

As ciências sociais no Brasil e suas formas de auto-observação: uma sociologia das comunicações científicas.

Antonio Brasil Jr. y Lucas Carvalho.

Cita:

Antonio Brasil Jr. y Lucas Carvalho (2021). *As ciências sociais no Brasil e suas formas de auto-observação: uma sociologia das comunicações científicas*. XIV Jornadas de Sociología. Facultad de Ciencias Sociales, Universidad de Buenos Aires, Buenos Aires.

Dirección estable: <https://www.aacademica.org/000-074/76>

O “pensamento social” visto de longe: desafios para uma nova agenda de pesquisas

Antonio Brasil Jr. (PPGSA/UFRJ)

Lucas Carvalho (PPGS/UFF)

Numa definição minimalista, podemos arriscar dizendo que a área de pesquisa em “pensamento social no Brasil” se debruça sobre as interpretações da sociedade brasileira e seus produtores, bem como sobre os seus efeitos no processo social. Se uma certa visão estereotipada sobre a área ainda vê nesta especialização acadêmica um foco talvez excessivo em um conjunto muito limitado de textos e autores da primeira metade do século XX – aparentemente vistos como em si mesmos “importantes” –, a verdade é que, nas duas últimas décadas, ampliou-se e se diversificou consideravelmente não só o que se entende por “interpretações da sociedade brasileira” mas também o que se considera “intelectual”. O que levou, como não poderia deixar de ser, a um debate bastante sofisticado, feito em diálogo com outras áreas de pesquisa, a respeito de novos modos de se pensar a relação entre texto e contexto, sobre o alargamento da noção da própria noção de intelectual e sobre o efeito (variado e em múltiplos sentidos) das ideias na modelagem de relações sociais. E, atravessando tudo isso, questões mais amplas, como a crise do Estado-nação – e a conseqüente problematização do “brasileiro” ou “no Brasil” – e a consciência mais aguda quanto ao caráter estruturante da ordem racial e de gênero, por exemplo, vem suscitando ângulos novos para a leitura dos textos “clássicos” ou criando o contexto para a discussão franca de textos pouco lidos ou até então quase desconhecidos.

Trata-se, pois, de uma área de pesquisa que vem sabendo se renovar e fazer frente aos desafios teóricos, metodológicos e empíricos postos pelas ciências sociais contemporâneas (Botelho, 2015).

Mais recentemente, um novo tipo de desafio foi colocado aos pesquisadores da área: como lidar com o chamado *big data*, isto é, pelo acesso a um conjunto de dados e informações sobre a vida social em escala até pouco tempo atrás inacessível aos cientistas sociais (Carvalho & Brasil Jr, 2020). Grandes organizações públicas ou

privadas geram diariamente volumes cada vez maiores de informações sobre os mais variados aspectos da vida social, o que vem suscitando o surgimento e a popularização crescente de ferramentas computacionais que permitem tratar esse material empírico através da anotação e da codificação de textos. Graças ao tipo de relação e de interação entre disciplinas e instituições propiciado pelo surgimento da Biblioteca Virtual do Pensamento Social, colocando lado-a-lado cientistas sociais, historiadores, biblioteconomistas, cientistas de dados etc., algumas formas coletivas e colaborativas de se enfrentar esse desafio começaram a ser ensaiadas (Brasil Jr. & Carvalho, 2018; Josiowicz & Brasil Jr, 2019). É o que vamos tratar aqui nesse texto.

Mas antes de seguirmos para a discussão de alguns estudos de casos que temos explorado a partir do grupo de pesquisa “Pensamento Social em dados” (*Data_PS*), que permitirão exemplificar algumas possibilidades de ligar o *big data* ao “pensamento social no Brasil”, vale explicitar o que talvez seja um pressuposto teórico-metodológico compartilhado: a “leitura distante” (*distant reading*), na acepção dada por Franco Moretti (2013) ao termo. Levar a sério a concepção de que a *distância* em relação ao texto pode ser uma condição do conhecimento na área de pesquisa que é o campo por excelência da “leitura cerrada” ou “em profundidade” (*close reading*) certamente parece provocador. Contudo, não se trata de traçar uma linha de ruptura ou muito menos sugerir qualquer superioridade da leitura distante em relação aos protocolos mais estabelecidos de pesquisa com textos. Pelo contrário, a leitura distante não está em posição de concorrência com a leitura em profundidade, pois ela constrói um novo objeto de pesquisa, que não possui qualquer comensurabilidade com os textos tomados individualmente. O que está em jogo, diante da tarefa de pesquisar *em conjunto* mil, dez mil, cem mil ou mesmo um milhão de textos não é a leitura exaustiva de cada um deles – decerto frustrante e impossível –, mas buscar estratégias para simplificá-los e codificá-los, permitindo o manejo de ferramentas de visualização de dados e de mensuração estatística de seus conteúdos.

Revolução informacional, *web 2.0* e novos modos de auto-observação

Como assinalado por Franco Moretti no livro justamente intitulado *Distant reading* (2013), a leitura à distância, ao selecionar aspectos muito específicos ou codificáveis dos textos – para dar alguns de seus próprios exemplos: tamanho dos títulos, presença de certos elementos estilísticos, frequência de determinados tempos verbais, dentre vários outros –, permite, ao tratar em conjunto estes aspectos selecionados, ir para muito além do próprio texto. Pois, ao tornar manejáveis e mensuráveis as relações formadas entre estes aspectos textuais, cuja seleção e homogeneização reduzem brutalmente a complexidade inerente a cada texto considerado individualmente, torna-se igualmente possível construir um objeto novo de conhecimento que é o sistema de relações formado por esses textos. Numa formulação paradoxal cara à teoria dos sistemas sociais, a redução de complexidade é condição para o aumento da complexidade (Luhmann, 2012). Pois a ultrassimplificação de um texto, de modo a codificá-lo a partir de determinados elementos, é a condição de possibilidade para se analisar *em conjunto* mil, dez mil, cem mil textos, um milhão de textos...

Um exemplo bem assentado deste tipo de procedimento é a análise de cocitação (Gmür, 2006), técnica há muito utilizada nos estudos bibliométricos. Trata-se de identificar as referências bibliográficas presentes em um conjunto selecionado de textos e, a partir daí, criar relações entre os nomes de autores(as) ou das referências citadas conjuntamente em um mesmo texto. Isso criará relações, de ordens crescentes de complexidade de acordo com o volume de informações disponíveis (desde que manejáveis computacionalmente), entre autores(as) ou textos citados conjuntamente a partir da atividade *citante* dos(as) autores(as) dos documentos selecionados na amostra. Com isso, torna-se possível assinalar que os(as) autores(as) A e B, ou as referências X e Y, encontram-se forte ou frouxamente acopladas (ou ainda totalmente dissociadas), o que se revela estratégico para detectar a presença de especializações e/ou cruzamentos disciplinares quando o material em análise são artigos ou livros acadêmicos, por exemplo. É claro que o(a) leitor(a) atento(a) de uma determinada área do conhecimento, ao se familiarizar com a literatura especializada, saberá dizer facilmente quais são os textos incontornáveis e poderá sugerir algumas relações preferenciais entre estes textos no interior de sua especialização. Porém, conforme

discussão em curso sobre balanços bibliográficos, não só o acréscimo da produção científica, mas sobretudo o acesso a escala de análise até pouco tempo atrás inacessível aos produtores desses balanços vem exigindo o uso de novas ferramentas de visualização do conhecimento científico disponível. Cada vez mais há a consciência de que o foco em poucos textos, mesmo que importantes, traz o risco de deixar de fora coisa demais (Botelho et al., 2019).

Não se trata, mais uma vez, de confrontar as virtudes da leitura à distância de um campo científico, para seguir no exemplo da análise de citação, face às eventuais limitações da leitura em profundidade – que, por definição, só pode se concentrar em poucos textos lidos em sua densidade e complexidade próprias. Ambas as formas de leituras são potentes e devem servir como ferramentas indispensáveis. E podem inclusive trabalhar em registro complementar: a partir de critérios propiciados pela leitura distante – exemplos: quais são os textos que se ligam a mais textos, quais conectam diferentes áreas do conhecimento, quais indicam campos emergentes etc. –, podemos *selecionar* os textos que serão lidos em profundidade. Não é outra coisa a importância crescente dos índices ou rankings de citação da produção acadêmica, graças à popularização de bases indexadoras como o Google Scholar: um princípio de seleção (discutível e altamente questionável, vale ressaltar) para tornar manejável o volume impressionante de documentos de pesquisa disponível para qualquer especialização acadêmica hoje (Brasil Jr & Carvalho, 2020).

