

QUANDO É QUE OS ELOGIOS E O HUMOR PODEM SER CONSIDERADOS DISCURSO DE ÓDIO? UMA PERSPETIVA DOS GRUPOS-ALVO EM PORTUGAL

Cláudia Silva

Laboratory for Robotics and Engineering Systems, Interactive Technologies Institute, Lisboa, Portugal/Departamento de Engenharia Informática, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal
Concetualização, curadoria dos dados, análise formal, investigação, metodologia, administração do projeto, visualização, redação do rascunho original, redação – revisão e edição

Paula Carvalho

Instituto de Engenharia de Sistemas e Computadores — Investigação e Desenvolvimento, Lisboa, Portugal/
Departamento de Engenharia Informática, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal
Concetualização, análise formal, aquisição de financiamento, investigação, administração do projeto, visualização, redação do rascunho original, redação – revisão e edição

RESUMO

Este artigo apresenta um estudo qualitativo acerca do discurso de ódio (DO) a partir das perspetivas de três comunidades minoritárias representativas em Portugal: afrodescendentes, pessoas de etnia Roma (ou cigana) e LGBTQ+. A investigação empírica envolveu a realização de três grupos focais, um com cada comunidade ($n=17$), tendo como principal objetivo investigar o modo como os membros desses grupos percebem e experienciam o DO no contexto social e geopolítico português. Os resultados deste estudo mostram que o DO indireto (ou encoberto) pode ser mais prejudicial do que o DO direto (ou explícito), pois inibe os mecanismos de reação por parte dos alvos, o que leva à perpetuação de sistemas de opressão e decadência social. Além disso, os dados revelam que as formas encobertas de DO manifestam-se muitas vezes como elogios e humor, sendo comuns no local de trabalho, nos meios de comunicação e no ambiente online. De facto, o elogio e o humor são estratégias comunicativas, ancoradas em estereótipos positivos ou negativos, que contribuem para a normalização do DO na sociedade portuguesa. Neste sentido, esta investigação alerta para a importância de identificar formas encobertas de DO e de desenvolver estratégias eficazes, como contranarrativas, para combatê-las. Por fim, destaca a necessidade de conceber sistemas de deteção automática capazes de identificar o DO indireto, dada a sua prevalência nas plataformas digitais.

PALAVRAS-CHAVE

discurso de ódio encoberto, estereótipos, elogio, humor, alvos do discurso de ódio

WHEN CAN COMPLIMENTS AND HUMOUR BE CONSIDERED HATE SPEECH? A PERSPECTIVE FROM TARGET GROUPS IN PORTUGAL

ABSTRACT

This article presents a qualitative study that examines hate speech (HS) from the perspectives of the most representative minority communities in Portugal, namely Afro-descendants, Roma, and LGBTQ+. The empirical research employed three focus groups, each with a representation of these communities ($n=17$), with the primary goal of investigating how members of these groups perceive and experience HS within the Portuguese social and geopolitical context. The results indicate that covert HS can be more detrimental than overt HS since the targeted individuals are often caught off guard and unsure how to respond, contributing to the continuation of systems of oppression and social decline. Additionally, our qualitative data show that covert forms of HS often manifest in the form of compliments and humour, common in various settings, such as the workplace, the media, and the online environment. These devices are communicative strategies anchored in both positive and negative stereotypes and may lead to the normalization of HS in Portuguese society. In this regard, this research draws attention to the significance of identifying covert forms of HS and devising strategies, such as counter-speech, to debunk them. Furthermore, it highlights the need for automatic detection of covert HS since these forms are prevalent in digital platforms.

KEYWORDS

covert hate speech, stereotypes, compliment, humour, hate speech targets

1. INTRODUÇÃO

O discurso de ódio (DO) não é um fenómeno exclusivo das redes sociais ou da internet. No entanto, vários estudos científicos têm fornecido evidências de que o DO é bastante prevalente no ambiente online devido ao conteúdo gerado pelos utilizadores, especialmente em plataformas como o Facebook, Twitter, YouTube e TikTok a nível global (Breazu & Machin, 2022b; Rieger et al., 2021). De facto, algumas pesquisas demonstraram que a incidência de DO tem aumentado constantemente em toda a Europa nos últimos anos (Bakowski, 2022), incluindo em Portugal, afetando particularmente populações e comunidades vulneráveis, tais como as comunidades afrodescendente, migrante e cigana ou Roma, sendo esta última um dos grupos-alvo mais afetados pelas mensagens de racismo e xenofobia, nomeadamente no paradigma português (Reynders, 2022; Silva, 2021).

