

COMO COMBATER O DISCURSO DANOSO: O USO DE REFORMULAÇÕES COMO CONTRADISCURSO*

How to Combat Harmful Speech: The Use of Rephrasing as Counterspeech

EDUARDA CALADO BARBOSA ^a

<https://orcid.org/0000-0002-2478-9670>

ecalado@unicamp.br

^a Universidade Estadual de Campinas, Campinas, Sao Paulo, Brasil.

Resumo

É amplamente aceito que o discurso intolerante e preconceituoso não tem lugar na comunicação democrática. Mas, como combatê-lo? Uma resposta natural é recorrer ao contradiscurso negativo, no qual o ouvinte resistente refuta o que foi dito pelo preconceituoso; e.g., negando um estereótipo negativo, como ‘mulheres são submissas’. Ultimamente, contudo, este tipo de estratégia vem sendo tratada com certo pessimismo por parte de filósofos que ressaltam que negar estereótipos reforça sua saliência e amplifica sua influência cognitiva. Isso pode fazer com que tentativas de refutação simplesmente saiam pela culatra como contradiscurso. Além disso, também já se argumentou, contra as refutações, que é injusto exigir deveres de resistência das vítimas típicas do discurso intolerante, já que sua fala costuma sofrer com silenciamentos e/ou lacunas hermenêuticas. Meu objetivo aqui é questionar essa visão pessimista sobre o contradiscurso negativo, argumentando que responder negativamente ao intolerante continua sendo uma opção adequada, desde que o falante resistente consiga mudar a QUD (*Question Under Discussion*) — que representa os objetivos da conversa — ou os termos problemáticos usados na interação. Proponho que as reformulações *in situ*, que são respostas essencialmente negativas, constituem boas alternativas às refutações. Sustento também que as reformulações realizadas na comunicação do Estado têm menos probabilidade de serem silenciadas, o faz delas ferramentas úteis para evitar problemas relacionados a questões deontológicas. Por fim, analiso um caso prático de uso de reformulação no âmbito do contradiscurso estatal.

Palavras-chave: Discurso danoso; Contradiscurso; Silenciamento; Reformulação; QUD.

Abstract

It is widely accepted that bigoted and prejudiced speech has no place in democratic communication. But how can it be fought? One response resorts to negative counterspeech,

* A autora agradece ao público do VII Congresso da ALFAN, realizado em Montevidéu, pelas contribuições e comentários recebidos, assim como ao apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo n° 2024/05386-9.

in which the resistant hearer typically denies what was said by the bigot; e.g., denying a harmful negative stereotype, like ‘women are submissive’. Lately, however, this type of negative strategy has been treated with some pessimism by philosophers who point out that denying stereotypes tends to reinforce their salience and to amplify their cognitive influence. This can cause refutations to backfire as counterspeech. It has also been argued, more broadly, that it is unfair to demand resistance duties from typical victims of bigotry, since their speech is often silenced and/or stumble upon hermeneutical gaps. My goal here is to question the pessimistic view of negative counterspeech by arguing that it is still an adequate choice in combating bigoted speech, provided the reply either changes the QUD (i.e., question under discussion) — which represents the goals of the conversation — or the terms of the conversation. I propose that rephrasings, which are negative replies in nature, are good alternatives to *in situ* denials. I also hold that rephrasings used in the State’s communication are less likely to be silenced, and this helps to avoid problems related to deontic concerns. Finally, I analyze a practical case of use of rephrasing in state counterspeech.

Key words: Harmful Speech; Counterspeech; Silencing; Rephrasing; QUD.

Introdução

Embora não haja um consenso acerca de como ou mesmo de se é possível conceitualizar inequivocamente o discurso de ódio¹, ele é comumente definido como discurso que discrimina, inferioriza, menospreza, nega a dignidade e incita ao ódio e à violência contra grupos identificados por categorias como raça, gênero, etnia, crença religiosa, etc., criando divisões sociais e ameaçando ideais de inclusão e participação democrática.² Alguns de seus usos mais paradigmáticos envolvem “palavras que ferem” (Matsuda et al., 2018), como insultos de grupo (ou *slurs*) para expressar atitudes e sentimentos negativos contra um grupo-alvo ou o endossamento de estereótipos perniciosos que os associam a propriedades negativas (Orlando & Saab, 2019; Barbosa, 2023; Barbosa & Fernandes, 2024). Usos públicos de frases como (1)-(3) abaixo são bons exemplos:

- (1) Mulheres são submissas.
- (2) Tutsis são baratas!

¹ Para Brown (2017a, 2017b), por exemplo, *discurso de ódio* é uma espécie de conceito guarda-chuva que engloba um “conjunto heterogêneo de fenômenos expressivos” e não um conceito unitário.

² Essa conceitualização é representativa das definições encontradas, por exemplo, no site das Nações Unidas: <https://www.un.org/en/hate-speech/understanding-hate-speech/what-is-hate-speech>; e do Conselho Europeu: <https://www.coe.int/en/web/combating-hate-speech/what-is-hate-speech-and-why-is-it-a-problem->. Data de acesso: 30 de julho de 2025.

(3) Essa sapatão não deveria estar aqui entre gente de bem.

Grande parte do debate filosófico sobre casos como (1)-(3), no entanto, gira em torno não apenas dos tipos de palavras usadas e das atitudes veiculadas, mas do tipo de *dano* que o discurso de ódio constitui e causa.³ Filósofas feministas da linguagem, como Rae Langton (1993), Ishani Maitra (2012) e Mary Kate McGowan (2019), argumentam que o discurso de ódio é constitutivamente danoso porque envolve atos de fala que subordinam e oprimem grupos-alvo ao classificá-los como inferiores, legitimar sua discriminação e privá-los de poderes importantes. Para McGowan, em particular, o discurso de ódio implementa normas opressivas em contextos conversacionais por meio de atos de fala paralelos. Outros, como Katharine Gelber (2021), consideram que o mal do discurso de ódio está em seu poder de excluir grupos-alvo da deliberação democrática. Por fim, dignitaristas, como Jeremy Waldron (2012), assumem que o discurso de ódio é danoso por comprometer a confiança dos indivíduos alvejados na proteção de sua *dignidade cívica*, causando exclusão, dentre outros males.⁴

Além disso, um ponto central do debate é se o discurso de ódio, entendido como discurso danoso, pode ser combatido simplesmente por meio de “mais discurso” — isto é, de *contradiscorso* — ou se o Estado precisa intervir com leis que sancionam agentes (e instituições ou plataformas⁵) por casos como (1)-(3). Para aqueles que favorecem o contradiscorso, uma tarefa importante é apontar que tipo de estratégia deve ser empregada: *in situ* ou estatal, preventiva ou reativa, positiva ou negativa? Minha intenção aqui é propriamente tratar dessa última questão⁶ tendo dois critérios como norteadores: um que diz respeito à eficácia das alternativas — i.e., quais mecanismos semântico-pragmáticos e táticas retóricas têm mais chances de sucesso em neutralizar e prevenir danos discursivos como resposta ao falante odioso — e outro que se volta para seus custos morais. Seguindo

³ Pode-se dizer que o discurso de ódio é constitutivamente danoso no sentido de que *o próprio ato de realizar uma fala odiosa* é danoso (ao silenciar, subordinar, discriminar, oprimir etc); e é causalmente danoso por gerar danos *em consequência* da sua realização (e.g., como afetar a auto-estima de uma vítima ou gerar medo de violência) (Gelber, 2021).

⁴ Outras posições destacam a relação entre o discurso de ódio e a materilização de condutas condenáveis dentro e fora do próprio discurso (e.g., o genocídio), como a de Lynne Tirrell (2012; 2017). Outras ainda, como Marques (2023), defendem que o discurso de ódio se caracteriza essencialmente por expressar ódio — ao pressupor que é adequado tratar membros de grupos-alvo com desprezo e como indignos de respeito básico.

⁵ Ver Popa-Wyatt (2023).

⁶ Irei, no entanto, permanecer neutra com relação à primeira pergunta. O que irei argumentar a seguir é compatível tanto com tomar o contradiscorso como uma alternativa a leis contra o discurso de ódio quanto como uma medida complementar a elas.

Cepollaro, Lepoutre & Simpson (2023), chamarei esses dois critérios respectivamente de *questão da eficácia* e *questão deôntica*.

Um lugar natural por onde começar é com as refutações diretas que negam os conteúdos expressos por afirmações como (1)-(3). Por exemplo, um falante odioso braveja um estereótipo desumanizador, enunciando, “Xs são baratas/parasitas/macacos” e um ouvinte resistente responde, “Não, isso é falso! Xs não são baratas/parasitas/macacos”.

A eficácia desse tipo de contradiscurso — que é *in situ*, reativo e negativo — entretanto, vem sendo vista com algum pessimismo por certa literatura especializada — fundamentada, principalmente, no estudo empírico de vieses cognitivos que alimentam certa narrativa irracionalista sobre a comunicação política.⁷ Por exemplo, recentemente, Maxime Lepoutre (2021), seguindo McGowan (2019) e Robert Simpson (2013), argumentou que refutações diretas que simplesmente negam o conteúdo explícito do discurso de ódio podem sair pela culatra ao reforçar a saliência das associações perniciosas feitas pelo falante odioso, tornando-as mais aderentes e cognitivamente influentes.