Aliás, a leitura distante não é somente um novo protocolo possível da pesquisa com textos. Podemos dizer que formas de leitura distante passam cada vez mais a fazer parte do estoque rotineiro de auto-observações recursivamente levado a cabo no processo social. Já há certa literatura disponível que enquadra esse fenômeno nos marcos de um “*knowing capitalism*” – um tipo de economia global que não apenas depende de tecnologias que geram grandes quantidades de dados digitais, mas igualmente torna esses mesmos dados mercadoria (*big data economy*) (Thrift, 2005), com consequências evidentes para a própria pesquisa em sociologia (Savage & Burrows, 2007) –, ou nos termos da generalização de uma espécie de “*audit culture*”, que multiplicaria em todas as organizações formas de avaliação e de mensuração quantitativa de atividades, performances, condutas, com amplos efeitos

potencialmente deletérios em termos da promoção dos valores da autonomia e da democracia (Shore & Wright, 2015). Para continuar no nosso exemplo do campo científico, o volume de produção de artigos atualmente disponível vem exigindo formas de auto-observação que possam lidar com esse universo complexo e potencialmente infinito de informações – e os estudos bibliométricos e cientométricos emergem justamente para lidar com esse problema. Que a forma preferencial de lidar com essa complexidade seja o fator de impacto da produção científica – medida a partir de critérios sempre muito seletivos (e discutíveis) de citações dos artigos – revela, é claro, um poderoso amálgama entre interesses comerciais dos grandes *publishers* e das principais bases indexadoras, que se combina com a difusão das “*audit cultures*” nos meios universitários de todo o mundo – amálgama este que reproduz de modo perverso desigualdades entre instituições, disciplinas, regiões etc. No entanto, não se trata simplesmente de se recusar essas formas de medições; pelo contrário, recusar as formas de leitura distante implicaria renunciar às possibilidades de manejar – por meio da codificação e da ultrassimplificação – o volume de conhecimento científico atualmente disponíveis. Um caminho possível parece estar presente nos movimentos atualmente existentes pelo “acesso aberto”, pelas “citações abertas” e pelos “dados abertos” dos artigos científicos, bem como iniciativas como o “Manifesto de Leiden” (Hicks et al., 2015) por um uso responsável das métricas e dos rankings científicos. Esses movimentos pretendem dar mais autonomia aos próprios cientistas ao democratizarem o acesso aos dados e às ferramentas capazes de gerar uma leitura distante da produção científica existente (e crescente em ritmo acelerado).

A profusão de modos de auto-observação da sociedade por meio da leitura distante está longe de ser privilégio do campo da ciência. Nas interações sociais atualmente remodeladas pelas mídias sociais, por exemplo, popularizaram formas próprias de leitura distante, como os *trending topics* e as *hashtags*, que permitem medir o volume de certos termos e fazer a ligação entre os textos (postagens) que os compartilham. Num contexto de proliferação inaudita de textos que circulam *online*, o uso de uma *hashtag* pode ser decisivo para criar ligações com outros textos – provavelmente escritos por pessoas desconhecidas (e que continuarão a lê-lo) – e,

talvez, alguma ressonância em alguma discussão. Fenômenos aparentados emergem em todos os campos da vida social.

Essa multiplicação de formas de leitura distante, que estão alterando os modos habituais da reflexividade social, revela que os princípios de seleção cultural até pouco tempo atrás “naturalizados” estão cedendo lugar a modos novos de seleção (Acerbi, 2019). A autoridade dos “especialistas” vem sendo deslocada por novos modos de produção, codificação e de seleção das informações, que dependem da agregação de rastros imensos de sinalizações (*likes*, *dislikes*, recomendações...) ou da atuação descentrada de compartilhamentos entre pessoas ou ainda de formas concertadas de operações de *bots* e afins (Bak-Coleman et al., 2021). Não por acaso, vemos a simultaneidade das crises de representação na política, na credibilidade dos veículos de mídia tradicionais, na crença na verdade científica etc. e, com ela, a urgência cobrada para que os “especialistas” se comuniquem de outros modos diante deste novo “público”.

O volume crescente de dados estruturados e semi-estruturados parece, enfim, colocar novos desafios às ciências humanas, de forma geral, e ao pensamento social brasileiro, em particular. Desafios, a bem dizer, teóricos e metodológicos. Afinal, a complexidade das relações internas a esses dados, que demandam forte investimento nas mais variadas técnicas e tecnologias computacionais, trazem informações que não seriam possíveis a partir de outras perspectivas de análise mais canônicas. Tamanha complexidade e variedade no tratamento analítico exige, nos parece, tomar o grande volume de dados formados por diferentes plataformas enquanto formadores de uma “semântica social” cujo alcance, influência e sentido importam estudar para compreender o modo pelo qual a sociedade é dotada de inteligibilidade para grupos sociais e indivíduos.

Quando reunidas em grandes conjuntos, seja formando redes complexas, seja codificadas em metadados, as variadas formas de textos que constituem os nossos projetos de pesquisa podem surpreender pelos novos sentidos que revelam. No que toca especificamente ao pensamento social brasileiro esse é um ponto caro. Afinal, a área acumula discussões muito sofisticadas e trabalhos clássicos sobre as relações entre ideias e vida social ou textos e contextos que nos parecem de especial interesse

quando se analisa as relações entre big data e sociedade. Embora seja inegável que as novas tecnologias comunicacionais, de gestão e armazenamento de dados impliquem formas ainda mais velozes e dinâmicas inéditas de envolvimento dos seus usuários, fato é que continua fundamental compreender o modo como as semânticas textuais se articulam aos usos recursivos realizados por atores individuais e coletivos

São sobretudo esses deslocamentos que permitem a formação de sistemas relativamente autônomos em seus sentidos, cujas lógicas precisamos decifrar. Traduzidas através do uso de estatísticas, gráficos, redes, dendogramas, essas lógicas expressam formas semânticas que circulam amplamente e que são recursivamente utilizadas por atores em diversos contextos. No fundo, o que queremos salientar é que as relações possivelmente existentes entre determinado conjunto de textos e fatores diversos de ordem econômica, política e social não devem ser postuladas *a priori*, sob o risco de reforçarmos reducionismos fáceis e de tomarmos como estabelecido o que é fundamentalmente uma questão empírica. Trata-se, no entanto, de incorporar a contingência como fator constitutivo do próprio enquadramento teórico que pretendemos dar aos casos aqui analisados.

Desse modo, importa acompanhar os movimentos de acoplamento e desacoplamento entre as semânticas textuais e os usos recursivos realizados por atores individuais e coletivos. Justamente por formarem um sistema relativamente autônomo, ainda que em rede, a partir do conjunto de textos é possível revelar padrões de interação formados a despeito das intenções dos atores. E mais: como na própria circulação, essas formas semânticas são ressignificadas, codificadas e decodificadas.

Da autonomia relativa da cultura às redes sócio-semânticas

A área de “pensamento social no Brasil” vem trazendo, ao longo das últimas décadas, por meio de suas reflexões teórico-metodológicas e pesquisas empíricas, um rico debate sobre as relações entre cultura e sociedade. Em termos esquemáticos, podemos dizer que, em alguma medida, os esforços recentes atuam no sentido de deslocar a perspectiva de uma “sociologia da cultura” de inspiração bourdieusiana em direção a uma “sociologia cultural” (não necessariamente modelada a partir de Jeffrey

Alexander), isto é, a uma abordagem que tome a cultura em sua complexidade constitutiva e em suas relações contingentes com o processo social (Brasil Jr, 2020). Na sociologia contemporânea, está emergindo uma discussão a respeito dos possíveis deslocamentos que o *big data* pode (e deve) promover nos modos assentados de análise sociológica da cultura (Lee & Martin, 2015; Mohr et al., 2015, 2020; Reed, 2015). Afinal, não se trata apenas de adequar os nossos recursos intelectuais disponíveis à “revolução informacional” em curso (Kallinikos, 2006b), mas repensá-los criativamente sob pena de não entender o que está realmente em jogo.

Ainda que geralmente mal interpretada, a distinção analítica trazida por Talcott Parsons entre os sistemas cultural e social é decisiva aqui, justamente por colocar a discussão da relação entre estes dois níveis (sistemas) em termos de suas complexidades e contingências próprias (Parsons, 1991). É certo que, quase sempre, Parsons não levou a sua própria inovação às últimas consequências, reduzindo a análise à identificação de formas de tradução da cultura em termos de “valores institucionalizados” no sistema social capazes de resolver os dilemas crônicos da interação social. Jeffrey Alexander, em sua proposta pós-funcionalista, chamou a atenção para essa limitação de Parsons, incorporando a contingência, o conflito e a reflexividade no cerne de sua teorização sobre as relações entre cultura e sociedade, a primeira concebida como analiticamente independente da última (Alexander, 1997).

Em outra chave, também em diálogo heterodoxo com a abordagem parsoniana, Niklas Luhmann concebeu as relações entre cultura (conceito que ele reputava analiticamente pouco produtivo) e sociedade nos termos das relações entre semântica e estrutura social (Luhmann, 1986). Semântica, para Luhmann, refere-se a condensações de sentido que orientam as comunicações (elementos básicos dos sistemas sociais, de acordo com a teoria dos sistemas sociais), tanto em suas formas mais simples ou convencionais (frases feitas, provérbios etc.) quanto em suas formas mais codificadas, como no caso dos “meios de comunicação simbolicamente generalizados” típicos da modernidade (exemplificados nos códigos simbólicos do dinheiro, do amor, do direito, da verdade científica etc.). As semânticas, portanto, são parte constitutiva das operações sociais básicas, posto que selecionam – e, com isso,

condicionam – os temas, os participantes, os conteúdos e as distinções que estruturam as comunicações e suas possibilidades de variação (Luhmann, 1995).