O aumento do DO nas redes sociais explica a razão pela qual os académicos têm desenvolvido métodos para automatizar a sua deteção (Fortuna & Nunes, 2018; Schmidt & Wiegand, 2019). Embora a deteção automática de DO online possa ser uma forma de mitigar este fenómeno, esta abordagem apresenta ainda muitas limitações. Primeiro, o DO é frequentemente confundido com outros conceitos subjetivos, como linguagem “abusiva”, “tóxica”, “perigosa”, “ofensiva” ou “agressiva”, levando à criação de recursos linguísticos e modelos de linguagem heterogéneos (Poletto et al., 2021). Além disso, a

maioria dos sistemas de deteção de DO não considera variáveis mais complexas, como a *prática social*, ou seja, o contexto social e cultural subjacente à produção e disseminação de DO (Assimakopoulos et al., 2017). Em segundo lugar, grande parte da investigação levada a cabo neste âmbito assenta na deteção do DO direto (ou explícito) em vez de formas indiretas (ou encobertas) de ódio (Baider & Constantinou, 2020; Bhat & Klein, 2020; Jha & Mamidi, 2017). O DO encoberto pode materializar-se de diversas formas, por exemplo, explorando termos duvidosos para ocultar o conteúdo comunicativo negativo ou através das *affordances* tecnológicas, tal como colocar um “gosto” num *post* que contém um *link* que redireciona para um website anti-imigração externo à rede social (Ben-David & Fernández, 2016). Este fenómeno é também frequentemente mascarado através do recurso a diferentes estratégias retóricas e discursivas, tais como ironia, sarcasmo e humor (Baider & Constantinou, 2020; Billig, 2001; Dynel, 2018), eufemismos (Magu & Luo, 2018) e perguntas retóricas (Albelda Marco, 2022; Krobová & Zàpotocký, 2021), tornando o seu reconhecimento mais difícil. Embora a maioria destas estratégias explore os estereótipos associados aos grupos-alvo (Buturoiu & Corbu, 2020; Chovanec, 2021; ElSherief et al., 2021), a pesquisa sobre a relação entre a manifestação do DO encoberto e os estereótipos associados aos grupos-alvo é ainda escassa.

Em contraste, o DO direto ou explícito é aparentemente mais fácil de ser reconhecido porque geralmente contém palavras e expressões ofensivas, agressivas e explicitamente discriminatórias. No entanto, fora do contexto, a presença desses termos não é, por si só, uma garantia de que os enunciados em que surgem possam ser automaticamente classificados como DO. Tal como refere Baider (2022), para além do nível locutório (i.e., a emissão da mensagem propriamente dita), é necessário considerar outras dimensões dos “atos de fala”, nomeadamente os níveis ilocutórios (i.e., a intenção do emissor da mensagem) e perlocutórios (i.e., o efeito da mensagem no recetor). Por outro lado, embora algumas formas encobertas de DO possam não ser identificáveis ao nível da sua materialização propriamente dita, é possível alterar essa perceção se considerarmos os níveis ilocutórios e perlocutórios.

No nosso estudo, concentrar-nos-emos principalmente no nível perlocutório, uma vez que estamos interessadas em compreender o impacto do DO nos grupos-alvo, tendo especificamente em conta a prática social portuguesa. Com este objetivo específico em mente e assumindo uma perspetiva do campo da comunicação, consideramos crucial abordar diretamente os destinatários do DO para identificar o impacto das mensagens (Paz et al., 2020), procurando dar resposta às seguintes questões de investigação:

1. Como é que os grupos-alvos de DO em Portugal percecionam este fenómeno?
2. Quais são as manifestações mais nocivas de DO na perspetiva dos grupos-alvo do discurso do ódio?

Para responder a estas questões, partindo das premissas enunciadas nesta introdução, adotámos uma abordagem qualitativa, que assenta na criação de três grupos focais (GF) que envolvem membros de comunidades frequentemente vítimas de DO em Portugal (Silva, 2021), nomeadamente as comunidades afrodescendente, Roma e LGBTQ+.

Como contribuição geral, apresentamos evidências empíricas de que o DO encoberto é mais prejudicial do que o DO direto ou explícito. Mais especificamente, examinamos o modo como o DO encoberto se pode manifestar como elogio e humor com base nas narrativas das experiências vivenciadas, relatos e percepções dos grupos-alvo do DO. Embora o humor tenha sido amplamente estudado na literatura sobre o DO, os elogios representam a nossa principal contribuição para os estudos e compreensão do DO.

2. CONCEITOS RELEVANTES

2.1. DISCURSO DE ÓDIO

Definir DO é uma tarefa complexa e desafiante, pois não há consenso universal sobre o seu significado (Brown, 2017; MacAvaney et al., 2019). De acordo com a última recomendação do Conselho da Europa (CM/Rec/2022/16), o DO é genericamente definido como

todos os tipos de expressão que incitam, promovem, espalham ou justificam a violência, o ódio ou a discriminação contra uma pessoa ou grupo de pessoas, ou que os denigrem, em função de suas características pessoais reais ou atribuídas, tais como “raça”, cor, língua, religião, nacionalidade, origem nacional ou étnica, idade, deficiência física, sexo, identidade de género e orientação sexual. (Council of Europe, 2022, Apêndice, Ponto 1.1.)

Nesta definição, há vários aspetos centrais a reter: o DO pode materializar-se sob diferentes formas (e.g., verbal vs não verbal; direto vs indireto); o autor do DO tem a intenção de atacar um grupo ou um dos seus membros; e, por fim, o grupo-alvo é atacado não por via de um comportamento ou ação específica, mas devido às suas características identitárias ou estatuto num determinado contexto social. A acrescentar à definição do Conselho da Europa, importa salientar que o DO está frequentemente assente num conjunto de estereótipos negativos, utilizados para discriminar estes grupos e, assim, reforçar a distância social entre os grupos dominantes (“endogrupos”) e os grupos vulneráveis (“exogrupos”; De Cillia et al., 1999; van Dijk, 1992;). Este aspeto é evidente na definição proposta pela *Recomendação de Política Geral N.º 15* da Comissão Europeia contra o Racismo e a Intolerância:

uso de um ou mais modos particulares de expressão – nomeadamente, a defesa, a promoção ou a incitação à discriminação, ódio ou desrespeito a uma pessoa ou grupo de pessoas, assim como qualquer forma de assédio, insulto, estereótipo negativo, estigmatização ou ameaça a essa pessoa ou grupo de pessoas, bem como qualquer forma de justificação de todos estes modos de expressão – que sejam baseados numa lista não exaustiva de características pessoais ou estatutos que incluem “raça”, cor, língua, religião ou crença, nacionalidade ou origem étnica ou nacional, bem

como ascendência, idade, deficiência, sexo, género, identidade de género e orientação sexual. (European Commission against Racism and Intolerance, 2016, p. 16)