Refutações também são vistas com certa reserva em virtude de considerações morais, já que as principais vítimas de discurso de ódio tipicamente estão mal posicionadas na hierarquia social para negar autoritativamente o que foi dito pelo falante odioso; nesse sentido, muitos questionam a razoabilidade e justiça de impor-lhes a carga da resistência (ver Lepoutre, 2021; Waldron, 2012; Langton, 1993).

No que concerne à questão da eficácia, mais especificamente, o intuito dessa literatura é trazer à tona os potenciais *efeitos consolidantes* das repetições de estereótipos no discurso público. Frente a essas considerações, minha proposta é analisar uma estratégia de contradiscurso ainda pouco discutida⁸ que busca justamente bloquear efeitos consolidantes ao mudar os rumos e/ou os termos da conversa: a *reformulação*. São casos como (4) abaixo, em que o ouvinte resistente responde a (1) com uma correção que busca mudar o rumo da conversa e bloquear a essencialização das mulheres em termos de propriedades negativas.

(4) Mulheres que são educadas em sociedades dominadas por homens podem ser submissas.⁹

⁷ Segundo a qual os problemas fundamentais da comunicação política tem suas raízes na irracionalidade humana e em sua susceptibilidade a padrões disfuncionais de pensamento motivado, enviesado, partidário, etc. (Dorst, 2020).

⁸ Uma exceção pode ser encontrada em Goffredo et al (2022). Barbosa (2023) e Barbosa & Fernandes (2024) discutem uma estratégia de correção semelhante.

⁹ Em Barbosa (ms), eu defendo que respostas como (4) podem bloquear leituras es-

O sentido em que (4) muda o rumo da conversa ficará mais claro quando a noção de “pergunta em discussão” ou *QUD* (do inglês *question under discussion*, que manterei aqui) for introduzida na seção 3.

Além disso, irei sugerir que, se aceitarmos que uma forma de responder à questão deontica é tomando o Estado como “falante resistente”, também podemos explorar o poder das reformulações de gerar mudanças de rumos e principalmente de termos de conversas públicas através “voz do Estado”.

Antes de avançar, porém, faço algumas ressalvas importantes sobre o uso *in situ* de reformulações. O foco da minha análise estará em casos corriqueiros de *asserções de estereótipos negativos*, como (1): representações enviesadas e preconceituosas de grupos-alvo que minam a confiança de seus membros na proteção de sua igualdade social (mais sobre isso na seção 1). Assumirei que enunciações desse tipo são danosas por transformarem a estrutura informacional de conversas em atividades inquisitivas que contribuem para a erosão do senso de dignidade das vítimas. Em razão disso, tomarei (1) e outros casos semelhantes como alvos legítimos de contradiscurso.

Ressalto, porém, que esse tipo de discurso danoso parece contrastar com o que se pode chamar de “discurso de ódio mais sério”, como a propaganda estatal desumanizadora — e.g., (2), no contexto que antecipou o genocídio de Ruanda (ver Tirrell, 2012) — ou a justificação, glorificação ou incitação à violência e ao crime (e.g., “O que fazemos com baratas é esmagá-las sem piedade!” no mesmo contexto) ou ainda o discurso que envolve *slurring* como uma “arma verbal”, a exemplo de (3).¹⁰

De fato, eu acredito que a etiqueta mais adequada para (1) seja a de *discurso preconceituoso*. São casos de discurso tipicamente alimentado por vícios epistêmicos que conduzem mais facilmente a crenças especialmente resistentes a contra-evidências sobre grupos sociais e que podem ser respondidos dentro de “jogos da persuasão” *in situ*.¹¹ O mesmo não pode ser

sencialistas de estereótipos formulados genericamente. Genéricos são semanticamente indeterminados e admitem diferentes leituras, e.g. enquanto generalizações estatísticas ou afirmações sobre a natureza ou essência de um tipo. O que (4) faz é bloquear essa determinação semântica — que, como Anderson, Haslanger & Langton (2012) apontam, torna difícil refutar o genérico porque lhe permite um comportamento semanticamente “escorregadio” — ao escolher a leitura essencialista e levantar alternativas que a negam. Isso possibilita ao falante resistente redirecionar a conversa colocando o essencialismo sobre grupos sociais em cheque sem uma confrontação arriscada. Dado que no presente artigo eu também trato do contradiscurso estatal, por questões de espaço, apresento apenas uma versão menos detalhada das reformulações *in situ*.

¹⁰ Agradeço a um(a) revisor(a) anônimo(a) por levantar esse ponto.

¹¹ Casos assim também envolvem sentimentos e atitudes intolerantes que podem ser

dito sobre (2)-(3) e o exemplo de incitação, já que reparar e prevenir danos nesses casos pode requerer o emprego de dispositivos legais, ação estatal, campanhas massivas e outras estratégias preventivas e punitivas que vão além, por exemplo, de meras refutações, reformulações, ou respostas sarcásticas *in situ*. Sendo assim, para evitar confusões sobre o tipo de fala ao qual minha discussão da reformulação se aplica, irei usar, no que se segue, apenas os termos ‘discurso danoso’ e ‘discurso preconceituoso’, deixando ‘discurso de ódio’ de lado.

Nas seções 1, 2 e 3 apresento a posição de Lepoutre como representante da “literatura pessimista” sobre a eficácia das refutações. São seções substancialmente expositivas, começando, na 1, pela interpretação do dignitarismo adotada por Lepoutre — que combina idéias de McGowan e Waldron. Em seguida, elaboro a questão deontica e recupero seu argumento em favor da solução do contradiscurso estatal. Na seção 3, passo à questão da eficácia, apresentando a crítica de Fraser (2023) e, em seguida, foco em sua sugestão de que o contradiscurso *in situ* deve buscar mudar os rumos de conversas perniciosas pela alteração de *QUDs*. Além disso, analiso criticamente sua estratégia de preferência, o sarcasmo, e, na seção 4, apresento a minha, a reformulação. Complementando a proposta de Fraser, ressalto que um olhar cuidadoso indica que há mais de uma forma de mudar conversas para fins de contradiscurso: uma delas é alterar os *termos* da conversa, mesmo quando a *QUD* permanece a mesma. Na subseção 4.1, considero brevemente um exemplo prático do uso de reformulações no contradiscurso estatal para mudar termos usados em conversas públicas sobre o assédio sexual.

1. Os desafios do contradiscurso no dignitarismo

Segundo o dignitarismo, enunciações como (1)-(3) são danosas principalmente porque negam enfaticamente o reconhecimento do status social igualitário dos membros dos grupos-alvo e minam a confiança em que suas liberdades, proteções e poderes serão os mesmos que os dos demais.

Como já foi mencionado, o principal expoente dessa posição é Jeremy Waldron (2012), que advoga pela necessidade de uma legislação contra o discurso de ódio, com base principalmente nas teses (i)-(vii):

- (i) No interesse do bem público da inclusividade, todos devem ter a garantia de que não terão que lidar com violência, discriminação ou exclusão;

combatidos por estratégias paralelas de foco mais afetivo, como o apelo a narrativas, por exemplo.

- (ii) Alguns grupos vêm de um passado recente no qual essa garantia encontrava-se ausente.
- (iii) Assegurar o status de seus membros como objeto de proteção e preocupação social significa reconhecer sua reputação básica (*basic good standing*) ou *dignidade cívica*. Por ser fundamentalmente um status, a dignidade gera demandas de reconhecimento;
- (iv) O discurso danoso/de ódio nega o reconhecimento da dignidade cívica de indivíduos ao tornar expressões de depreciação, menosprezo, vilipêndio, desumanização, uma parte duradoura e uma memória presentificada em seu entorno;
- (v) Essa não é a aparência de uma sociedade bem ordenada — no sentido de Rawls (ver Moon, 2014)¹². Os cidadãos de uma sociedade bem ordenada *sabem* que seus pares respeitarão sua dignidade.
- (vi) A aparência social (visual, sonora) importa porque é principalmente através dela que uma sociedade dá a conhecer seu compromisso com princípios de justiça.
- (vii) Uma sociedade que permite o discurso danoso/de ódio especialmente quando vista pelos olhos de membros dos grupos-alvo, é injusta e aterradora.
- (viii) Leis que restringem o discurso que envolve danos dignitários são desejáveis para garantir o bom ordenamento (i.e., justiça) social.¹³

Porém, nem todos aqueles que aceitam as teses (i)-(vii) concordam com a conclusão sugerida em (viii). Por exemplo, Lepoutre (2021) parte do dignitarismo, mas defende, na contramão de Waldron, que *certo tipo* de contradiscurso — que ele chama de “contradiscurso positivo” — envolve custos morais menores que soluções coercitivas (e.g., medidas que impõem sanções econômicas ou encarceramento) e, em consequência, é uma ferramenta preferível contra o discurso danoso.¹⁴

¹² Em uma sociedade bem-ordenada, basicamente, todos aceitam e sabem que todos os demais também aceitam os mesmos princípios de justiça. Para Waldron, esse conhecimento compartilhado é implícito e deve permanecer assim. Negar ou questionar a dignidade cívica de membros de um grupo social envolve torná-lo saliente, visível, etc.