No entanto, para além deste nível mais basal, em que semântica e estrutura aparecem como duas faces da mesma moeda, Luhmann também assinala, por meio da distinção entre operação (o enlace permanente entre as operações do sistema; neste caso, o enlace incessante entre comunicações) e observação (um tipo de comunicação improvável e que depende de certo limiar de complexidade social, que são as autodescrições da sociedade), uma outra forma de relacionar semântica e estrutura social, que nos interessa mais de perto aqui (Holmes, 2018; Luhmann, 2021; Stichweh, 2016; Waizbort, 2020). Esse estoque de observações constitui o momento de reflexividade nas operações dos sistemas sociais; quando as observações se fixam e se acumulam em textos, são chamadas de “autodescrições” – o prefixo “auto” remete ao caráter interno destas formas de (auto-)observação. As autodescrições conformam um acervo de ideias, conceitos e distinções que funcionam como uma espécie de patrimônio coletivo à disposição dos sistemas sociais, que precisam cronicamente se auto-observar a fim de manejar sua própria complexidade interna. Sim, os sistemas sociais, nesta chave teórica, se auto-observam e atualizam (de modo não linear e sempre imprevisível) seu aprendizado social por meio do recurso a textos, que são selecionados no processo de “evolução das ideias” (Luhmann, 2013). Tal “evolução” diz respeito à produção incessante de variações semânticas, processo que se dinamizou fortemente com a invenção da imprensa e, mais recentemente, ganhou tração inaudita com o advento da internet (e da “revolução informacional” que ela propiciou)¹. O aumento da variação disponível implica de modo concomitante a necessidade de mecanismos seletivos capazes de filtrar o que deve ser lido e o que deve ser esquecido, quais distinções devem ser mantidas e quais devem ser descartadas, quais conceitos se “plausibilizam” diante da mudança estrutural e quais perdem contato com o

¹ “In other words, additional facilities for massing and accelerating variation are needed (as where biochemical mutation is complemented in the evolution of life by bisexual reproduction). In societal evolution, this has happened in two ways: through the dissemination medium writing and by enhancing conflict competence and conflict tolerance in society (to put it another way, by renouncing the externalization of all conflicts as typical in segmentary societies)” (Luhmann, 2012, p. 279).

processo social e se tornam obsoletos. Como postula a teoria neodarwiniana da evolução, da qual essa discussão em parte é tributária (Stichweh, 2007, 2017), justamente por serem processos distintos, a variação – a produção incessante de “mutações” dentro do repertório cultural disponível – e a seleção – a fixação de determinado conceito/distinção como instância de auto-observação societal – não se relacionam de modo determinístico, mas contingente². Uma vantagem evidente de se relacionar variação e seleção de modo contingente é justamente colocar em destaque a imensa experimentação existente na produção das formas simbólicas – a maior parte desta produção não encontra maiores ressonâncias no processo social, sendo descartadas ou ignoradas indefinidamente – e suas relações potencialmente “frouxas” ou mesmo não inteiramente “funcionais” para os interesses materiais e simbólicos dos grupos sociais com poder efetivo de condicionar as formas de seleção cultural. Como diria Franco Moretti, igualmente precursor no campo emergente denominado de “evolução cultural” – que procura aplicar o enfoque darwiniano ao estudo da mudança cultural (Mesoudi, 2016) –, ao usar a distinção variação/seleção para o campo das transformações da forma literária:

A sequência de variações e seleção pode realmente funcionar como novo padrão esclarecedor, revelando o momento em que as relações sociais permitem ao sistema literário buscar os seus próprios experimentos livres e, pelo contrário, quando exigem dele uma função mais estrita e bem definida. Para mencionar o primeiro exemplo que vem à mente, um padrão desses pode ser a solução do problema que toda sociologia honesta da literatura teve necessariamente que enfrentar: por que as formas literárias nunca são tão “funcionais” quanto a história social nos leva a esperar? Por que estão sempre meio fora do alvo, desequilibradas, opacas? Mais uma vez virou moda acreditar que isso acontece devido a alguma vocação anárquica intrínseca à literatura, mas prefiro sugerir uma resposta mais prosaica: essa “imperfeição” sociológica se deve ao fato de que as pressões sociais (e daí também as pressões do bloco dominante) só estão ativas em *metade* da história literária. À moda darwinista, o contexto pode selecionar formas, mas não gerá-las. Assim, as formas dominantes, como as ideias dominantes, não são bem as

² “Only through the noncoordination of variation and selection, in other words, by avoiding communication on this relationship, can evolution become sufficiently probable and lead sufficiently rapidly to the development of an order that is improbable. It is in this sense that we can also describe the relationship between variation and selection as chance: the determinacy of variation tells us nothing about the chances of selection” (Luhmann, 2012, p. 279).

formas da classe dominante; são as formas que a classe dominante selecionou – mas sem tê-las produzido. Ninguém é onipotente (Moretti, 2007, p. 313).

As formas pelas quais a sociedade se auto-observa, recorrendo a um repertório de formas simbólicas disponíveis, se ligam de diversas maneiras a um determinado contexto social. Em termos metodológicos, a operacionalização da análise nesse nível em que as semânticas sociais se interrelacionam com as estruturas sociais nem sempre é uma tarefa fácil. Um ramo da análise de redes que vem enfrentando nos últimos anos esse desafio é o das chamadas redes sócio-semânticas, baseadas sobretudo nos trabalhos precursores de Harrison White (Godart & White, 2010; White, 2012) e John Mohr (Mohr, 1998). Seu enfoque recai sobretudo na relação entre padrões de organização de redes formadas por indivíduos e grupos sociais (as chamadas redes sociais) e os significados culturais (símbolos, narrativas, categorias e identidades, ou, numa palavra, redes semânticas) mobilizados nos processos de interação ou de comunicação. Através da análise de redes sócio-semânticas não apenas é possível observar como determinadas formas culturais emergem e se difundem, mas também como o próprio padrão de organização das redes sociais – grupos isolados ou muito conectados, com maior ou menor número de intermediários etc. – operam segundo certas expectativas, identidades e narrativas. Há ainda outro ganho metodológico particularmente importante quando se recorre à análise de redes sócio-semânticas: através dela é possível especificar conexões que não são homogêneas e bem delimitadas (como o conceito de grupo social sugere), dando-se destaque tanto aos laços fracos quanto aos laços mais densos que conectam os nós das redes em comunidades. Um bom balanço a respeito de como as redes sócio-semânticas, ao tomarem a cultura e sociedade como dois níveis analíticos dotados de complexidade próprias, porém interligados, permitem articular em outro patamar as dualidades básicas da teoria social – como estrutura e conteúdo, métodos quantitativos e qualitativos, análise formal e hermenêutica etc. – se encontra nos trabalhos de Nikita Basov, Ronald Breiger, Jan Fuhse, Paul DiMaggio, dentre vários outros(as) (Basov et al., 2020; DiMaggio, 2020; J. Fuhse et al., 2020).

Como já destacado por diversos especialistas, o interesse das ciências sociais pelo chamado *big data* não deve somente ao seu tamanho, mas à sua dimensão relacional. Diferentes tipos de dados estão integrados em grandes bases, gerando, a cada novo *input* e seleção, novas informações. Nessa espécie de reação em cadeia, as formas culturais de que falávamos vão ganhando contorno, tendo seus sentidos iniciais reforçados ou modificados. Esse movimento dinâmico e relacional se amplifica ainda mais com a chamada *Web 2.0*, termo que designa uma ampla rede de comunidades e serviços oferecidos na internet com ativa participação de seus usuários, a exemplo das mídias sociais e plataformas em formato wiki (Sajus et al., 2009). Nela, o material simbólico não é apenas disponibilizado, como na mídia impressa e sites convencionais, mas se constituem em instâncias de negociação dos sentidos envolvidos a partir da intensa participação dos seus usuários (J. A. Fuhse, 2018).