De uma maneira geral, os principais aspetos destas definições foram recuperados nas diversas explicações formais existentes do DO (Brown, 2017; Silva, 2021). Por exemplo, Richardson-Self (2018) resume estes pontos comuns da seguinte forma:

em primeiro lugar, o discurso do ódio é muitas vezes descrito como essencialmente hostil. Em segundo lugar, entende-se o discurso do ódio como aquele que *faz* determinadas ações: silencia, demoniza, deprecia, humilha, intimida, incita à violência, discrimina, difama, degrada, persegue, ameaça, e outros atos semelhantes. Finalmente, o discurso do ódio é no geral compreendido como um comportamento expressivo correspondente a traços identitários percebidos ou atribuídos a um determinado grupo. Esses traços incluem raça, religião, orientação sexual, deficiência física, género, e (cada vez mais) a identidade de género. (p. 256)

Com o objetivo de estabelecer o nosso posicionamento teórico, destacamos as dimensões de opressão e poder social que subjazem ao conceito de DO, documentadas em diversos estudos, tais como os de van Dijk (1992), De Cillia et al. (1999), Matsuda (1989) e Matamoros-Fernández e Farkas (2021), entre outros. Estes autores incluem as relações de poder assimétricas, a discriminação e o controlo, bem como a opressão histórico-social, incluindo o racismo sistémico, como aspetos fundamentais do DO enquanto fenómeno sociopolítico.

Neste artigo, concentramo-nos na expressão verbal do DO, nomeadamente o DO encoberto, e, sobretudo, no impacto que este tem nos grupos-alvo.

2.1.1. DISCURSO DE ÓDIO ENCOBERTO

Estudos recentes têm afirmado que “hoje a maior parte do discurso de ódio é encoberto e expressa racismo, sexismo, homofobia” (Baider, 2022, p. 2353). Uma explicação para a prevalência do DO encoberto é a implementação do Código de Conduta de 2016, que limita/restringe o DO online nas redes sociais¹ (Baider, 2022). Como forma de contornar esta questão, os utilizadores têm procurado utilizar estratégias comunicativas subtis para espalhar o ódio sem serem detetados e punidos.

O DO direto ou explícito é expresso literalmente e inclui, com frequência, epítetos raciais e linguagem racista, preconceituosa ou derogatória, amplamente considerada como expressão pública inaceitável (Daniels, 2008). Por outro lado, o DO encoberto também conhecido como discurso “camuflado”, “subtil”, “soft” ou “implícito” (Assimakopoulos et al., 2017; Daniels, 2008), utiliza normalmente linguagem figurativa

¹ “Para prevenir e combater a propagação do discurso de ódio ilegal online, em maio de 2016, a Comissão celebrou um acordo com o Facebook, Microsoft, Twitter e YouTube um ‘Código de conduta para combater o discurso de ódio ilegal online’” (European Commission, s.d., para. 1).

e pode ser interpretado como uma forma de disfarçar ou camuflar mensagens com motivações racistas ou outros tipos de DO encoberto. Por exemplo, num estudo comparativo de comentários online de gregos e cipriotas gregos, Baider e Constantinou (2020) demonstraram que estratégias retóricas específicas, como a ironia verbal e o humor, são formas eficazes de difundir o DO encoberto.

Foram identificadas na literatura outras estratégias de comunicação para expressar o DO encoberto, tais como “dog-whistling”, uma forma de comunicação simbólica que envolve símbolos, palavras-chave e linguagem codificada, aparentemente inócua para o público geral, porém carregada de significado implícito, sendo compreendida somente por uma parte da audiência (normalmente, um grupo-alvo). Essa técnica, que tem sido usada para contornar a eliminação de DO online direto através da moderação automática, possibilita a disseminação de DO encoberto através de imagens, *memes*, eufemismos, símbolos e linguagem camuflada (Bhat & Klein, 2020). Neste artigo, analisamos especificamente a manifestação do DO encoberto com recurso ao elogio e ao humor, ambos ancorados em estereótipos positivos e negativos.

Estereótipos e discurso de ódio. Na psicologia social, “geralmente, os estereótipos são definidos como crenças acerca de certos grupos, avaliações ou atitudes preconceituosas em relação a um determinado grupo e comportamentos discriminatórios que sistematicamente beneficiam ou prejudicam um grupo social” (Jussim & Rubinstein, 2012, p. 1). Os estereótipos também são definidos como generalizações acerca de grupos sociais, em que certas características são atribuídas a todos os membros do grupo, ignorando as diferenças individuais dentro do grupo (Seiter, 1986). Nos estudos da comunicação, os estereótipos são frequentemente abordados no contexto da representação mediática (e.g., McInroy & Craig, 2017; Trebbe et al., 2017) e são muitas vezes vistos como mecanismos discursivos que servem para manter o *status quo* e justificar desigualdades e diferenças sociais, servindo assim um propósito ideológico (Seiter, 1986).

