). *Etnografia e consumo midiático: novas tendências e desafios metodológicos*. Rio de Janeiro: E-papers, pp. 11-27.
- Hine, Christine. 2015. *Ethnography for the Internet: embedded, embodied and everyday*. London: Bloomsbury Academic.
- Hymes, Dell. 1996. *Ethnography, linguistics, narrative inequality. Toward an understanding of voice*. New York: Taylor & Francis.
- Maly, Ico. 2024. *Metapolitics, Algorithms and Violence: New Right Activism and Terrorism in the Attention Economy*. New York: Routledge.
- Marwick, Alice E.; boyd, danah. 2010. I tweet honestly, I tweet passionately: twitter users, context collapse, and the imagined audience. *New Media & Society*, 13 (1): 114-133.
- McLuhan, Marshall. 1964. *Understanding media: The extensions of man*. New York: McGraw-Hill Book Company.
- Moita Lopes, Luiz Paulo da; Fabrício, Branca F. 2019. Por uma “proximidade crítica” nos estudos em Linguística Aplicada. *Calidoscópio* 17 (4):711-723. doi:10.4013/cld.2019.174.03.
- O’Neil, Cathy. 2016. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Crown Publishers.
- Pinto, Joana Plaza. 2019. ‘É só mimimi?’ Disputas metapragmáticas em espaços públicos online. *Interdisciplinar* 31: 221–36.
- Rajagopalan, Kanavillil. 2003. *Por uma linguística crítica: linguagem, identidade e a questão ética*. São Paulo: Parábola Editorial.

- Santos, Vinícius V. V. dos. 2016. *Big Data, Meio e Linguagem: novas tecnologias e práticas linguísticas*. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- Silva, Carolina F. S. 2022. “*Nós (não) podemos fazer isso!*”: *Fios emaranhados das trajetórias feministas e antifeministas do cartaz “Rosie, a Rebitadora”*. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- Silva, Diego G. M. 2018. “*Qualidade é insubstituível!*”: *ideologias linguísticas em manuais de legenders independentes de equipes de fansubbing*. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- Silva, Karoline de S. S. 2020. *Padrões de beleza, hipersexualização e racismo: a língua e o processo socio-histórico na imagem da mulher negra em ambiente virtual*. Monografia (Bacharelado em Letras:Linguística) – Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- Tarnoff, Ben. 2022. *Internet for the people: the fight for our digital future*. Verso: New York.
- Thibout, Charles. 2022. Les GAFAM et l’État: Réflexion sur la place des grandes entreprises technologiques dans le champ du pouvoir. *La Revue Internationale et Stratégique* 125 (1): 75–88. <https://doi.org/10.3917/ris.125.0075i>.
- Vallada, Amanda D. 2021. *Inventando a diferença: Ideologias linguísticas e história natural dos discursos do novo biologismo*. Dissertação (Mestrado em Letras e Linguística) – Faculdade de Letras, Universidade Federal de Goiás, Goiânia.
- Varis, Piia. 2014. Digital Ethnography. *Tilburg Papers in Cultural Studies* 104: 1-21.
- Varis, Piia; Hou, Mingyi. 2020. Digital approaches in Linguistic Ethnography. En K. Tusting (ed.). *The Routledge Handbook of Linguistic Ethnography*. New York: Routledge, pp. 229-240.
- Wang, Xuan; Spotti, Massimiliano; Juffermans, Kasper; Cornips, Leonie; Kroon, Sjaak; Blommaert, Jan. 2014. Globalization in the margins: toward a re-evaluation of language and mobility. *Applied Linguistics Review* 5 (1): 23-44.

Notas

1. Este artigo é fruto de debates realizados na Coletiva Ciborga, grupo de pesquisadoras críticas no campo da etnografia digital do qual fazemos parte. Sendo assim, agradecemos e reconhecemos a importância de Amanda Diniz Vallada, Bianca Alencar Vellasco, Karoline de Sousa Soares Silva e Maria Rosaria Colangeli de Souza para as formulações que apresentamos aqui. Qualquer equívoco ou imprecisão é de responsabilidade das autoras deste artigo.
2. Disponível em: https://pt.wikipedia.org/wiki/Rosie_the_Riveter. Acesso em: 21 jun. 2024.
3. Disponível em: <https://datareportal.com/global-digital-over-view#:~:text=There%20are%205.44%20billion%20internet,higher%20in%20many%20developing%20economies>. Acesso em: 17 mai. 2024.
4. Cf. <https://outraspalavras.net/tecnologiaemdisputa/teoria-e-se-estivermos-a-beira-do-platafor->

mesmo/.

5. Nesse sentido, a *Stanford Network Analysis Platform* (SNAP), por exemplo, é uma biblioteca contendo modelos de análise para redes massivas com centenas de milhões de nós e bilhões de arestas, “manipulando gráficos grandes com eficiência, calculando propriedades estruturais, e gerando gráficos regulares e aleatórios”. Disponível em: <https://snap.stanford.edu/>.