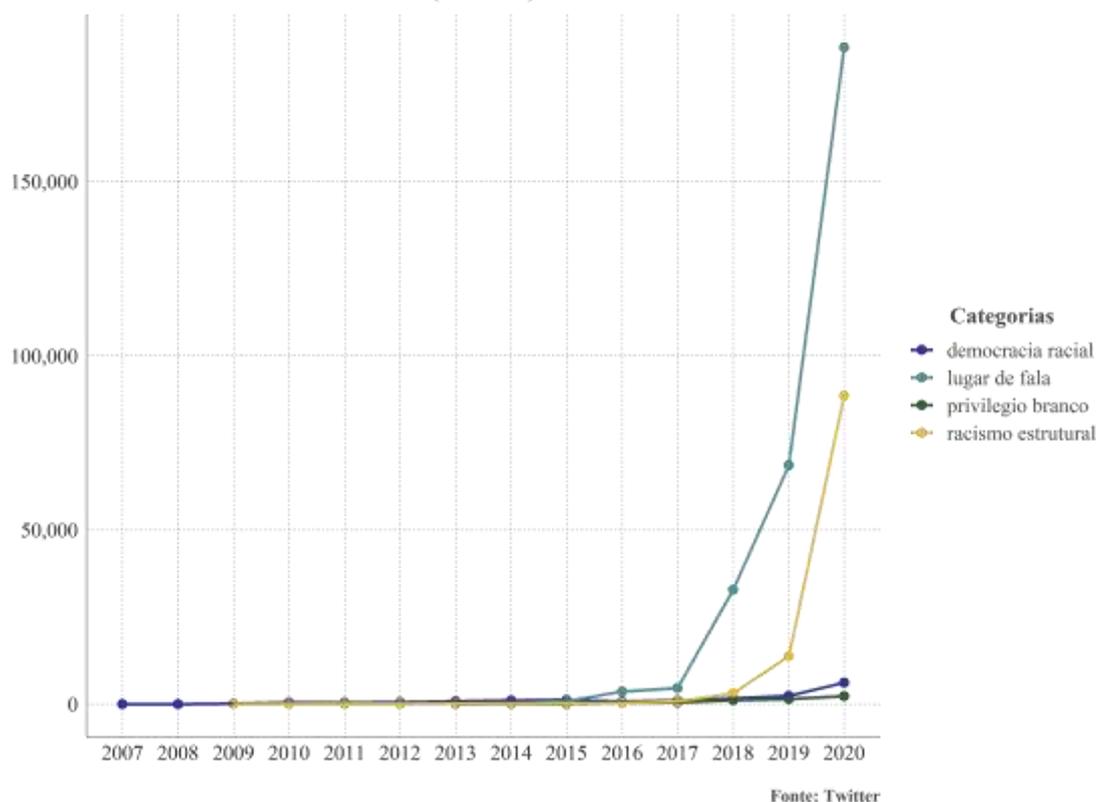
Fato é que, em termos metodológicos e operacionais, as formas culturais que estão disponíveis nessas plataformas (de material e formatos diversos, indo de pequeníssimos a grandes textos, passando por *threads*, *likes*, *dislikes*, recursos visuais como *emojis* e *memes*) são fontes ricas de análise para os cientistas sociais e, em particular, para os pesquisadores da área de “pensamento social no Brasil”, justamente porque capazes de condensar conteúdos e visões da sociedade que rapidamente podem circular e “viralizar”. A compreensão de que as ideias e as formas culturais são forças sociais em disputa pela orientação das condutas individuais e coletivas é uma questão em relação em torno da qual esta área de pesquisa vem se debruçando há décadas e com trabalhos que se tornaram verdadeiros clássicos das ciências sociais brasileiras (Bastos & Botelho, 2010; Schwarcz & Botelho, 2011). O que se abre agora é a possibilidade de acompanharmos de perto e quase simultaneamente a formação e circulação daquelas ideias e o modo como são recursivamente utilizadas por diferentes atores em interação e disputa. Tendo em vista os desafios trazidos pelo *big data* em especial à área de pensamento social, que, em 2020, foi criado o grupo de pesquisa “Pensamento Social em Dados” (*Data_PS*) composto por pesquisadores da Universidade Federal Fluminense (UFF) e Universidade Federal do Rio de Janeiro (UFRJ). Nele estamos trabalhando com dados de diversas plataformas digitais, sobre os quais falaremos a seguir.

Questões de pesquisa e desafios do *big data* à área de pensamento social brasileiro (um exercício de pesquisa)

Os dados que apresentaremos a seguir fazem parte de um projeto em desenvolvimento no *Data_PS* que pretende monitorar a circulação de categorias, temas e autores/as da tradição do pensamento social brasileiro em diversas mídias sociais. Como ponto de partida do projeto, coletamos dados no *Twitter* utilizando como termo de busca “democracia racial”, uma das categorias fundamentais no léxico da área de pensamento social e talvez uma das mais presentes no debate público. Nosso intuito inicial era o acompanhar de que modo essa categoria estaria sendo acionada pelos usuários do *Twitter* na explicação e compreensão de fenômenos sociais contemporâneos. Em uma primeira análise, constatamos que “democracia racial” de fato se fazia presente no debate entre os usuários da plataforma, no entanto, sua compreensão parecia incompleta quando dissociada de outras categorias de uso mais recente. De modo a melhor capturar os sentidos dos usos da categoria “democracia racial”, optamos, em uma segunda etapa, por incluir na coleta de dados os termos “privilégio branco”, “lugar de fala” e “racismo estrutural”. A busca utilizando as quatro categorias retornou 1.151.519 tweets e retweets entre 2007 e 2020. Optamos por retirar os retweets do conjunto de dados, totalizando 487.883 tweets. Isso porque ao empregarmos as técnicas de análise textual, o conjunto de dados com retweets aparecia muito marcado por acontecimentos sobre os quais as categorias que serviram para a nossa busca se debruçaram para compreender ou caracterizar, a exemplo de times de futebol, programas televisivos, artistas e etc.

O gráfico 1 mostra a frequência dos termos ao longo do tempo. Observa-se o forte crescimento recente das categorias “lugar de fala”, “racismo estrutural” e “privilégio branco”, sobretudo a partir de 2017. “Democracia racial” tem tido leve alta a partir de 2017 e alta mais expressiva em 2020. Nossa análise, ainda exploratória, começará com a categoria “democracia racial”.

Frequência das categorias ao longo dos anos
(2007-2020)



Talvez o exemplo mais paradigmático do modo pelo qual as ideias podem se tornar verdadeiras forças sociais é o debate ensejado em torno da categoria “democracia racial”. A ideia de que no Brasil teriam convivido harmônica e sincreticamente diferentes etnias e culturas e, por conseguinte, dirimido divisões raciais ganhou em Gilberto Freyre sua formulação clássica. O autor pernambucano é tido frequentemente como criador do termo, muito embora ele mesmo nunca o tenha usado de forma sistemática. O que importa, no entanto, é o fato de que Freyre é dono de uma formulação intelectual que fincou raízes em diversos campos – inclusive político – e ajudou a fomentar certo imaginário de que na sociedade brasileira não haveria linhas raciais nítidas e tampouco conflitos raciais institucionalizados. No limite, as tensões raciais se resolveriam em ajustes pessoais, familistas – privados, numa palavra. Desde então movimentos sociais negros e diversos trabalhos das ciências sociais institucionalizadas – os mais emblemáticos o de Florestan Fernandes (2008) – chamam a atenção para o fato de que a convivência pacífica entre

raças na sociedade brasileira seria, na verdade, um “mito”, que tornaria muitas vezes os conflitos velados, embora não menos cruéis.

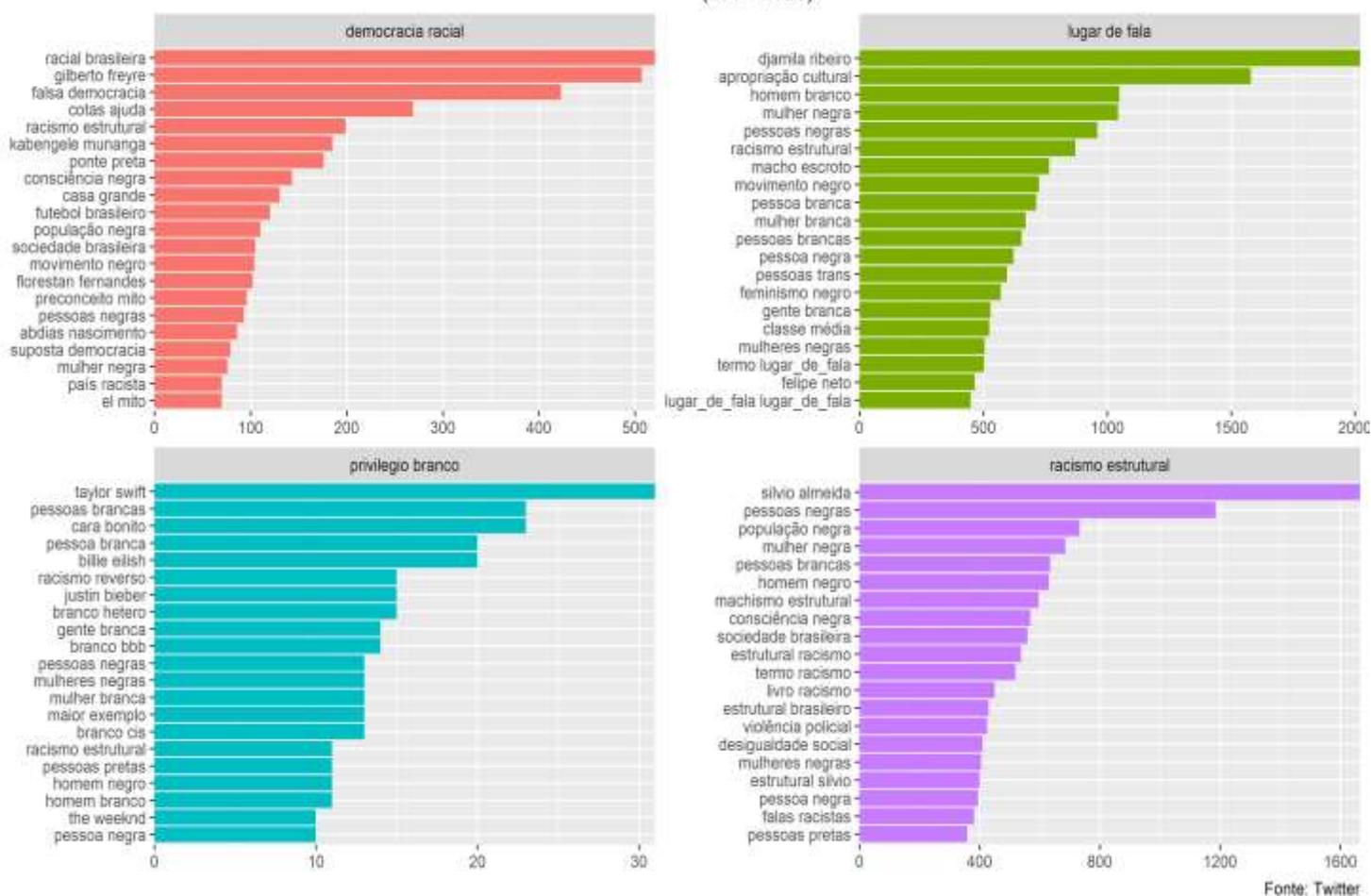
Não se trata aqui de retomarmos as discussões sobre a categoria “democracia racial” no pensamento social, cujas raízes podem inclusive ser remontadas a período anterior à obra de Gilberto Freyre, mas de destacar como em anos recentes foi se constituindo em torno dela uma semântica que tem atualizado seu caráter de denúncia das desigualdades raciais e também ressignificado os seus sentidos. Chama a atenção inclusive que o uso da categoria, embora persistente (em volumes relativamente baixos) desde 2007, cresce nos últimos anos, o que parece contraintuitivo, uma vez que outros termos – como “racismo estrutural”, por exemplo – parecem remeter melhor ao arco de problematizações da desigualdade racial no país. No entanto, como muitos já salientaram, a eficácia do “mito” depende justamente de sua natureza percebida como autoevidente, e não de sua necessária articulação discursiva. Durante muito tempo atuou entre nós uma espécie de “pacto de silêncio” em torno da “democracia racial brasileira”, que deslegitimou a discussão pública e franca da questão racial no Brasil como uma “não-questão”, já que supostamente não haveria problemas a resolver numa sociedade que se autocompreende como “racialmente democrática”. Portanto, a subida recente do termo, em vez de indicar a força da tese da “democracia racial” no debate público, parece justamente revelar a tendência contrária, o que análise dos “bigramas” mais frequentes no interior do *corpus* de tweets que mencionam o termo vai demonstrar mais adiante.