Embora a articulação entre o DO encoberto e os estereótipos não tenha sido amplamente explorada, alguns investigadores defendem que a utilização de estereótipos negativos deve ser incluída na definição de DO (ver Haladzhun et al., 2021). De facto, os estereótipos negativos, incluindo os raciais, são frequentemente usados nas redes sociais para depreciar ou humilhar membros de uma comunidade vulnerável com base em generalizações negativas e falaciosas (Paz et al., 2020; Sanguinetti et al., 2018). Por exemplo, as principais representações das comunidades ciganas ou Roma giram em torno dos estereótipos de pessoas que agem perpetuamente fora da lei, vivem em condições degradantes e optam voluntariamente por depender de benefícios sociais, em vez de lutarem pelo seu próprio sustento (Buturoiu & Corbu, 2020).

Adicionalmente, e em consonância com os objetivos gerais deste artigo, é de salientar que os estereótipos sociais também têm sido considerados no âmbito da investigação em deteção automática de DO online, uma vez que pesquisas anteriores evidenciaram que os classificadores automáticos de DO aprendem e reproduzem estereótipos sociais semelhantes aos humanos (Davani et al., 2021).

Neste trabalho, argumentamos que os estereótipos estão subjacentes ao DO encoberto, especialmente sob a forma de elogios e humor, tendo como base conceptual o modelo de conteúdo de estereótipo fornecido por Fiske et al. (2002). Este modelo propõe que as percepções sociais e estereótipos são formados ao longo de duas dimensões: “afeição” (*warmth*, no seu original em inglês), onde se incluem, por exemplo, os conceitos de confiança, simpatia e bondade, e “competência”, onde se incluem, por exemplo, os conceitos de capacidade, assertividade e inteligência. Alguns exemplos ilustrativos das dimensões psicológicas das formações de estereótipos de grupo são oferecidos por Davani et al. (2021):

na sociedade contemporânea dos Estados Unidos, as pessoas cristãs e heterossexuais são comumente percebidas como altamente capazes de expressar afeição e competência, o que leva as pessoas a expressarem orgulho e admiração por estes grupos sociais. Já as pessoas asiáticas e aquelas ricas são frequentemente estereotipadas como competentes, mas não afetuosas. Por outro lado, as pessoas idosas e aquelas portadoras de deficiência física são estereotipadas como afetuosas, mas não competentes. Por fim, as pessoas que moram na rua, e os afrodescendentes de baixos rendimentos, juntamente com os árabes, são estereotipados como frios e incompetentes (Fiske et al., 2007). (p. 6)

Elogios como mensagens nocivas. A noção de que os elogios podem ser interpretados como negativos não é nova. Pelo contrário, vários estudos em psicologia social têm analisado os elogios como uma forma de ofensa aos grupos-alvo (Alt et al., 2019; Czopp, 2008; Czopp & Monteith, 2006; Siy & Cheryan, 2013). Embora a pesquisa tradicional se tenha concentrado na prevalência de estereótipos negativos, estudos mais recentes têm investigado o modo como estereótipos positivos sobre grupos marginalizados conduzem a elogios estereotipados (e.g., Alt et al., 2019; Czopp et al., 2015). A literatura apresenta exemplos de estereótipos positivos de grupos sociais, como a ideia de que os asiáticos são habilidosos em matemática e, por exemplo, os afro-americanos são vistos como atléticos ou musicalmente talentosos, ou ainda como pessoas sociáveis e sexuais. Argumentamos que os estereótipos positivos devem ser considerados no contexto do DO por três razões principais. Em primeiro lugar, os estereótipos são restritivos porque se baseiam apenas na filiação ao grupo, sem considerar as diferenças individuais (Czopp, 2008). Em segundo lugar, vários estudos evidenciaram que, apesar de serem elogios, os estereótipos positivos podem ter efeitos negativos (ainda que não intencionais) sobre os alvos desses estereótipos (Czopp & Monteith, 2006). Em terceiro lugar, os estereótipos positivos complementam-se frequentemente com estereótipos negativos para desacreditar ou diminuir os membros dos grupos-alvo. As mulheres, por exemplo, são vistas como calorosas, mas fracas; as pessoas asiáticas como competentes, mas frias; e os negros como atléticos, mas pouco inteligentes (Czopp, 2008). Na verdade, Czopp e Monteith (2006) mostram que as pessoas que elogiam negros pela suposta capacidade atlética e musical (estereótipos positivos) também são propensas a depreciá-los pela suposta preguiça e criminalidade (estereótipos negativos).

O que é relevante enfatizar na literatura sobre elogios estereotipados para o presente artigo é a evidência científica de que as pessoas que são alvos desses “elogios” e aqueles que elogiam (estereotipistas) têm percepções distintas dos estereótipos positivos. Embora os recetores dos elogios possam considerá-los inaceitáveis e reagir negativamente, as percepções do grupo dominante (e.g., homens em relação às mulheres, brancos em relação aos negros e asiáticos) podem ser que esses elogios são, de facto, benéficos, “pois refletem um enaltecimento genuíno dos pontos fortes do grupo” (Czopp, 2008, p. 414).