¹³ A justificativa de Waldron para sugerir a necessidade de leis é principalmente tomar o discurso danoso/de ódio como um caso de difamação/injúria contra grupos (*group libel*). Outra razão é sua preocupação com a questão deontológica, que apresento em mais detalhe na próxima seção.

¹⁴ Eu não terei muito espaço para tratar do contradiscurso positivo aqui, mas ele se caracteriza basicamente por afirmar uma visão correta de mundo incompatível com o dis-

Por ora, é importante entender apenas que Lepoutre se opõe a (viii) principalmente por pensar que grande parte dos que vociferam visões intolerantes e preconceituosas são indivíduos socioeconomicamente insatisfeitos que se sentem — e, em alguns casos, realmente são — politicamente marginalizados. E como a capacidade de contraposição à dominação desses indivíduos — i.e., à subjugação ao poder arbitrário e sem controle — depende em parte da possibilidade de comunicar publicamente seus interesses e preocupações àqueles que detêm poder, proibi-los pode reduzir suas liberdades significativamente, gerando exclusão e contribuindo para a intensificação dos seus pontos de vista mais perniciosos.¹⁵

Mas isso não significa que Lepoutre ignore que a existência de assimetrias sociais profundas refletidas “no que podemos fazer usando palavras” — para remeter ao famoso slogan austiniano — possam limitar consideravelmente a capacidade de membros de grupos-alvo de combatem o discurso danoso usando apenas a própria voz. Filósofas austinianas contemporâneas a cuja visão ele adere, como Rae Langton, Jennifer Hornsby e Ishani Maitra, em vários trabalhos¹⁶, têm chamado atenção para o fato de que a existência de uma hierarquização de grupos sociais em sistemas desiguais de distribuição de poder podem afetar o que Austin (1962) identificou como *condições de felicidade* e gerar falhas sistemáticas na realização de certas ações discursivas. O resultado é uma espécie de silenciamento.

Essa é uma ameaça clara à premissa inicial de Lepoutre de que o contradiscurso é preferível a medidas coercitivas: se alguns indivíduos es-

curso danoso, acarretando sua negação. Um exemplo dado pelo autor é o famoso discurso “Eu tenho um sonho”, de Martin Luther King Jr. (1963), que se opunha à cultura racista imperante nos EUA da época, não ao negar explicitamente seus estereótipos, mas ao promover uma visão positiva e contrastante da igualdade racial. O objetivo primordial do contradiscurso positivo é colocar associações positivas em evidência, sem propriamente “mudar o assunto” da pauta pública. Isso tem como efeito diminuir a aderência do discurso danoso e aumentar a aderência de discursos que promovem valores democráticos. Eu concordo que o contradiscurso positivo, usado como ferramenta preventiva, é de fato útil e poderoso. Usado *in situ*, no entanto, ele pode ser considerado evasivo. Por exemplo, uma resposta a (1) com algo como “Mulheres são a maioria da força de trabalho em setores essenciais da economia de serviços, como educação e saúde”, embora útil para dar saliência a certas informações sobre o grupo-alvo, pode falhar em gerar oportunidades de persuasão racional e melhoramento epistêmico porque evita a fricção epistêmica que é tipicamente necessária para tal fim. Mais sobre isso na seção 3. Caponetto & Cepollaro (2022) levantam um ponto semelhante sobre o uso do que elas chamam de *bending* como contradiscurso.

¹⁵ Waldron considera essa objeção (o chamado “*freedom-based argument*”), que é popular na tradição liberal, mas ele acha que esse custo é superado pelos custos de gerar danos à dignidade de membros de grupos-alvo.

¹⁶ Ver Hornsby e Langton (1998); Langton (1993), Maitra (2009).

tão mais bem posicionados socialmente que outros para refutar, protestar, denunciar, por exemplo, — atos de fala que são parte da matéria-prima do contradiscurso —, então a sugestão de que o contradiscurso é suficiente (ou menos custoso, pelo menos) para combater o discurso danoso esbarra no que chamei na seção anterior de questão deontica e a recomendação deixa de ser generalizável. Na próxima seção, explico melhor essa afirmação e como Lepoutre soluciona o problema.-

2. Quem está em posição de usar “mais discurso”?

Para responder à pergunta no título, convém começar (re)tomando a distinção de Austin entre três dimensões ou aspectos diferentes dos atos de fala: a *locução*, ou seja, a emissão de palavras e orações com certas propriedades linguísticas (sintáticas, semânticas e fonológicas), a *ilocução*, a ação realizada tipicamente a partir da locução, e a *perlocução*, o ato de produzir efeitos a partir da fala. Além disso, tenhamos em mente o que Austin (1962/1990, p. 30) chamou de “doutrina das infelicidades”, que estabelece tipos de falhas na realização de ilocuções. A mais importante para nossos propósitos será a categoria do *desacerto*, que eu explico a seguir.

Consideremos o ato de casar duas pessoas pela enunciação da locução performativa, “Eu as declaro casadas”. O objetivo dessa ação é gerar o efeito convencional de mudar o estado civil de duas mulheres. Mas, como observado por Austin, isso dependerá da satisfação de um conjunto de condições que tornam o ato de casar o que ele é (i.e., que lhe dão sua força ilocucionária característica), algumas das quais estão convencionalmente estipuladas e institucionalmente cristalizadas, como: a) que o ato seja realizado por uma autoridade institucional; b) que seja dirigido a duas pessoas adultas não-casadas; c) que a ação conte com a presença de certo número de testemunhas, etc.

A não satisfação de pelo menos uma dessas condições poderá levar à falha de procedimento ou “infelicidade” que Austin (1962/1990, p. 32) chamou de desacerto. Imagine, por exemplo, uma situação na qual uma juíza declara o casal casado, mas não há testemunhas presentes ou que uma das noivas já é casada ou ainda que todas as condições anteriores estejam dadas, mas que o país proíba o casamento entre pessoas do mesmo sexo.

Nem todas as ilocuções, no entanto, são atos institucionais, como casar. Segundo a interpretação da teoria dos atos de fala de Strawson (1964), em particular, uma ação como alertar a alguém de algo, por exemplo, requer apenas a produção de apreensão (*uptake*) da intenção comunicativa¹⁷ do

¹⁷ Strawson adota a definição griceana de significado não-natural (Grice, 1957) para

falante por parte do ouvinte e não necessariamente depende de procedimentos convencionais. Suponha que eu estou caminhando com minha amiga Moama e vejo uma bicicleta avançando rapidamente em nossa direção. Eu digo, “Aquela bicicleta está vindo rápido”, com o intuito de alertá-la e levá-la a se proteger do risco iminente de ser atropelada. Se Moama, por alguma razão, não reconhecer que minha enunciação foi realizada não para asserir algo sobre a bicicleta, mas para alertá-la sobre o risco de um atropelamento, meu ato será malogrado e um desacerto será produzido.

Com esses conhecimentos básicos sobre atos de fala em mente, podemos passar agora ao que defendem algumas das filósofas feministas mencionadas na seção anterior: que alguns desacertos são produzidos graças ao pertencimento do falante a um grupo-alvo e isso pode gerar seu silenciamento e a impossibilidade de uso de “mais discurso”.

Consideremos, para começar, o ato de *protestar*, tomando um exemplo dado por Langton (1993): o caso da catalogação da autobiografia de Linda Marchiano, também conhecida como Linda Lovelace, *Ordeal*, como leitura pornográfica (Langton, 1993, p. 321). *Ordeal* conta a experiência de Marchiano como atriz na indústria pornográfica norte-americana dos anos de 1970 e, em particular, narra como a autora foi vítima de violências abomináveis durante as gravações do filme *Deep throat*: violência física, hipnose involuntária e tortura. Segundo a interpretação de Langton, a intenção da autora era denunciar as práticas profundamente condenáveis da indústria e, principalmente, protestar contra a pornografia dominante, mas o livro terminou listado como “leitura adulta” em um catálogo que vendia pornografia.

Nesse contexto, o ato de protesto de Marchiano sofreu um desacerto causado pelo que Langton chama de *descapacitação ilocucionária* da sua fala: “Algo sobre quem ela é, algo sobre o papel que ela exerce, impede que ela satisfaça as condições de felicidade do protesto¹⁸, pelo menos aqui” (p. 322). Em um sentido, seu protesto é tornado “indizível” (*unspeakable*, no original): ela pode usar as palavras corretas, realizar as locuções necessárias, mas seu protesto (i.e., a força ilocucionária própria do ato de protestar) contra a pornografia é interpretado como o seu exato oposto, como uma jogada no jogo de linguagem da pornografia. Essa descapacitação da sua agência discursiva não é apenas um dano em si, mas também um obstáculo

explicar a apreensão: um ouvinte apreende o que um falante teve a intenção de comunicar quando ele passa a ter um estado cognitivo (e.g. uma crença) como resultado de seu reconhecimento da intenção do falante de produzir tal estado.