Por óbvio, a discussão crítica em torno da tese da “democracia racial” não implica afirmar sem mais que hoje em dia é hegemônica a autocompreensão da sociedade brasileira como racista e estruturalmente desigual. Trata-se, aliás, de uma interessante agenda de pesquisa rastrear a semântica conservadora que está se constituindo quando a tese da “democracia racial” vai se desencantando (perdendo a eficácia enquanto “mito”), isto é, deixando de ser autoevidente aos atores sociais. Como os processos sociais de mudança (inclusive no plano da mudança cultural) jamais são lineares e progressivos, podemos acompanhar formas de acomodação entre a atual percepção de que a sociedade brasileira é racista e a negação simultânea de que práticas racistas estejam na base das condutas dos agentes sociais. Como já apontou a

antropóloga Lilia Schwarcz, tendemos a nos conceber como uma “ilha de democracia racial”, cercado de racistas por todos os lados” (Schwarcz, 2007, p. 12). Haveria racismo no Brasil, mas ele é um problema do *outro*, o que dificulta, mais uma vez, a discussão pública e franca do tema.

Os dados sugerem que a categoria, portanto, é acionada sobretudo por parte daqueles que estão criticando – e denunciando – a imagem das relações raciais a ela associada. É exatamente isso que o gráfico 2 aponta através da frequência de “bigramas” (dupla de termos que coocorrem em determinado corpus textual). No caso dos tweets que mencionam a categoria “democracia racial”, destacam-se os bigramas “democracia racial brasileira”, “falsa democracia”, “suposta democracia”, “preconceito mito”, denotando exatamente o caráter “falso” de uma convivência racial igualitária no país. Em polo complementar, a recorrência de bigramas como “consciência negra”, “população negra”, “movimento negro” apontam para agentes envolvidos no esforço em descortinar essa realidade que o “mito” pretende acobertar. Encontram-se ainda, nesse conjunto, termos já tradicionalmente presentes no campo semântico da categoria, como “Gilberto Freyre”, “Abdias do Nascimento”, “Florestan Fernandes” e “mito”, e outros novos, em referência a acontecimentos relacionados ao futebol e a cotas. Nesse último caso, é revelador o modo como a categoria “democracia racial” é acionada pelos usuários do Twitter, já que pretende criticar a base do argumento daqueles que seriam contrários às cotas, principalmente o de que elas seriam potenciais causadoras de conflitos raciais antes tidos como inexistentes.

Frequência de bigramas por categoria (2007-2020)



Fonte: Twitter

Merece especial atenção a categoria “racismo estrutural”, muito citada entre os bigramas associados à “democracia racial”. Ao analisarmos esses dados, constatamos que, para a compreensão do debate recente em torno da categoria “democracia racial”, seria inescapável recorrer a outras categorias. A lista de bigramas associados à “democracia racial” é extensa, no entanto, três deles nos chamaram a atenção embora não apareçam entre os mais citados, sobretudo porque revelam a importância das ciências sociais na constituição do debate sobre questão racial. São eles: “racismo estrutural”, “privilégio branco” e “lugar de fala”. A segunda rodada de extração de dados foi justamente, como já ressaltado acima, dedicada a coletar tweets que mencionam ao menos uma dessas três categorias. Com os bigramas de todas as quatro categorias coletadas, fica bem evidente a importância que “racismo estrutural” adquire no debate recente sobre questão racial. Na tabela 1, temos os dados relativos aos tweets

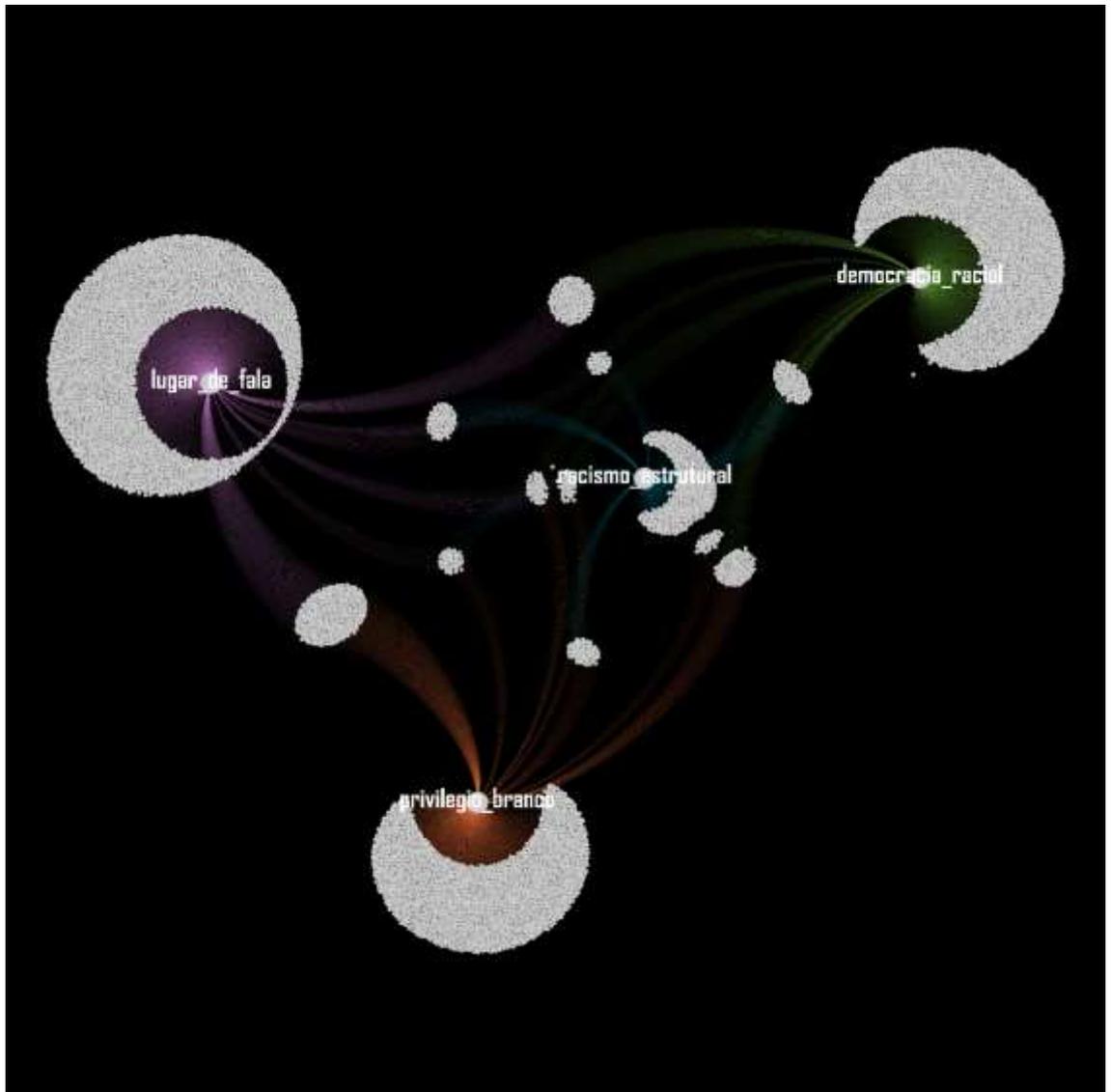
coletados – sem retweets – e as categorias citadas. Nota-se como os tweets que mencionam “democracia racial” (10569) são aqueles que mais mencionam outras categorias, seja contendo ao menos duas categorias (4944), seja ainda contendo três categorias (2389) ou todas as categorias (877). “Privilégio branco” aparece em segundo lugar entre as que mais citam outras categorias, seguida por “racismo estrutural” e “lugar de fala”.