Por fim, entendemos que esses supostos elogios são prejudiciais para os grupos-alvo, porque como Siy e Cheryan (2013) argumentam, os estereótipos positivos impõem uma identidade social aos seus recetores e fazem com que se sintam despersonalizados ou “aglutinados” com outras pessoas do seu grupo social, pelo estereotipista.

Humor e discurso de ódio. O humor é um recurso retórico complexo que pode ser classificado em diversas categorias, apresentando diferentes funções. Dentre outras categorias, a literatura distingue o “humor afiliativo” do “humor agressivo” (Martin et al., 2003). De acordo com os autores, enquanto o primeiro pode ser visto como uma forma de humor não hostil, que normalmente tem como objetivo promover o sentimento positivo e facilitar as relações, o segundo baseia-se recorrentemente no sarcasmo, ironia e sátira, com o objetivo de depreciar os grupos-alvo e de os expor frequentemente ao ridículo. Estes dois tipos de humor estão entre as categorias mais predominantes e exploradas em conteúdos multimodais (particularmente *memes*), partilhados em diferentes plataformas de redes sociais, como o Facebook (Taecharungroj & Nueangjamnong, 2015). Entre os *memes* agressivos, é de salientar a profusão de *memes* racistas, comumente utilizados por comunidades marginais da web (Zannettou et al., 2018).

Embora alguns investigadores tenham enfatizado as funções positivas do humor na diminuição da distância social entre grupos, vários estudos têm demonstrado o caráter depreciativo do humor na estigmatização de grupos vulneráveis percebidos como social e culturalmente inferiores (Breazu & Machin, 2022a). De facto, o humor agressivo pode ser um recurso retórico eficaz para exprimir uma hostilidade dissimulada, porque é geralmente visto como uma piada (Billig, 2001), protegendo o autor de acusações de intenção discriminatória (Woodzicka et al., 2015). No entanto, apesar de algumas formas de mensagens humorísticas que transmitem hostilidade poderem não ser identificadas como tal pela maioria dos utilizadores, os grupos-alvo visados podem sentir-se especialmente afetados pelas mesmas (Schmid et al., 2022). Como demonstrado por Woodzicka et al. (2015), as piadas racistas são de facto muitas vezes vistas como uma forma de manifestação de racismo pelas comunidades atingidas.

O humor agressivo é, pois, expresso sem considerar o seu potencial impacto sobre os outros, e é frequentemente baseado em estereótipos negativos, sendo o alvo um objeto de ridicularização. Com base numa análise discursiva das piadas antimuçulmanas e antisemitas, Weaver (2013) demonstrou que os estereótipos e a inferiorização são usados de forma combinada ou independente para formar imagens “aceitáveis” e inclusivas de piadas para mascarar o racismo. Estas estratégias promovem a construção do

“exogrupo” como um grupo social inferior (Breazu & Machin, 2022a) e contribuem para reforçar e normalizar os estereótipos negativos associados às comunidades atacadas.

3. MÉTODOS

Este artigo adota uma abordagem qualitativa para compreender o modo como os grupos-alvo do DO em Portugal (particularmente as comunidades afrodescendentes, ciganas ou Roma e LGBTQ+) conceptualizam o DO. O estudo, centrado na recolha e análise das perceções e experiências vivenciadas pelos grupos-alvo mencionados, foi desenvolvido no âmbito do projeto de investigação financiado intitulado HateCovid-19.PT, que tem por objetivo último o desenvolvimento de um protótipo para detetar automaticamente o DO nas redes sociais em português europeu. Para esse fim, foram realizados três GF, no verão de 2021, cada um envolvendo uma das comunidades acima mencionadas ($n=17$). O GF é um método apropriado para conduzir esta pesquisa, pois ajuda a compreender os fenómenos de DO a partir da experiência dos grupos-alvo, “cujas vozes são frequentemente marginalizadas na sociedade em geral” (Given, 2008, p. 352).

Todas as sessões ocorreram online, no Zoom, e foram gravadas. O GF com afrodescendentes decorreu a 19 de maio de 2021, com nove participantes (duração de 03:03:24). O GF com LGBTQ+ ocorreu a 5 de julho de 2021, com cinco participantes (duração de 02:28:57). O GF com a comunidade cigana ocorreu a 21 de julho de 2021, com três participantes (duração de 02:01:05). Para fins de organização, os participantes são citados ao longo do artigo através da letra “P” seguida de um número (identificador do participante) e do GF em que participaram (e.g., P1-Afro, P1-LGBTQ+, P1-Roma). O nosso objetivo inicial era ter entre seis e nove participantes em cada GF. No entanto, esse critério não foi integralmente cumprido, devido a vários desafios de recrutamento relacionados com os participantes LGBTQ+ e pessoas de etnia cigana durante o período estabelecido para o trabalho empírico. Especificamente, no caso do GF Roma, tivemos dificuldade em encontrar participantes disponíveis para partilharem as suas opiniões e experiências vividas abertamente com as pesquisadoras. O desafio de envolver os participantes de etnia cigana na pesquisa qualitativa já foi, aliás, reportado por outros estudiosos (Condon et al., 2019). As barreiras ao recrutamento podem ser várias, incluindo a desconfiança e o receio de prejudicar as crenças culturais da comunidade, especialmente em relação a temas sensíveis, como é o caso do DO. Além disso, os membros desta comunidade têm enfrentado uma história de discriminação racial, perseguição e exclusão política e social em Portugal (Cádima et al., 2020; Comissão para a Igualdade e Contra a Discriminação Racial, 2020) e em toda a Europa (Delcour & Hustinx, 2017; Maeso, 2021), o que pode justificar a recusa em participar no âmbito de estudos de natureza qualitativa, que implicam a partilha de experiências vividas.