¹⁸ Para Langton (1993), a apreensão de uma ilocução é uma condição de felicidade de sua realização, mas ela não oferece uma explicação detalhada dos mecanismos da apreensão. Maitra (2009) preenche essa lacuna ao explicar a apreensão em termos griceanos.

à sua capacidade de recorrer a “mais discurso” contra seu vitimário, a indústria pornográfica (Langton, 1993, p. 325).

E esse tipo de impedimento não é o único que afeta membros de grupos-alvo. Miranda Fricker (2007), Rebecca Mason (2021) e outros nos falam, por exemplo, de casos nos quais conceitos e palavras (ou significados) que dão sentido e *comunicam* adequadamente certas experiências de membros de grupos-alvo estão ausentes do léxico coletivo ou são deficientes, graças à sua exclusão das trocas epistêmicas. Fricker chamou esse mal de *injustiça hermenêutica*. Alguns exemplos conhecidos discutidos pela autora são os dos conceitos de assédio sexual e depressão pós-parto, cuja ausência, por muito tempo, impossibilitou as mulheres (particularmente) de se referirem adequadamente às investidas de assediadores e à depressão pós-parto como experiência socialmente compartilhada.

No caso específico do assédio sexual, que nos interessa em particular, a lacuna hermenêutica foi com frequência preenchida por palavras como “brincadeira”, “elogio”, “cantada”, “flerte”, que decididamente careciam da significância normativa adequada para capturar o caráter reprovável e transgressor da conduta em questão (Romdenh-Romluc, 2017). Meu ponto ao trazer o conceito de injustiça hermenêutica aqui é mostrar como a falta de poder hermenêutico pode impactar o uso pleno de “mais discurso” por parte de certos indivíduos. Exemplos famosos como os das americanas Carmita Wood e Anita Hill, mencionados por Fricker (2007), ilustram bem como a falta ou a incipiência de recursos hermenêuticos em torno do assédio sexual obstaculizou os processos de *denúncia* dessas práticas em instituições democráticas importantes, como o poder judiciário estadunidense e as universidades daquele país. Parece ser o caso então, que, sem os recursos (palavras e conceitos) corretos, denúncias se tornam inescrutáveis ou esvaziadas da significância normativa adequada e que, mais uma vez, isso impacta negativamente a capacidade de membros de grupos-alvos de usar “mais discurso”.

Mas, mesmo quando os recursos hermenêuticos necessários estão disponíveis e a apreensão ilocucionária está alinhada com a intenção comunicativa do falante, o discurso de um membro de um grupo-alvo pode não ter os efeitos *perlocucionários* pretendidos, agora em virtude da interferência de normas culturais dominantes. Vejamos um exemplo.

Comentário objetificador na aula de lógica: Maria é uma jovem estudante do primeiro ano de filosofia. Em uma aula de lógica, o professor convida os alunos da sala a resolverem uma prova no quadro. Maria se oferece diligentemente, resolve com sucesso o problema e é parabenizada pelo professor, que pergunta ao resto da turma “Todos

entenderam?”. João, um aluno no último ano do curso que Maria viu poucas vezes, responde com um sorriso, “Não, professor. Um corpo desses me distrai de qualquer lógica”. Muitos riem (alguns embaracosamente).¹⁹

Em um caso assim, que opções de “mais discurso” tem Maria, o alvo do comentário objetificador? Uma delas é certamente confrontar e criticar João, talvez até condenar explicitamente o comportamento compactuante dos demais colegas e do professor. Em princípio, essa opção permite a Maria não apenas reequilibrar a situação ao expor seu desconforto, mas também bloquear a possibilidade de que o contexto da aula se torne aberto a expressões de objetificação feminina que respaldam visões sexistas.

Uma filósofa contemporânea que explica como a estrutura informacional de uma conversa pode acabar favorecendo visões de mundo opressivas, como o sexismo, é McGowan. Ela conceitualiza trocas conversacionais — e.g. a aula de lógica — aos moldes de Lewis (1979): como atividades cooperativas a serem entendidas em analogia com a cinética de jogos de *scorekeeping*, cujos placares determinam a permissibilidade de jogadas atuais a partir das jogadas anteriormente registradas. Nessa visão, cada enunciação realizada influencia o que é permissível no futuro da interação. McGowan (2019) explica essa característica em termos do que ela chama de *normas-s* (*s-rules*, no original), regras locais que instauram fatos de permissibilidade por meio de exercitativos conversacionais.

No caso da fala acima, quando João trata Maria como um objeto de desejo sexual, segundo McGowan, ele gera, paralelamente, um fato de permissibilidade ou norma-s segundo o qual se torna permitido falar de mulheres como meros objetos de desejo sexual. Se antes não era permitido tratar mulheres como objetos — porque, suponhamos, havia regras igualitárias imperando no ambiente da sala de aula — a partir da fala de João, a permissibilidade de falas objetificadoras se torna um fato. E, embora os outros participantes não precisem compactuar pelo simples fato de que a enunciação foi registrada, sua realização gera demandas de resposta, tácita ou articulada, dadas expectativas naturais de cooperatividade entre falantes.

Outra forma de entender o modelo de *scorekeeping* é incluindo *QUDs* no placar (Roberts, 2015). Nessa interpretação, a estrutura informacional de uma conversa evolui em certa direção não como resultado de atos de fala implícitos, mas porque as jogadas futuras devem responder à *QUD* mais imediata para contar como contribuições cooperativas. Eu explico essa alternativa melhor apenas na próxima seção, já que ambas interpretações

¹⁹ Esse exemplo está adaptado de McGowan (no prelo).

são compatíveis com o que direi a seguir sobre respostas possíveis ao exemplo do comentário objetificador.²⁰

Voltando ao exemplo: uma opção para um ouvinte resistente, e.g., Maria, é impedir que a enunciação de João conte como parte das pressuposições válidas no contexto ao refutar o que foi dito. Maria pode responder, por exemplo,

(5) Cara, o que você acabou de dizer é tão equivocado e em tantos níveis que nem sei por onde começar, mas eu vou ressaltar o mais danoso de todos: você foi absolutamente sexista e me objetificou!

Um dos efeitos de (5) será eliminar o conteúdo da contribuição de João do terreno comum (*common ground*) da conversa, ou seja, daquilo que é aceito (ou tratado como aceito) por todos os participantes. A conversa que se segue não poderá, então, tomar a fala de João como parte do seu pano de fundo.

Mas, será que o ato de resistência discursiva de Maria é justo *com Maria*? É bastante provável que não, já que o tom direto e confrontacional de (5) pode inverter o julgamento de outros membros do grupo acerca de quem, entre Maria e João, foi a verdadeira vítima de um mau comportamento. Isso se dá, porque, *qua* atividades sociais, conversas incluem normas culturais, — subtipos de “Normas-G”, no modelo de McGowan — algumas das quais envolvem critérios de polidez.

Em particular, uma resposta confrontacional como a de Maria, que não se esforça por preservar a face (*save face*) de seu adversário, pode terminar saindo pela culatra como contradiscurso. Ressalte-se também que regras de polidez na chamada cultura ocidental costumam ser mais rigorosas com as mulheres (McGowan, no prelo). Assim, alguns efeitos perlo-

²⁰ Como adiantei na introdução, irei usar a noção de *QUD* para explicar como e por que (1) (e.g.) é danosa, favorecendo a interpretação de Roberts e o dignitarianismo. Embora McGowan ofereça *insights* interessantes sobre como conversas podem se tornar atividades sociais de discriminação, sua teoria traz alguns problemas que debilitam um pouco seu apelo. Marques (2023), por exemplo, levanta dois pontos importantes. Primeiramente, que o conceito de norma-s não explica satisfatoriamente porque uma asserção como (1), por exemplo, é *constitutivamente* danosa. Afinal, implementar uma norma de tratamento diferencial, por si só, não constitui necessariamente um dano. Se uma empresa implementa a política de deixar de contratar pessoas com predisposição comprovada ao câncer de pele para um posto de trabalho como salva-vidas numa praia tropical, essa norma não é danosa ou injusta. Desse modo, parece ser que o que determina se uma norma é ou não danosa são condições externas ao ato ilocucionário. Essa é uma má explicação para o que é o discurso *constitutivamente* danoso e para como ele opera, como conclui corretamente Marques. O outro problema diz respeito a como uma norma-s se instala em um contexto. Por exemplo, se um grupo de pessoas responder a uma pergunta impertinente ignorando-a, torna-se ou não permitido fazer perguntas impertinentes na continuação da conversa?

cucionários de (5), como o arrependimento e retratação de João e a solidarização dos demais, podem terminar perdidos, limitando o uso de “mais discurso” por parte de Maria. Mais especificamente, ela sofreria um tipo de silenciamento do tipo perlocucionário, ou seja, que a impede de obter certas consequências desejadas por meio da fala. Langton (1993) os denomina “frustração perlocucionária”.

Os três casos dos quais tratei nesta seção — a saber, de descapacitação ilocucionária, da frustração perlocucionária e da existência de lacunas hermenêuticas — sugerem algumas limitações preocupantes para o uso de contradiscurso *in situ* contra o discurso danoso por parte das vítimas; limitações essas que precisam ser tratadas pelo defensor da viabilidade do contradiscurso, especialmente tendo-se em vista o contraste com medidas de restrição legal, que parecem proteger membros de grupos desprivilegiados do fardo da autodefesa em situações desvantajosas (Waldron, 2012).