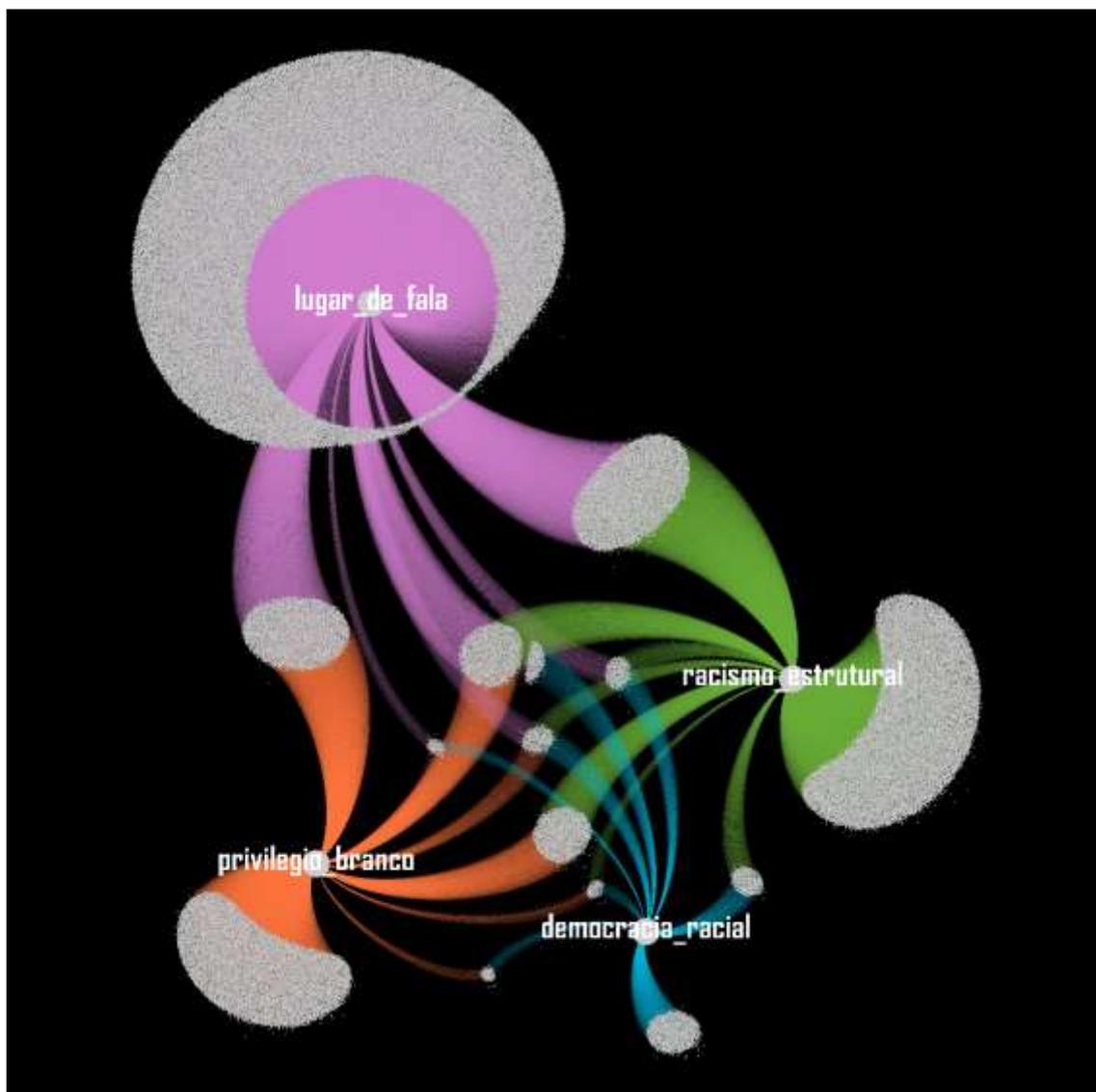
Citações das categorias (2007-2020)								
Usuários	Democracia racial	Democracia racial (%)	Lugar de fala	Lugar de fala (%)	Privilégio branco	Privilégio branco (%)	Racismo estrutural	Racismo estrutural (%)
Citam a categoria	10569	100.0%	174253	100.0%	45020	100.0%	64176	100.0%
Citam pelo menos duas categorias	4944	46.8%	30607	17.6%	19227	42.7%	24778	38.6%
Citam pelo menos três categorias	2389	22.6%	6164	3.5%	5577	12.4%	6142	9.6%
Citam todas as categorias	877	8.3%	877	0.5%	877	1.9%	877	1.4%

Fonte: Twitter

De modo complementar à análise dos bigramas, mas trabalhando agora com os dados dos usuários, vemos igualmente uma transformação no nexos semântico destas quatro categorias e suas relações recíprocas. Modelando uma rede sócio-semântica – uma rede que liga atores sociais e categorias culturais –, podemos ver como os usuários da plataforma fazem uso, em seus tweets, desses termos. Lembrando aqui que as redes sócio-semânticas têm a vantagem de capturar as dimensões sociais e culturais de maneira interligada, porém sem reduzi-las uma à outra. Com isso, podemos ver como a movimentação dos usuários redistribuem a posição relacional das categorias, suas proximidades e distâncias, de acordo com os acoplamentos (e desacoplamentos) semânticos operados ao longo do tempo. Modelamos duas redes sócio-semânticas de dois modos (*2-mode* ou rede bipartida) (*Graph, bipartite - Encyclopedia of Mathematics*, [s.d.]) – uma que só admite relações entre dois tipos distintos de elementos; no nosso caso, usuários e categorias (as relações são os textos dos tweets que contém alguma das quatro categorias). A imagem 1 se refere ao período de tempo 2007 a 2017 e a imagem 2 ao período 2018 a 2020, uma vez que a explosão de

ocorrências se dá exatamente nos últimos três anos. Os usuários estão identificados apenas por suas *ids* e estão na cor cinza – daí que os nomes dos usuários não estejam visíveis nos redes a seguir.

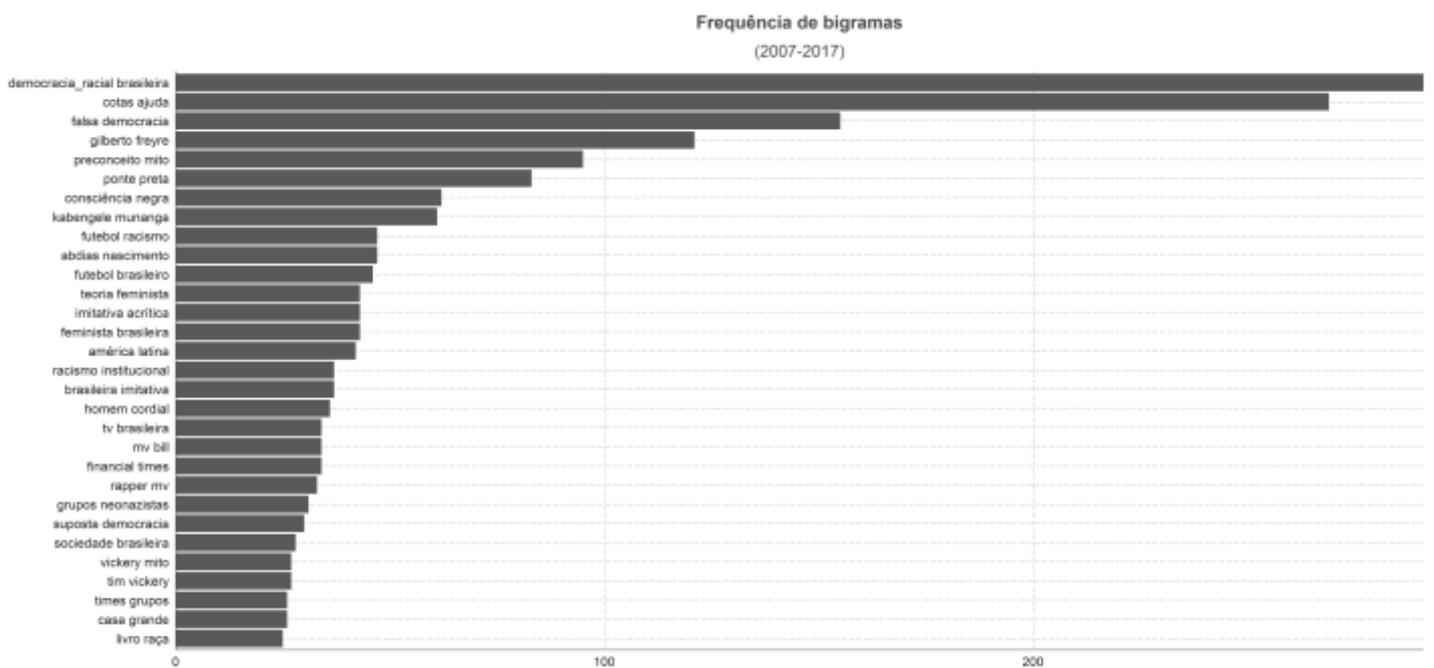




A espessura das arestas – as relações – é distinta nas duas redes em função do volume de tweets, que é muito desigual – de fato, a quantidade de menções aos quatro termos aumenta muito visivelmente depois de 2017 (lembramos que as relações estão definidas pelo número de textos que ligam usuários às categorias). Mas vejamos como os termos de maior menção, que organizam a estrutura da rede, triangulando-a, mudam nos dois períodos. Em 2007-2017, o número de usuários que usam unicamente os termos “democracia racial”, “lugar de fala” e “privilegio branco” se distribui de modo relativamente equilibrado; com menor volume, “racismo estrutural” se encontra na região central da rede, uma vez que os usuários que usaram o termo neste período usaram com relativa frequência também os demais termos. Já no período 2018-2020, a menção às demais categorias sobe muito mais rapidamente que “democracia racial”,

que passa a se localizar intermediando “privilégio branco” e “racismo estrutural”, revelando o nexos semântico que passou a ressignificar o seu significado cultural. De certo modo, “democracia racial” e “racismo estrutural” trocam de posição estrutural em suas relações sócio-semânticas, o último em parte subsumindo o contexto de sentido da primeira.