4. PROCEDIMENTOS

Todos os GF seguiram o mesmo protocolo e foram conduzidos pela mesma investigadora, que se autoidentifica como uma mulher negra, cisgénero, heterossexual e migrante, que se mudou do Brasil para Portugal. O GF foi transcrito por uma jovem estudante de mestrado em migração, também ela uma mulher negra. Este nível de partilha sobre o posicionamento social e identitário é relevante para a prática da reflexividade, uma vez que vários estudos atuais sobre ética e conceção de novas tecnologias têm recomendado esta prática (Bardzell & Bardzell, 2011; Schlesinger et al., 2017). Mais relevante, esta prática de reflexão sobre o lugar de fala da autora revela que uma das investigadoras faz parte dos grupos-alvo do DO em Portugal, facto que poderá ter contribuído para a construção de *rappport* com os participantes. Para completar este exercício de “posicionalidade”, convém acrescentar que a segunda autora se autoidentifica como uma mulher portuguesa branca, cisgénero, heterossexual.

Nos GF, a investigadora começou por apresentar o projeto em que se insere este estudo (*HateCovid-19.PT*; <https://hate-covid.inesc-id.pt>), introduzindo a equipa de investigação e o objetivo geral do projeto: contribuir para a análise e deteção automática de DO online no contexto português. Em seguida, foi feita uma breve descrição dos objetivos específicos do projeto: (a) investigar estratégias linguísticas e retóricas subjacentes ao DO direto (explícito) ou indireto (implícito ou encoberto); (b) construir um *corpus* a partir das redes sociais para estudar o DO online e fenómenos estreitamente relacionados, tais como contranarrativas e discurso ofensivo; e (c) desenvolver um protótipo para deteção automática de DO.

Após esta introdução, a investigadora convidou os participantes a iniciar a discussão pedindo-lhes que se identificassem (relativamente à identidade social) e tentassem formular a forma como os outros os viam, particularmente no contexto português. Em seguida, dada a diversidade de definições associadas ao conceito de DO, que podem variar de acordo com a área de estudo, o contexto geopolítico e a perspetiva dos atores envolvidos, a investigadora pediu explicitamente aos participantes que elaborassem este conceito, trazendo o seu ponto de vista e partilhando com o grupo a sua própria experiência vivida. Especialmente no que diz respeito ao DO direto, que frequentemente se baseia numa linguagem ofensiva e depreciativa, os participantes foram convidados a partilhar exemplos de palavras ou expressões de ódio que podem ser mais impactantes ou prejudiciais, de acordo com a sua compreensão. Esta informação é especialmente útil para questionar o uso de abordagens baseadas em léxico para identificar DO online automaticamente. Na verdade, a maioria dos léxicos existentes são repositórios de palavras e expressões ofensivas, insultuosas e depreciativas, que podem ser interpretadas de forma diferente dependendo do contexto linguístico e pragmático em que ocorrem. Além disso, esses recursos geralmente não são validados pelos potenciais grupos-alvo a que essas palavras se aplicam, sendo difícil avaliar o seu grau de precisão e cobertura. Em relação ao DO encoberto, procuramos compreender a perspetiva dos grupos-alvo sobre este fenómeno, abordando aspetos relacionados com a sua materialização e severidade em comparação com o DO direto. Uma vez que o humor tem sido identificado na literatura

como uma estratégia retórica eficaz para divulgar o DO encoberto, esta investigação também procurou analisar as perspetivas das comunidades-alvo sobre este tópico. A discussão foi acompanhada pela partilha de tweets contendo DO, como pontos de partida para a conversa (ver Figura 1, Figura 2 e Figura 3). Devido aos procedimentos éticos, tais *tweets* foram anonimizados, e os seus respetivos links não são apresentados, como se vê abaixo.



Figura 1. Exemplo de tweet usado no grupo focal com afrodescendentes

Fonte. Twitter

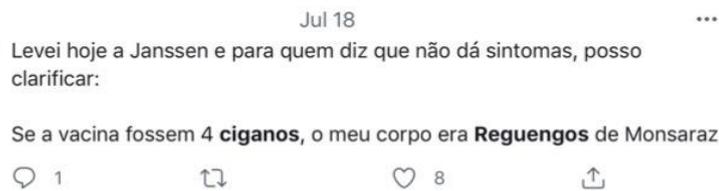


Figura 2. Exemplo de tweet usado no grupo focal com pessoas de etnia Roma

Fonte. Twitter



Figura 3. Exemplo de tweet usado no grupo focal com pessoas LGBTQ+

Fonte. Twitter

5. PARTICIPANTES

Os participantes foram recrutados por meio das redes de contacto da equipa de investigação conforme os requisitos estabelecidos: (a) autoidentificarem-se como pessoas negras ou afrodescendentes, como pessoas de etnia cigana ou Roma ou como pessoas LGBTQ+ e (b) serem capazes de falar sobre o DO a partir das suas próprias experiências.