Uma resposta a essa questão é propor que o Estado, em sua comunicação, seja o principal responsável por combater o discurso danoso.²¹ Afinal, o Estado possui autoridade, poder e uma plataforma de grande alcance, além de dispor de uma ampla gama de recursos e possibilidades de ação (Brettschneider, 2012; Gelber, 2002). Em particular, ele é menos suscetível aos tipos de silenciamento discutidos anteriormente.²² E já que o Estado tem autoridade para determinar os termos e rumos das conversas públicas e promover novas normas culturais, ele pode ser visto como uma espécie de “interlocutor ideal do falante preconceituoso”. Essa opção soa intuitivamente menos injusta do que cobrar de vítimas comunicativamente desempoderadas deveres de resistência ao preconceito.

Além disso, o Estado está *prima facie* em melhores condições de levar a cabo mecanismos preventivos de contraposição, especialmente por meio de campanhas educativas, da fala de políticos influentes, do financiamento de medidas e instituições dedicadas a combater o discurso danoso, etc. De fato, para autores como Brettschneider, o contradiscurso é um dos *deveres* do Estado liberal, democrático e plural, cujo discurso deve inclusive ignorar demandas de neutralidade para proteger irrestritamente princípios de igualdade cívica.

Voltando brevemente a Lepoutre, ele também vê o contradiscurso estatal como uma solução à questão deontica (Lepoutre, 2021, pp. 92-97).

²¹ Os três casos mostram limitações para o discurso das *vítimas* do dano discursivo, mas ainda podemos assumir que espectadores que não pertencem aos grupos-alvo contam como bons candidatos a ouvintes resistentes. Essa é uma alternativa da qual não tratarei aqui.

²² Com exceção do problema das lacunas hermenêuticas, que afetaria também a fala do Estado.

Porém, como ele destaca, essa é apenas metade da solução ao problema de como combater o discurso danoso. Afinal, o discurso do Estado é, em princípio, tão passível de sair pela culatra por razões que têm a ver com a eficácia das estratégias de comunicação escolhidas quanto o do cidadão.

3. Como usar “mais discurso”?

O que venho chamando aqui de “refutações diretas” e que Lepoutre (2019; 2021) chama de “contradiscurso negativo” — e.g., (6) abaixo em resposta a (1) — são basicamente *correções* que sinalizam a rejeição do conteúdo de atos de fala antecedentes e, conseqüentemente, de sua inserção no terreno comum.

(6) Não, mulheres não são submissas.

A qualificação “diretas” capta o fato de que essas correções negam os estereótipos asseridos, e não pressupostos ou inferidos por implicatura pelo ato de fala, como fazem o bloqueio por explicitação (ver Langton, 2018) ou as negações metalinguísticas (ver Haslanger, 2010).

Como mencionei na introdução, embora refutações diretas pareçam ser candidatas naturais ao papel de contradiscurso, certa literatura recente, que inclui Lepoutre, é pessimista sobre sua eficácia. Uma das razões para isso é o que eles chamam de problema da *maleabilidade assimétrica dos contextos*: que é mais fácil tornar uma jogada conversacional permissível através de uma ação verbal do que revogar tal permissão pelos mesmos meios (McGowan, 2019).

Essa propriedade aparece originalmente em observações de Lewis (1979), e é tomada por McGowan em sua teoria sobre a relação entre discurso, normas e opressão (Simpson, 2013, pp. 571-575). Como vimos na seção anterior, para a autora, afirmações como (1) geram norma-s que tornam permissível discriminar mulheres no contexto em questão. Assim, por exemplo, se antes de (1) não era permitido tratar mulheres como inferiores, depois de (1), passa a ser; se mulheres estavam asseguradas de seu status social igualitário antes de (1), com (1), elas deixam de estar.

Mas revogar esse tipo de norma-s é praticamente impossível porque elas geram atualizações concomitantes nos fatos de saliência do contexto. E, pelo menos no modelo de *scorekeeping* da autora, essas mudanças são geradas automaticamente, sem a necessidade de aceitação mútua dos participantes (McGowan, 2019, pp. 44-47). Como resultado, uma enunciação perniciosa como (1) poderá tornar uma associação sexista saliente e evocar, no processo, o sexismo enquanto visão de mundo (Simpson, 2013, pp.

572-575). Nesse sentido, uma refutação direta — e.g., (6) —, que repete a associação contestada, pode, inadvertidamente, reforçar a saliência daquilo que deseja combater (a saber, a associação perniciosa e a visão de mundo da qual ela é parte).

Os problemas com as refutações diretas não param por aí. Simpson (2013) e Lepoutre (2021), resgatando estudos de psicologia cognitiva, chamam atenção para os efeitos da chamada Heurística da Fluência, que prevê que a saliência de associações pode gerar disposições doxásticas que favorecem o discurso preconceituoso. Por um lado, a saliência de uma associação reforça sua familiaridade e, por outro, a familiaridade de uma informação pesa cognitivamente em favor de sua aceitação em re-exposições (Lewandowsky et al, 2012). Assim, um indivíduo para o qual a associação entre mulheres e submissão é familiar, estaria mais disposto a aceitar enunciações como (1) em re-exposições do que indivíduos que nunca escutaram tal associação.²³ Para esses dois autores, portanto, as refutações diretas são não apenas ineficazes como contradiscurso, mas também potencialmente danosas.

3.1. Uma explicação alternativa da ineficácia das refutações diretas

O argumento contra as refutações diretas — e principalmente a versão sustentada por Lepoutre —, recebeu uma crítica bastante debilitante de Rachel Fraser (2023). Ela ataca especificamente sua confiança na psicologia associativa para fundamentar um diagnóstico de ineficácia.

A primeira observação interessante de Fraser nesse sentido é que associações são *semanticamente transparentes*, ou seja, elas não representam que tipo de relação se dá entre os conteúdos associados. Por exemplo, ter uma associação psicológica entre sal e pimenta em mente não significa necessariamente ter a disposição de acreditar nem que o sal é pimenta nem que o sal *não é* pimenta. Similarmente, uma mera associação entre, digamos, *mulheres* e *submissão*, não explicaria disposições doxásticas específicas acerca das mulheres, já que a mesma associação serviria para explicar tanto a disposição para crer que *Mulheres são submissas* quanto a disposição para crer que *Mulheres não são submissas*. “[A associação], sendo semanticamente cega, ‘não se importa’ com a presença ou ausência do ‘não’”

²³ Esse efeito foi confirmado por replicações do paradigma testado por Johnson e Seifert (1994), que mostra que retrações de desinformação raramente têm o efeito pretendido de eliminar sua fiabilidade (Lewandowsky et al, 2012). Schwarz et al (2007) explicam esse fenômeno com base na experiência metacognitiva da fluência em re-exposições à informação retratada. Eles concluíram que, quando não temos razões para duvidar da veracidade de uma proposição, a avaliação de sua plausibilidade dependerá fortemente de sua familiaridade e compatibilidade com um modelo mental coerente.

(Fraser, 2023, p. 322); isso indicaria que uma associação em si mesma não teria o poder de minar a dignidade cívica de alguém ou de um grupo (e.g., as mulheres), mesmo quando altamente saliente.

Essa dificuldade sugere que a *mera* saliência de estereótipos não seria a ferramenta mais adequada para explicar a falha ou sucesso do contradiscurso. Fraser sugere uma alternativa centrada no conceito de *QUD*, que é uma ferramenta analítica usada para explicar a relação entre frases ou sequências de frases e o discurso (Benz & Jasinskaja, 2017). A ideia geral é que uma *QUD* representa as metas compartilhadas pelos participantes de uma conversação e captura, intuitivamente, aquilo *sobre o que uma conversa é em um dado momento*.

Na visão de Craige Roberts (2012), uma das principais proponentes desse quadro teórico, uma conversa deve ser entendida como um jogo inquisitivo entre participantes racionais e cooperativos, cujo propósito comum, a cada momento da interação, é responder a uma pergunta implícita aceita por todos: a *QUD* mais imediata. Os participantes, então, estipularão conjuntamente uma *estratégia de investigação* que seleciona uma pilha de sub-perguntas estruturadas hierarquicamente, a serem respondidas e, conseqüentemente, dispensadas da pilha, na medida em que o discurso avança rumo a uma resolução. Será um requerimento racional — seguindo máximas griceanas, como Relevância e Quantidade (Grice, 1975) — que todas as novas contribuições (i.e., cada nova jogada) respondam à *QUD*, contando ou como uma resposta completa ou parcial — tipicamente uma asserção — ou uma nova pergunta. Nesse modelo, então, participar de uma conversa requer tanto entender frases quanto recuperar as questões implícitas que elas respondem como parte de uma estratégia de investigação compartilhada (Kehler & Rhode, 2016).

Suponha que eu inicie uma conversa afirmando (7) no contexto de uma interação com colegas de profissão em um *coffee break*.

(7) Mary Kate McGowan é uma filósofa contemporânea importante.