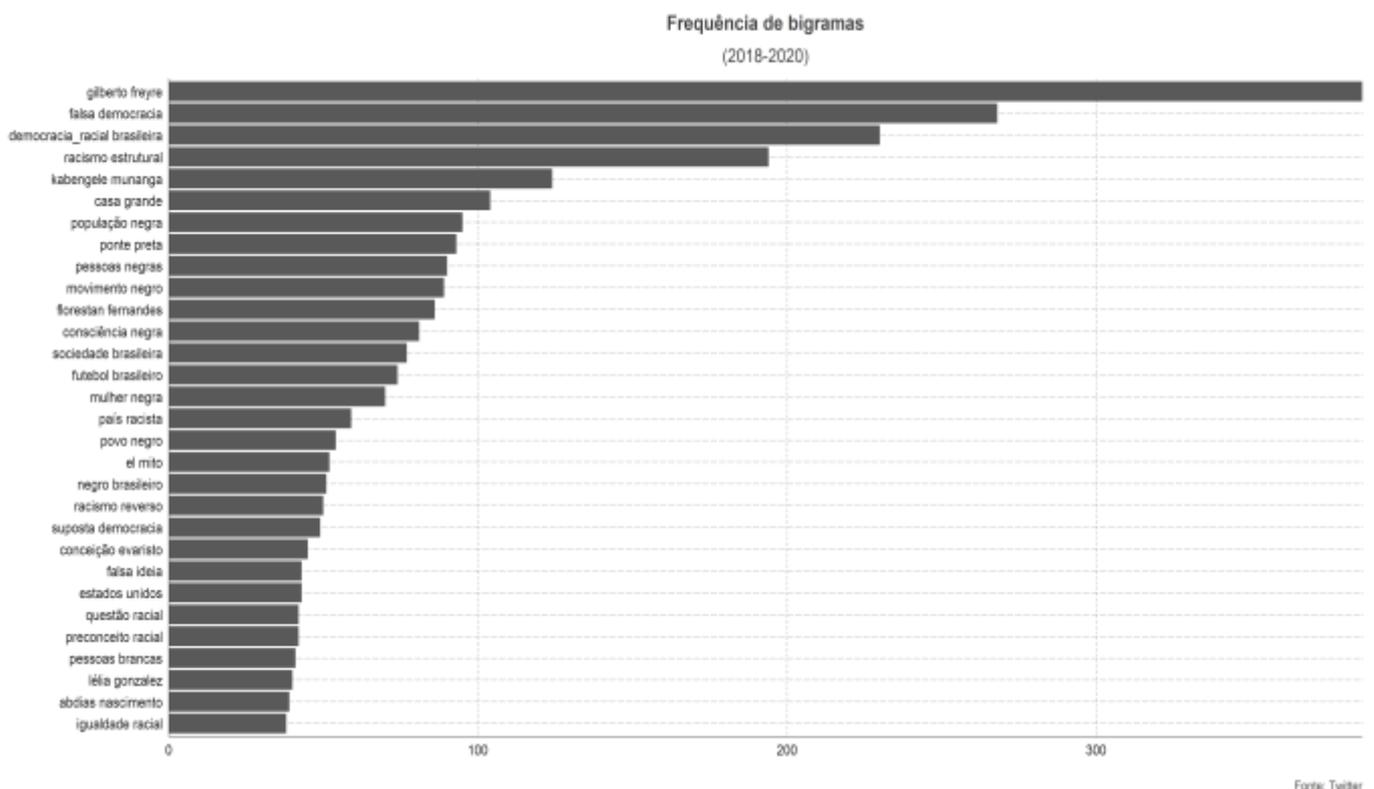
Repartindo os bigramas por faixas de tempo, também é possível observar mais detidamente o papel que a categoria “racismo estrutural” tem na reorganização dos discursos sobre a questão racial. No gráfico 3, referente aos bigramas mais frequentes presentes nos tweets que citam “democracia racial” entre 2007 e 2017, observa-se termos em alusão a acontecimentos que desencadearam o uso da categoria, como programas de tv, futebol e notícias, mas também debates mais amplos sobre teorias feministas. Destaca-se nesse gráfico o bigrama “racismo institucional”, que, aparentemente, tem conotação muito próxima a “racismo estrutural”.



Fonte: Twitter

No gráfico 4, referente aos bigramas mais frequentes presentes nos tweets que citam “democracia racial” entre 2018 e 2020, é possível perceber como “racismo estrutural” consegue capitanear o debate sobre questão racial no Twitter que, até

então, se dava de forma um tanto quanto fragmentada, embora potente e já apontando para o racismo como marca da sociedade brasileira. Importante destacar como “racismo estrutural” vem ganhando o debate público sobretudo a partir do livro de autoria de Silvio Almeida em 2019 e que traz a categoria no título (Almeida, 2019). O livro, por sua vez, faz parte da coleção intitulada “Feminismos Plurais”, coordenada por Djamila Ribeiro, autora de outro grande sucesso na mesma coleção intitulado *Lugar de fala*, também lançado em 2019.



A chamada “revolução informacional” (Alaimo & Kallinikos, 2017; Kallinikos, 2006a), que vem remodelando as ordens interacionais e produzindo dados relacionais sobre a vida social em escala até pouco tempo atrás inacessível (Housley, 2021; Housley et al., 2014), permite desdobrar uma agenda muito cara à área de “pensamento social no Brasil”: quais são os efeitos das ideias no processo social? É claro que a recorrência de certos termos em uma plataforma de *microblogging* como o Twitter, termos que condensam e codificam certas interpretações da sociedade brasileira, depende de um processo mais amplo e de múltiplas mediações, de vão desde

a produção e a seleção de ideias por parte de intelectuais e cientistas, suas formas de comunicação, bem como as suas possibilidades de difusão e rotinização em produtos culturais de ampla difusão (livros escolares, programas audiovisuais, canções e, mais recentemente, postagens em mídias sociais). No entanto, o volume, a distribuição temporal e os nexos semânticos (como as análises das redes sócio-semânticas e dos bigramas ajudam a identificar por meio de uma “leitura distante”) que alguns termos conhecem no Twitter – ou na Wikipédia, no YouTube, nas mídias impressas e digitais, para darmos alguns exemplos de outras bases de dados exploráveis por esta agenda – são bons indicadores de como algumas categorias culturais circulam, conhecem significado cultural e, portanto, produzem efeitos sociais e políticos os mais variados. Em termos metodológicos, os desafios postos pelo *big data* também não são nada desprezíveis, como discutimos nas primeiras partes do texto. Afinal, lidamos agora com “textos” em que os “contextos” se constituem em escala e dinamicidade que demandam monitoramento constante dos dados processados e que nem sempre os sentidos dos discursos mobilizados são colocados de maneira evidente. Acompanhar a produção desses discursos ao longo do tempo é também uma atividade fundamental para os estudiosos do pensamento social, não só para capturar as mudanças das categorias que circulam no debate público, mas também para compreender como elas se constituem em uma espécie de repertório cognitivo que dão sentido a acontecimentos do passado e do presente (Botelho, 2010). Diante das novas possibilidades trazidas pelo *big data*, é fundamental avançar em frentes de pesquisa que identifiquem de que forma diferentes circuitos de ideias se entrelaçam ao compartilhamento, discussão e controvérsias intelectuais e também aos debates levantados por grupos e movimentos sociais (Keuchenius & Mügge, 2021). O exercício aqui feito com o termo “democracia racial”, que será aprofundado em publicações futuras, pretendeu mostrar a produtividade deste tipo de investigação.