5.1. AFRODESCENDENTES

Todos os participantes se autoidentificam como negros ou afrodescendentes. Quatro são mulheres (cisgénero), dois são homens (cisgénero), e três identificam-se como *queer*, género não-normativo. A faixa etária varia entre 19 e 51 anos de idade, com uma média de 34,8 anos. A maioria dos participantes ($n=8$) trabalha na indústria dos média, quer nos média tradicionais, nomeadamente como jornalistas em canais de televisão nacionais ($n=2$), quer em meios digitais, nomeadamente como produtores de conteúdo e entrevistadores em canais centrados nos afrodescendentes em Portugal ($n=3$), como produtores de conteúdo/influenciadores nas redes sociais ($n=1$), como atores de cinema/teatro ($n=1$) ou como jornalistas *freelancers* ($n=1$). Todos eles utilizam as redes sociais para divulgar o seu trabalho e se envolver com as suas audiências. À exceção de um, que se mudou para Portugal como adulto (P7-Afro), todos os participantes nasceram em Portugal ou mudaram-se de um país de língua portuguesa quando eram ainda crianças.

5.2. LGBTQ+

Os participantes deste estudo autoidentificam-se como (a) mulher transexual, branca; (b) homem cisgénero (pansexual), branco; (c) *queer*, negra, e mulher lésbica; (d) branca, não binária e não heterossexual; e (e) não binária. Todos eles estão envolvidos em organizações que coordenam direitos e ativismo LGBTQ+ em Portugal. A maioria mantém uma presença ativa nas redes sociais, como o Facebook e o Twitter. Por razões de privacidade dos participantes, os nomes das organizações não são divulgados neste estudo. As idades dos participantes variam entre 22 e 49 anos. A idade média é de 34,6 anos.

5.3. PESSOAS DE ÉTNIA CIGANA OU ROMA

Todos os participantes se autoidentificam como pessoas de etnia cigana, sendo dois do sexo feminino com cerca de 30 anos de idade, e um do sexo masculino com 22 anos de idade. Assim como outros participantes deste estudo, estão envolvidos em organizações relacionadas com o ativismo cigano, como mediadores culturais em centros de emprego ou educação. Neste caso, todos eles se conheciam previamente.

6. ANÁLISE DE DADOS

A transcrição dos GF gerou um documento com 71 páginas de dados textuais, com um total de 38.719 palavras. Para extrair resultados temáticos relevantes e inovadores deste conjunto de dados, utilizamos uma abordagem mista, o que significa que analisamos os dados de forma dedutiva e indutiva. A análise temática é uma técnica amplamente utilizada para identificar, analisar e relatar padrões dentro de dados qualitativos (Braun & Clarke, 2006). As principais características deste método são a natureza iterativa do processo e as discussões entre investigadores em paralelo com a revisão bibliográfica. Neste sentido, as autoras discutiram os códigos e os resultados temáticos várias vezes.

Embora a análise de dados tenha gerado uma multiplicidade de códigos e temas, nem todos acrescentaram novas perspetivas ao estudo do DO. Com isto em mente, a nossa análise de dados foi refinada e orientada pela revisão da literatura e pelas questões de investigação anteriormente mencionadas.

Em primeiro lugar, familiarizamo-nos com os dados. A primeira autora, que conduziu os GF, leu todas as transcrições, atribuindo manualmente códigos indutivamente aos dados textuais, sem quaisquer condicionalismos teóricos ou objetivos específicos em mente. A investigadora iterou este processo, procurando aspetos específicos nas transcrições que pudessem responder às perguntas feitas no GF, e escreveu um resumo de cada GF para ser discutido com a segunda autora.

Como segundo passo, geramos códigos e temas alinhados com os objetivos da investigação e perguntas de investigação. Com base na discussão entre as duas autoras, e numa segunda iteração de leitura das transcrições e códigos iniciais, a primeira autora inseriu todos os dados no software NVivo e codificou qualitativamente as transcrições, procurando a prevalência de resultados nos GF. Nesta fase, é de ressaltar que a relevância dos resultados temáticos não se baseou em medidas quantificáveis, mas sim na captura de algo importante que pudesse contribuir para responder às nossas perguntas de investigação (Braun & Clarke, 2006).

Adicionalmente, para aplicar uma abordagem mista (raciocínio dedutivo e indutivo) na análise de dados, procurámos nos dados aspetos relacionados com os objetivos do projeto de investigação, tais como investigar estratégias linguísticas e retóricas subjacentes ao DO direto (explícito) ou indireto (implícito ou encoberto).

Após este processo, foram selecionados seis temas abrangentes: (a) *o DO encoberto manifesta-se como um elogio* (presente em todos os grupos-alvo); (b) *o humor é considerado como DO encoberto* (presente em todos os grupos-alvo); (c) *o DO é legitimado pelos principais meios de comunicação social tradicionais* (não-consensual); (d) *o DO online é considerado mais prejudicial devido ao seu alcance e impacto coletivo* (não-consensual); (e) *o DO aumentou devido ao crescimento dos partidos de extrema-direita na Europa e em Portugal*; e (f) *a filtragem das palavras é complexa, comprometendo a deteção automática do DO*.