Se quisermos representar a estrutura informacional da conversa após a ocorrência de (7) usando *QUDs*, alguns candidatos naturais são: ‘Mary Kate McGowan é uma filósofa contemporânea importante?’, chamemo-la Q1, e ‘Quais filósofas contemporâneas são importantes?’, chamemo-la, Q2. Se um colega responder algo como ‘Não concordo que ela seja importante’, ele terá tratado Q1 como a *QUD*; na ausência de divergência, Q1 será selecionada como tal e uma estratégia de investigação para respondê-la será estabelecida. Se o meu colega responder, em contrapartida, “Sally Haslanger e Rae Langton também”, Q2 passa a ser a melhor candidata.

Na visão de Fraser, a saliência que realmente importa para explicar a eficácia ou ineficácia das refutações diretas, e de correções em geral, não é a saliência de associações danosas, mas a saliência que é capturada por *QUDs* (Lepoutre, 2023). Nessa interpretação, (1) é danosa não porque torna a associação negativa entre mulheres e submissão saliente, mas porque responde à pergunta “Mulheres são submissas?”, uma *QUD* que decididamente ameaça a confiança das mulheres de que sua dignidade cívica estará protegida no contexto em questão. A forma mais adequada de combater esse tipo de enunciação *in situ*, sugere Fraser, é, então, mudando a *QUD*.

Mudanças de *QUD* se dão com frequência e tipicamente de modo implícito. Suponha que, no contexto da conversa acima, meu colega me responde primeiramente, “Sim, ela é uma filósofa importante”, e, em seguida, diga, “Sally Haslanger e Rae Langton também”; nesse caso, a *QUD* muda rapidamente de Q1 para Q2. Isso representa uma alteração na estrutura informacional do discurso, que deixa de responder a uma pergunta polar (i.e., do tipo sim/não) para responder a uma pergunta que divide o espaço lógico em múltiplas células.

Fraser sustenta que esse mecanismo pode ser explorado no caso do contradiscurso. Em particular, ela apresenta o sarcasmo como uma forma de mudar a *QUD* para blindar a evolução da estrutura informacional dos efeitos do discurso danoso. Seu principal exemplo é (8) em resposta a (1),

(8) Ah, claro! Mulheres são *tão* submissas! Minha mulher faz tudo que eu mando.

Para ela, essa resposta muda a *QUD* de “Mulheres são submissas?” para “Por que alguém seria estúpido o suficiente para pensar que mulheres são submissas?”, um movimento que transforma o jogo: os participantes saem de uma conversa na qual o status social das mulheres está em questão para uma conversa na qual a estupidez do sexismo está em questão. Um sexista se sentirá bem menos à vontade na segunda conversa do que na primeira. A primeira pode ser definida como um jogo sexista, a segunda como um jogo de resistência conversacional (especificamente, ao sexismo).

Note que uma resposta como (6) falha em gerar esse efeito. (6) responde à mesma *QUD* que (1), a saber, “Mulheres são submissas?”. Tal constatação leva Fraser a concordar com Lepoutre sobre a ineficácia das refutações, mas partindo de razões diferentes. Em sua visão, refutações falham não porque reforçam acidentalmente associações negativas, mas sim porque falham em mudar a *QUD*, aceitando que a estratégia de investigação seja confortável para o falante preconceituoso.

Eu concordo com o espírito da proposta de Fraser, desde sua desconfiança do diagnóstico de Lepoutre acerca da tendência à ineficácia das refutações diretas, até sua sugestão de que uma boa estratégia para proteger contextos de enviesamento em favor da opressão, do ódio e da intolerância envolve mudar a *QUD* e manipular a saliência de perguntas. Entretanto, eu irei argumentar a seguir, primeiramente, que o uso do sarcasmo como contradiscurso tem limitações interpretativas e epistêmicas que Fraser não discute e que podem comprometer o apelo de sua proposta. E em segundo lugar, eu chamo atenção para outros efeitos curativos que “mudar a conversa” podem ter sobre contextos, até mesmo quando a *QUD* não muda.

4. A reformulação como contradiscurso

O sarcasmo é o uso de ironia verbal para ridicularizar algo ou alguém. Tipicamente, atos de fala desse tipo são considerados rudes e até ofensivos, porque ameaçam a face positiva do ouvinte, ou seja, seu desejo de que seus desejos sejam também desejáveis para outros (Brown & Levinson, 1987). Quão representativo é o risco de colapso cooperativo em virtude desse perfil sociolinguístico, entretanto, irá depender de fatores como a familiaridade entre os falantes.

Algumas pesquisas indicam que a familiaridade entre interlocutores tanto aumenta a frequência de uso da ironia quanto conduz com mais frequência à leitura humorística em vez de agressiva do sarcasmo por parte dos ouvintes (Kotthoff, 2003; Eisterhold et al., 2006). Em contrapartida, na ausência de familiaridade, o uso de sarcasmo é mais comumente interpretado como crítico. É de se esperar, então, que, em contextos com maior grau de familiaridade (e.g., conversas informais entre amigos e parentes), uma resposta indireta como (8) seja interpretada como menos ameaçante à cooperatividade do que (6) — ou uma versão mais crítica, como: “Não seja ridículo! Mulheres não são submissas!” — principalmente por sinalizar um sentimento de afiliação entre o falante e o ouvinte.

Imagine que Victor, um antissexista convicto, se encontra semanalmente com seu cunhado sexista, João, com quem, apesar das diferenças, tem uma relação cordial. Em várias conversas prévias com João, Victor tentou levá-lo a mudar sua visão de mundo, falando positivamente sobre a igualdade de gênero, etc., mas seus esforços foram inúteis. Caso Victor finalmente se canse, ele poderá usar o sarcasmo para proteger o contexto da influência de frases como (1), “cortando o mal pela raiz”, na expectativa de que sua jogada soe menos confrontacional do que uma resposta direta como (6) e, ao mesmo tempo, elimine a pergunta “Mulheres são submissas?” do espaço lógico da conversa.

Esse é um resultado desejável.²⁴ Porém, mesmo em contextos nos quais o sarcasmo é à primeira vista recomendável, os custos interpretativos do emprego de meios indiretos de comunicação, como é o caso da ironia, podem levar a interpretações equivocadas (McGowan, no prelo). Suponha que João afirma (1) e Victor responde com (8) tendo como objetivo principal comunicar sua condenação do endosso de João de uma visão de mundo que ele considera opressiva e injusta e impedir que a conversa passe a ameaçar ideais dignitários. Seu objetivo principal, contudo, pode ser obscurecido pelo uso do humor e se perder na complexidade da interpretação. Por exemplo, João pode entender a crítica como pessoal e falhar completamente em entender que Victor desejava condenar o sexismo em sentido mais geral, como supõe Fraser.

Por fim, uma resposta indireta como (8) também pode falhar em deixar claro o que exatamente está sendo julgado equivocado no conteúdo de (1). Um falante pouco virtuoso epistemicamente — e.g., ingênuo, presunçoso, ou mal informado — pode, como resultado, falhar em encontrar uma forma melhor de pensar sobre os gêneros. Compare (8) com (4), por exemplo: (4) também muda a *QUD*, de “Mulheres são submissas?” para “Por que mulheres são submissas?” ou “Mulheres são submissas por natureza?”, ambas perguntas que tanto desviam os rumos da conversa quanto criam oportunidades de aprendizado social a partir do que Medina (2013, p. 50) chama de *fricção epistêmica benéfica*. Trata-se de um tipo de resistência entre pontos de vistas diferentes que pode conduzir à superação da “ignorância social” — i.e., a ausência de conhecimento sobre grupos-alvo que tem como base a opressão (Mills, 2007) — e seus vícios epistêmicos típicos.

Meu principal argumento em favor das vantagens dessa estratégia é de fato epistêmico. Em particular, me parece que embora reformulações não gerem resultados como eliminar ou silenciar localmente perguntas perniciosas — aos moldes de (8) —, elas geram oportunidades de fricção epistêmica benéfica.

Considerando o cenário da conversa entre Victor e João, imagine agora que João pertence a um grupo social privilegiado e está acostumado a não se perguntar sobre a experiência, crenças e motivações de grupos-alvo. Em particular, João nunca se perguntou se haveria outras razões pelas quais as mulheres em seu meio se comportam de modo submisso, além de sua suposta natureza. Segundo Medina (2013), essa atitude reflete vícios

²⁴ Bartlett & Krasodonski-Jones (2015) mostraram também que o uso de humor em resposta ao discurso danoso na internet costuma gerar mais dados de engajamento em redes sociais — como, likes e compartilhamento no X (antigo Twitter) e no Facebook — do que respostas antagonicas — como (5) e suas variantes. Isso sugere que, na internet pelo menos, o uso de sarcasmo funciona melhor que o uso de refutações.

de *preguiça epistêmica e de mentalidade fechada* — i.e., resistência a buscar contra-evidências às próprias crenças e uma atitude de “não precisar saber”, respectivamente. Podemos imaginar, então, um cenário em que (4), além de gerar uma mudança de *QUD*, ajuda a colocar em curso um processo compartilhado de questionamento e superação da ignorância social de João, guiado por uma indicação clara de como a crença contestada pode ser substituída. Esse resultado é melhor, de um ponto de vista epistêmico, do que aquele esperado no uso do sarcasmo, principalmente porque não gera interpretações equivocadas que potencialmente frustram os propósitos do ouvinte resistente.