Referências

- Acerbi, A. (2019). *Cultural Evolution in the Digital Age*. Oxford University Press.
- Alaimo, C., & Kallinikos, J. (2017). Computing the everyday: Social media as data platforms. *The Information Society*, 33(4), 175–191.
<https://doi.org/10.1080/01972243.2017.1318327>
- Alexander, J. C. (1997). *Neofunctionalism and After*. John Wiley & Sons.
- Almeida, S. (2019). *Racismo Estrutural* (1ª edição). Editora Jandaíra.
- Bak-Coleman, J. B., Alfano, M., Barfuss, W., Bergstrom, C. T., Centeno, M. A., Couzin, I. D., Donges, J. F., Galesic, M., Gersick, A. S., Jacquet, J., Kao, A. B., Moran, R. E., Romanczuk, P., Rubenstein, D. I., Tombak, K. J., Bavel, J. J. V., & Weber, E. U. (2021). Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27).
<https://doi.org/10.1073/pnas.2025764118>
- Basov, N., Breiger, R., & Hellsten, I. (2020). Socio-semantic and other dualities. *Poetics*, 78, 101433. <https://doi.org/10.1016/j.poetic.2020.101433>
- Bastos, E. R., & Botelho, A. (2010). Horizontes das Ciências Sociais: Pensamento social brasileiro. In *Horizontes das ciências sociais no Brasil: Sociologia*. ANPOCS.
- Botelho, A. (2010). Passado e futuro das interpretações do país. *Tempo Social*, 22(1), 44–66.
- Botelho, A. (2015). Un programa fuerte para el pensamiento social brasileño. *Prismas: revista de historia intelectual*, 19, 151–162.

- Botelho, A., Brasil Jr., A., & Hoelz, M. (2019). Tão longe, tão perto: Sociologia & antropologia no limiar de uma década. *Sociologia & Antropologia*, 9(3), 717-739. <https://doi.org/10.1590/2238-38752019v931>
- Brasil Jr, A. (2020). Intelectuais, cultura e poder: Uma agenda brasileira de pesquisa. In J. A. de Lima & R. de C. Fazzi (Orgs.), *Campos das ciências sociais* (1ª Edição). Editora Vozes.
- Brasil Jr., A., & Carvalho, L. (2018). Biblioteca Virtual do Pensamento Social: Conhecimento, democratização e reflexividade das interpretações do Brasil. *Parcerias Estratégicas*, 22(45), 143-155.
- Brasil Jr, A., & Carvalho, L. C. (2020). O impacto da sociologia: Cultura de citações e modelos científicos / The impact of sociology: citation culture and scientific models. *Revista Brasileira de Sociologia - RBS*, 8(20), 248-269. <https://doi.org/10.20336/rbs.700>
- Carvalho, L., & Brasil Jr, A. (2020). Mapeando a área de pensamento social no Brasil: Uma análise preliminar de sua produção em artigos. *Revista Eletrônica de Comunicação, Informação e Inovação em Saúde, no prelo*.
- DiMaggio, P. (2020). Duality and relationality: The cultural matrix of John W. Mohr. *Poetics*, 78, 101438. <https://doi.org/10.1016/j.poetic.2020.101438>
- Fernandes, F. (2008). *A integração do negro na sociedade de classes, vol. 1*. Editora Globo.
- Fuhse, J. A. (2018). New Media and Socio-Cultural Formations. *Cybernetics & Human Knowing*, 25(4), 73-96.

- Fuhse, J., Stuhler, O., Riebling, J., & Martin, J. L. (2020). Relating social and symbolic relations in quantitative text analysis. A study of parliamentary discourse in the Weimar Republic. *Poetics*, 78, 101363. <https://doi.org/10.1016/j.poetic.2019.04.004>
- Gmür, M. (2006). Co-citation analysis and the search for invisible colleges: A methodological evaluation. *Scientometrics*, 57(1), 27–57. <https://doi.org/10.1023/a:1023619503005>
- Godart, F. C., & White, H. C. (2010). Switchings under uncertainty: The coming and becoming of meanings. *Poetics*, 38(6), 567–586. <https://doi.org/10.1016/j.poetic.2010.09.003>
- Graph, bipartite—Encyclopedia of Mathematics*. ([s.d.]). Recuperado 16 de abril de 2021, de https://encyclopediaofmath.org/index.php?title=Graph,_bipartite
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431. <https://doi.org/10.1038/520429a>
- Holmes, P. (2018). A semântica da complexidade e sua estrutura social: O materialismo “pós-estruturalista” da teoria dos sistemas. *Sociologias*, 20(47), 338–368.
- Housley, W. (2021). *Society in the Digital Age: An Interactionist Perspective* (1ª edição). Sage Publications Ltd.
- Housley, W., Procter, R., Edwards, A., Burnap, P., Williams, M., Sloan, L., Rana, O., Morgan, J., Voss, A., & Greenhill, A. (2014). Big and broad social data and the sociological imagination: A collaborative response. *Big Data & Society*, 1(2), 205395171454513. <https://doi.org/10.1177/2053951714545135>

- Josiowicz, A., & Brasil Jr, A. (2019). Pensamento social e pesquisa informacional: O caso da Biblioteca Virtual do Pensamento Social (BVPS) | Social Thought and Informational Research: the case of the Biblioteca Virtual do Pensamento Social (BVPS). *Revista Brasileira de Sociologia - RBS*, 7(16), Article 16. <https://doi.org/10.20336/rbs.459>
- Kallinikos, J. (2006a). *The consequences of information: Institutional implications of technological change*. Edward Elgar. <http://www.loc.gov/catdir/toc/ecip0613/2006014313.html>
- Kallinikos, J. (2006b). Information out of information: On the self-referential dynamics of information growth. *Information Technology & People*, 19(1), 98–115. <https://doi.org/10.1108/09593840610649989>
- Keucheniuss, A., & Mügge, L. (2021). Intersectionality on the go: The diffusion of Black feminist knowledge across disciplinary and geographical borders. *The British Journal of Sociology*, 72(2), 360–378. <https://doi.org/10.1111/1468-4446.12816>
- Lee, M., & Martin, J. L. (2015). Coding, counting and cultural cartography. *American Journal of Cultural Sociology*, 3(1), 1–33. <https://doi.org/10.1057/ajcs.2014.13>
- Luhmann, N. (1986). *Love as Passion: The Codification of Intimacy*. Harvard University Press.
- Luhmann, N. (1995). *Social Systems*. Stanford University Press.
- Luhmann, N. (2012). *Theory of Society, Volume 1* (R. Barrett, Trad.). Stanford University Press.
- Luhmann, N. (2013). *Theory of Society, Volume 2*. Stanford University Press.

- Luhmann, N. (2021). *Teoria dos sistemas na prática vol. III: História, semântica e sociedade*. Editora Vozes.
- Mesoudi, A. (2016). Cultural Evolution: A Review of Theory, Findings and Controversies. *Evolutionary Biology*, 43(4), 481–497. <https://doi.org/10.1007/s11692-015-9320-0>
- Mohr, J. W. (1998). Measuring Meaning Structures. *Annual Review of Sociology*, 24(1), 345–370. <https://doi.org/10.1146/annurev.soc.24.1.345>
- Mohr, J. W., Bail, C. A., Frye, M., Lena, J. C., Lizardo, O., McDonnell, T. E., Mische, A., Tavory, I., & Wherry, F. F. (2020). *Measuring Culture*. Columbia University Press.
- Mohr, J. W., Wagner-Pacifici, R., & Breiger, R. L. (2015). *Toward a computational hermeneutics*. 2(2). <https://doi.org/10.1177/2053951715613809>
- Moretti, F. (2007). *Signos e estilos da modernidade: Ensaio sobre a sociologia das formas literárias*. Civilização Brasileira.
- Moretti, F. (2013). *Distant Reading* (1ª Edição). Verso.
- Parsons, T. (1991). *The Social System*. Routledge.
- Reed, I. A. (2015). Counting, interpreting and their potential interrelation in the human sciences. *American Journal of Cultural Sociology*, 3(3), 353–364. <https://doi.org/10.1057/ajcs.2015.7>
- Sajus, B., Cardon, D., Levrel, J., Vatant, B., Bermès, E., Oury, C., & Sussan, R. (2009). Web 2.0, et après ? *Documentaliste-Sciences de l'Information*, Vol. 46(1), 54–66.

- Savage, M., & Burrows, R. (2007). The Coming Crisis of Empirical Sociology. *Sociology*, 41(5), 885–899. <https://doi.org/10.1177/0038038507080443>
- Schwarcz, L. M. (2007). Lilia Schwarcz: Quase pretos, quase brancos. *Pesquisa Fapesp*, 137, 10–15.
- Schwarcz, L. M., & Botelho, A. (2011). Simpósio: Cinco questões sobre o pensamento social brasileiro. *Lua Nova: Revista de Cultura e Política*, 82, 139–159.
- Shore, C., & Wright, S. (2015). Audit Culture Revisited: Rankings, Ratings, and the Reassembling of Society. *Current Anthropology*, 56(3), 421–444. <https://doi.org/10.1086/681534>
- Stichweh, R. (2007). Evolutionary Theory and the Theory of World Society. *Soziale Systeme*, 13(1–2), 528–542. <https://doi.org/10.1515/sosys-2007-1-245>
- Stichweh, R. (2016). Estructura social y semántica: La lógica de una distinción sistémica. *MAD*, 35, 1–14. <https://doi.org/10.5354/rmad.voi35.42794>
- Stichweh, R. (2017). Evolution and Control of Society. *NAL NF Nr.* 419, 109.
- Waizbort, L. (2020). Cada um por si e nenhum por todos. Elementos da Covariação nos domínios de estrutura social e semântica. *Sociologia & Antropologia*, 10, 423–448.
- White, H. C. (2012). *Identity and Control: How Social Formations Emerge - Second Edition*. Princeton University Press.