No que concerne à estrutura do discurso, respostas como (4) representam jogadas *conversacionais melhorativas* (ver Caponetto & Cepollaro, 2023). “Mudar a conversa” ao mudar a *QUD* é uma forma de levar a tarefa de melhorar contextos conversacionais moralmente a cabo, mas outras, como mudar os *termos* da conversa, sem necessariamente alterar a *QUD*, também são possíveis. Por exemplo, (9) B. em resposta a (9)A. abaixo.

- (9) A. Maria ficou chateada porque o chefe elogiou os seios dela.
B. Maria ficou chateada porque o chefe a *assediou sexualmente*.

Suponha que (9) A. foi dita em um contexto empresarial no qual se pode assumir que todos os participantes conhecem o conceito de assédio sexual. Para entender como (9) B. opera, assumamos que a *QUD* em (9) A. é “Por que Maria ficou chateada?”, chamemo-la, Qm. Uma reformulação como (9) B., em resposta, não altera a *QUD* — já que (9) B. também é uma resposta a Qm —; porém, (9) B. melhora os termos da conversa. Em particular, a expressão ‘*assediou sexualmente*’ captura a significância normativa correta — i.e., a valência negativa — que permite classificar a ação em questão como uma infração condenável.

O termo reformulado ‘*elogiou*’, em contrapartida, obscurece a possibilidade dessa classificação em vistas da sua aceção positiva; e mais importante: ele obstaculiza a cristalização do uso cotidiano de um termo que preenche uma lacuna hermenêutica. Medina (2023) chama distorções desse tipo de *bloqueios hermenêuticos*. Tal como eu usarei o termo aqui, ele denomina a substituição de termos que preenchem lacunas hermenêuticas por outros que obscurecem, distorcem ou deflacionam a experiência referida. A reformulação aos moldes de (9) B. surge, em face disso, como uma *tática de desbloqueio hermenêutico* e, em última instância, de empoderamento hermenêutico *in situ*.

Para Langton (2010, p. 463), uma forma de combater a marginalização hermenêutica é buscar nomes adequados para descrever as experiên-

cias das vítimas e aproveitar o *momentum* social para *consolidar o nome dado*. Eu acrescentaria, pegando emprestado o conceito de Medina, que outra forma importante é seguir resistindo continuamente a comportamentos que buscam bloquear hermeneuticamente o conceito conquistado e o nome dado à experiência em questão. O diálogo em (9) pode ser interpretado às luzes dessa sugestão.

O uso de reformulações serve para evitar esses efeitos e tentar garantir que conteúdos que conceitualizam inadequadamente uma conduta não cheguem a formar parte do terreno comum. Isso ajuda a blindar os rumos de conversas correntes e futuras. E essa é uma tarefa que o Estado também pode assumir no âmbito das conversas públicas.

4.1. *Mudando os termos de conversas públicas: o caso da cartilha do Chega de Fiu Fiu*

Como vimos anteriormente, reformulações aos moldes de (4) e (9) B. são exemplos de contradiscurso reativo, *in situ* e negativo. São respostas corretivas dadas em reação a falas preconceituosas e/ou injustas, negando parcialmente o que foi dito pelo falante. Reformulações como (4) mudam o que está em questão, desviando a rota de conversas para potencialmente gerar oportunidades de aprendizado social. Já reformulações do tipo de (9) B. buscam basicamente desbloquear um bloqueio hermenêutico.

Como sugiro a seguir, pelo menos as reformulações do segundo tipo (9) B. podem ser usadas na comunicação do Estado para gerar efeitos preventivos de empoderamento hermenêutico. Começarei assumindo, seguindo Brettschneider (2012, p. 71), que o Estado tem poder *expressivo* além de coercitivo. Isso significa que o Estado pode, ou melhor, deve participar ativamente de *jogos de persuasão* em torno de valores democráticos — i.e., deve falar em prol da proteção de direitos e tentar não-coercitivamente mudar crenças antidemocráticas. Em particular, o Estado deve reforçar seu compromisso com tais valores, apresentando-os como “razões que fundamentam direitos”.

Brettschneider cita algumas formas pelas quais isso pode ser realizado: nas justificativas de decisões da Corte Suprema (nos EUA) ou do STF (no Brasil), no currículo escolar, na preservação da memória de movimentos e personalidades que representam ideais democráticos. Outra forma está na publicação de materiais como cartilhas, que contém informação útil e acessível sobre direitos e deveres.

Há muitos exemplos reais de uso da “voz” do Estado desse tipo, mas me concentrarei em um caso ilustrativo aqui: a cartilha sobre o assédio

sexual lançada em 2014 pela Defensoria Pública do Estado de São Paulo.²⁵ Esse material apresenta o conceito de assédio sexual em espaços públicos de um modo que salienta a igualdade e a liberdade das mulheres como valores endossados pelo poder judiciário do estado de São Paulo. Por exemplo, quando responde à pergunta “Por que devemos denunciar o assédio?”, a cartilha diz: “[para] mostrar que podemos *igualar* a voz e o poder da mulher na sociedade” (meu itálico). Ao considerar as consequências do assédio, o material foca na perda de liberdades, ao dizer: “muitas [mulheres] *podam sua própria liberdade e seu direito de escolha* — deixando de usar uma roupa ou de cruzar uma praça, por exemplo — por medo de sofrer tais abordagens” (meu itálico).

Mas, o que torna essa cartilha especialmente interessante para os propósitos presentes é o fato de que ela fala em favor da igualdade das mulheres *reformulando* aplicações problemáticas dos conceitos de paquera, elogio e cantada em termos de assédio sexual. O texto começa perguntando: “Andar pelas ruas e ouvir um comentário obsceno sobre o seu corpo é um elogio? Ouvir uma cantada no ambiente de trabalho é algo natural?”. E segue explicando que elogios e paqueras não causam medo e angústia e que formular assédios nesses termos é uma distorção que confunde as vítimas e blinda a ação de ser vista como imoral e danosa.

A cartilha faz algo muito parecido com o que o falante de (9) B. faz: busca impedir que a formulação problemática de assédios como cantadas ou elogios se torne ou siga sendo parte do terreno comum das conversas públicas acerca do lugar das mulheres na sociedade. Está, além disso, seguindo a recomendação de consolidar o nome correto do assédio e de impedir bloqueios hermenêuticos.

5. Conclusão

Na contramão de certa literatura recente, eu busquei defender aqui que o contradiscurso negativo não precisa ser abandonado ou visto de modo pessimista no combate ao discurso danoso. Opor-se ao falante preconceituoso ao reformular as ideias que ele expressa — que é uma forma essencialmente negativa de engajamento — não apenas desvia o rumo de conversas perniciosas para destinos melhores, como pode gerar oportunidades importantes de superação conjunta da ignorância ativa; uma vantagem sobre outras formas de contradiscurso *in situ*, como o sarcasmo. E mesmo quando o

²⁵ Essa cartilha foi feita em parceria com a ONG Think Olga, e com base na campanha *Chega de Fiu Fiu* da jornalista Juliana de Freitas. Ver https://thinkolga.com/wp-content/uploads/2020/04/ThinkOlga_Cartilha-Assedio.pdf. Acesso em 30 de julho de 2025.

rumo das conversas (entendidas estruturalmente em termos de *QUDs*) não são propriamente alterados, o uso de reformulações para mudar práticas linguísticas pode ajudar a blindar casos do que chamei, seguindo Medina (2023), de bloqueios hermenêuticos.

Essa segunda vantagem é algo que o Estado, em seu poder expressivo, pode explorar para reforçar seu compromisso com a igualdade cívica. Segundo a minha breve análise da subseção 4.1., a campanha *Chega de Fiu Fiu* da Defensoria Pública de São Paulo exemplifica justo essa possibilidade de reformular conceitos e cristalizar termos corretos para o assédio sexual e, assim, falar em favor da igualdade das mulheres. Isso é condizente com a tese de que o Estado, por estar mais protegido de certas formas de silenciamento (discutidas em detalhe na seção 2), é também um ouvinte resistente especialmente eficaz no combate ao discurso danoso.

Referências bibliográficas

- Anderson, L., Haslanger, S., & Langton, R. (2012). Language and race. Em *The Routledge companion to philosophy of language* (pp. 753-767). Routledge.
- Austin, J. L. (1962/1990). *Quando dizer é fazer: Palavras e ação* (D. Marcondes de Souza Filho, Trad.). Artes Médicas.
- Barbosa, E. C. (2023). Code words and (re) framing. *Manuscrito*, 46(3), e-2023.
- Barbosa, E. C., & Fernandes, M. T. (2024). Resistência conversacional e as variedades de contradiscurso. *Principia: An international journal of epistemology*, 28(3), 5.
- Bartlett, J., & Krasodomski-Jones, A. (2015, 19/12/2024). Counter-speech: Examining content that challenges extremism online. <https://www.dangerousspeech.org/libraries/counter-speech-examining-content-that-challenges-extremism-online>
- Benz, A., & Jasinskaja, K. (2017). Questions under discussion: From sentence to discourse. *Discourse Processes*, 54(3), 177-186. <https://doi.org/10.1080/0163853X.2017.1316038>
- Brettschneider, C. (2012). *When the state speaks, what should it say?: How democracies can protect expression and promote equality*. Princeton University Press.
- Brown, A. (2017a). What is hate speech? Part 1: The myth of hate. *Law and philosophy*, 36(4), 419-468.
- Brown, A. (2017b). What is hate speech? Part 2: Family resemblances. *Law and Philosophy*, 36(5), 561-613.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language*

- usage*. Cambridge University Press.
- Caponetto, L., & Cepollaro, B. (2023). Bending as counterspeech. *Ethical theory and moral practice*, 26, 577-593. <https://doi.org/10.1007/s10677-022-10334-4>
- Cepollaro, B., Lepoutre, M., & Simpson, R. M. (2023). Counterspeech. *Philosophy Compass*, 18(1), e12890. <https://doi.org/10.1111/phc3.12890>
- Dorst, K. (2020, Mar 12). The rational question. *The Oxonian Review*: <https://www.oxonianreview.com/articles/the-rational-question>
- Eisterhold, J., Attardo, S., & Boxer, D. (2006). Reactions to irony in discourse: Evidence for the least disruption principle. *Journal of Pragmatics*, 38(8), 1239-1256. <https://doi.org/10.1016/j.pragma.2004.12.003>
- Fraser, R. (2023). How to talk back: Hate speech, misinformation, and the limits of salience. *Politics, Philosophy & Economics*, 22(3), 315-335. <https://doi.org/10.1177/1470594X231167593>
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Gelber, K. (2002). *Speaking back: The free speech versus hate speech debate*. John Benjamins.
- Gelber, K. (2021). Differentiating hate speech: A systemic discrimination approach. *Critical Review of International Social and Political Philosophy*, 24(4), 393-414. <https://doi.org/10.1080/13698230.2019.1576006>
- Goffredo, P., Basile, V., Cepollaro, B., & Patti, V. (2022). Counter-TWIT: An Italian corpus for online counterspeech in ecological contexts. In K. Narang, A. Mostafazadeh D., L. Mathias, B. Vidgen & Z. Talat (Eds.), *WOAH 2022-6th Workshop on Online Abuse and Harms, Proceedings of the Workshop* (pp. 57-66). Association for Computational Linguistics (ACL).
- Grice, H. P. (1957). Meaning. *The Philosophical Review*, 66(3), 377-388.
- Grice, H. P. (1975). Logic and conversation. Em *Speech acts* (pp. 41-58). Brill.
- Haslanger, S. (2010). Ideology, generics, and common ground. Em C. Witt (Ed.), *Feminist metaphysics: Explorations in the ontology of sex, gender and the self* (pp. 179-207). Springer.
- Hornsby, J., & Langton, R. (1998). Free speech and illocution. *Legal Theory*, 4(1), 21-37. <https://doi.org/10.1017/s1352325200000902>
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1420-1436. [ANÁLISIS FILOSÓFICO 46\(1\) - \(mayo 2026\)](https://doi.org/10.1037/0278-</p></div><div data-bbox=)

- 7393.20.6.1420
- Kehler, A., & Rohde, H. (2016). Evaluating an expectation-driven question-under-discussion model of discourse interpretation. *Discourse Processes*, 54(3), 219-238. <https://doi.org/10.1080/0163853X.2016.1169069>
- Kotthoff, H. (2003). Responding to irony in different contexts: On cognition in conversation. *Journal of Pragmatics*, 35(9), 1387-1411, [https://doi.org/10.1016/S0378-2166\(02\)00182-0](https://doi.org/10.1016/S0378-2166(02)00182-0)
- Langton, R. (1993). Speech acts and unspeakable acts. *Philosophy & Public Affairs*, 22(4), 293-330.
- Langton, R. (2010). Epistemic injustice: Power and the ethics of knowing. By Miranda Fricker. Oxford, Oxford University Press, 2007. *Hypatia*, 25(2), 459-464. <https://doi.org/10.1111/j.1527-2001.2010.01098.x>
- Langton, R. (2018). Blocking as counter-speech. Em D. Fogal, D. W. Harris & M. Moss (Eds.), *New work on speech acts* (pp. 144-164). Oxford. <https://doi.org/10.1093/oso/9780198738831.003.0006>
- Lepoutre, M. (2017). Hate speech in public discourse: A pessimistic defense of counterspeech. *Social Theory and Practice*, 43(4), 851-883. <https://doi.org/10.5840/soctheorpract201711125>
- Lepoutre, M. (2019). Can more speech counter ignorant speech. *Journal of Ethics and Social Philosophy*, 16(3), 155-191. <https://doi.org/10.26556/jesp.v16i3.682>
- Lepoutre, M. (2021). *Democratic speech in divided times*. Oxford University Press.
- Lepoutre, M. (2023). Discursive optimism defended. *Politics, Philosophy & Economics*, 22(3), 357-374. <https://doi.org/10.1177/1470594X231179665>
- Lewandowsky, S., Ecker, U., Seifert, C., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106-131. <https://doi.org/10.1177/1529100612451018>
- Lewis, D. (1979). Scorekeeping in a language game. Em R. Bauerle, U. Egli & A. von Stechow (Eds.), *Semantics from different points of view* (pp. 172-187). Springer.
- Maitra, I. (2009). Silencing speech. *Canadian Journal of Philosophy*, 39(2), 309-338. <http://doi.org/10.1353/cjp.0.0050>
- Maitra, I. (2012). Subordinating speech. Em I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 94-120). Oxford Academic.
- Marques, T. (2023). The expression of hate in hate speech. *Journal of Applied Philosophy*, 40(5), 769-787.
- Mason, R. (2021). Hermeneutical injustice. Em *The Routledge Handbook of social and political philosophy of language* (pp. 247-258). Routledge.

- Matsuda, M. J., Iii, C. R. L., Delgado, R., & Crenshaw, K. W. (2018). *Words that wound: Critical race theory, assaultive speech, and the first amendment*. Routledge.
- McGowan, M. K. (2019). *Just words: On speech and hidden harm*. Oxford University Press.
- McGowan, M. K. (en prensa). On constitutively harmful speech, the riskiness of *in situ* counter-speech, and silencing. Em S. Haslanger, L. Schroeter, K. Jones, G. Restall, & F. Schroeter (eds.), *Mind, language, and social hierarchy: Constructing a shared social world*. Oxford University Press.
- Medina, J. (2013). *The epistemology of resistance: Gender and racial oppression, epistemic injustice, and the social imagination*. Oxford University Press.
- Medina, J. (2023). *The epistemology of protest: Silencing, epistemic activism, and the communicative life of resistance*. Oxford University Press.
- Mills, C. (2007). White ignorance. Em S. Sullivan & N. Tuana (Eds.), *Race and epistemologies of ignorance* (pp. 11-38). State University of New York Press.
- Moon, J. D. (2014). Well-ordered society. Em J. Mandle & D. A. Reidy (Eds.), *The Cambridge Rawls Lexicon* (pp. 874-877). Cambridge University Press.
- Orlando, E., & Saab, A. (2019). Términos peyorativos de grupo, estereotipos y actos de habla. *Crítica*, 51(153), 31-58.
- Popa-Wyatt, M. (2023, Jul 27). The challenges of regulating online speech. <https://blog.policy.manchester.ac.uk/posts/2023/07/the-challenges-of-regulating-online-speech/>
- Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6), 1-69. <https://doi.org/10.3765/sp.5.6>
- Roberts, C. (2015). Accommodation in a language game. Em B. Loewer & J. Schaffer (Eds.), *A companion to David Lewis* (pp. 345-366). Wiley Blackwell.
- Romdenh-Romluc, K. (2017). Hermeneutical injustice and the problem of authority. *Feminist Philosophy Quarterly*, 3(3). <https://doi.org/10.5206/fpq/2017.3.1>
- Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, 39, 127-161. [https://doi.org/10.1016/S0065-2601\(06\)39003-X](https://doi.org/10.1016/S0065-2601(06)39003-X)
- Simpson, R. (2013). Un-ringing the bell: McGowan on oppressive speech and

- the asymmetric pliability of conversations. *Australasian Journal of Philosophy*, 91(3), 555-575. <https://doi.org/10.1080/00048402.2012.704053>
- Strawson, P. F. (1964). Intention and convention in speech acts. *The philosophical review*, 73(4), 439-460.
- Tirrell, L. (2012). Genocidal language games. Em I. Maitra & M. K. McGowan (Eds.), *Speech and harm: Controversies over free speech* (pp. 174-221). Oxford University Press.
- Tirrell, L. (2017). Toxic speech: Toward an epidemiology of discursive harm. *Philosophical topics*, 45(2), 139-162.
- Waldron, J. (2012). *The harm in hate speech*. Harvard University Press.

Recebido em 20 de dezembro de 2024; revisado em 5 de agosto de 2025; aceito em 21 de agosto de 2025.