Como terceiro passo, procedemos à revisão e aperfeiçoamento dos nossos temas. Verificámos que o primeiro e o segundo tema (a, b) estavam estreitamente relacionados, uma vez que ambos os fenómenos são suportados por estereótipos e são entendidos pelos participantes como formas eficazes de encobrir o DO. Excluímos o terceiro tema (c), pois esta descoberta está representada na discussão sobre humor. Embora relevantes, o quarto e o quinto temas (d, e) foram apenas abordados por alguns participantes e não estavam diretamente relacionados com as nossas perguntas de investigação. O sexto tema (f) foi elaborado para reconhecer as opiniões dos participantes sobre o objetivo geral da automatização da deteção de DO, mas foi retirado da lista final de resultados temáticos porque este aspeto não foi sistematicamente abordado nem discutido exaustivamente em todos os GF.

Com base nesta fundamentação, as duas autoras reviram os temas e memorandos e, após discussão sobre as questões de investigação, chegaram a dois temas-chave finais,

tendo em consideração a sua relevância e o consenso gerado entre os grupos-alvo: o *DO encoberto manifestado como elogio* e *DO encoberto disfarçado de humor*. Além disso, estes temas foram os mais inovadores a nível teórico, revelando como os grupos-alvo conceitualizam o DO, nomeadamente as formas encobertas. Ao contrário do humor, que foi codificado dedutivamente, os elogios foram codificados indutivamente, uma vez que esta descoberta emergiu em todos os GF. Reconhecendo a novidade de abordar os elogios como uma estratégia eficaz subjacente ao DO encoberto, um aspeto frequentemente negligenciado na literatura sobre DO, decidimos analisar este resultado em profundidade.

7. RESULTADOS

7.1. PRIMEIRO TEMA: DISCURSO DE ÓDIO DISFARÇADO DE ELOGIO

Ao contrário do que se observa no DO encoberto, os participantes neste estudo desenvolveram frequentemente estratégias para lidar com o DO direto. Por um lado, os enunciados de DO direto, tais como os ilustrados nos exemplos a seguir apresentados, são considerados práticas diárias por estes participantes. Tais expressões de DO direto provocam um efeito anestésico ou mecanismo de reação por parte dos participantes. Em particular, os participantes de todos os GF mencionaram que desenvolveram “formas de se protegerem” (P2-LGBTQ+) contra este tipo de DO explícito ou “habituararam-se” (P3-Roma) a estas expressões: “preto, vai para a tua terra” (P4-Afro), “os ciganos merecem morrer” (P2-Roma) e “suas putas!” (P4-LGBTQ+).

Por outro lado, com base nas experiências vivenciadas pelos participantes, comparativamente ao DO explícito, o DO encoberto pode ser considerado mais danoso para a autoestima e dignidade deste grupo. Isto acontece principalmente devido aos valores subjacentes geralmente associados a mensagens encobertas baseadas na superioridade branca e na discriminação estrutural (ou institucional). Ademais, o DO encoberto faz com que os grupos-alvo se sintam impotentes, uma vez que essas mensagens são frequentemente disfarçadas ou transmitidas sob a forma de elogios que são prejudiciais aos alvos. Como defendido pelos participantes, quando o DO é apresentado como elogio (intencional ou disfarçado), o alvo desse elogio sente-se desarmado, ficando sem argumentos para se defender. Embora a noção de que o DO encoberto será mais prejudicial nem sempre seja óbvia, os dados revelam que os participantes destacam esta diferença de intensidade. Por exemplo, P3-Roma afirma:

para mim, é tão nocivo o discurso de ódio explícito como o velado. Às vezes, o discurso velado é pior porque nós não estamos à espera e nem sabemos o porquê de não termos sido aceites, mas é só porque existem pessoas preconceituosas atrás.

Os dados demonstram ainda que os elogios se expressam sob, pelo menos, duas formas interligadas: (a) a interseção entre nacionalidade e “raça” e (b) a normatividade.

A primeira diz respeito à relação entre a identidade nacional portuguesa e a branquitude (Fanon, 1952/2017), definida “como um sistema de dominação baseado na perpetuação da subordinação dos povos identificados e definidos como não brancos pelos povos definidos historicamente como brancos/ocidentais ou dentro da europeidade” (Maeso, 2021, p. 29). Esta forma contemporânea de racismo associada à nacionalidade foi também documentada por Kilomba (2018). Como refere Kilomba, estas novas formas de racismo raramente inferem “inferioridade racial” como no passado; elas remetem antes para “diferença cultural” ou de “religiões” e da sua incompatibilidade com a cultura nacional.

Neste sentido, P4-Afro relatou que frequentemente lhe dizem que ele “parece muito português”, o que ele acredita ser uma tentativa de “branqueamento” da sua identidade, uma vez que cresceu numa família branca privilegiada após ter sido adotado. Ele citou o livro *Um Preto Muito Português* como uma possível explicação de como os outros o veem. P4-Afro exemplifica o tipo de enunciados que reforçam esta interpretação: “‘Para mim, tu és como se fosses branco’, isto é igualmente ofensivo, é uma categorização de que os pretos são inferiores. É como quando dizem: ‘Esqueci-me que és preto’ ou ‘para um preto, és muito inteligente’”.

A segunda forma sob a qual o DO se manifesta tem a ver com as experiências partilhadas pelos participantes LGBTQ+. Vários participantes partilharam como o DO encoberto disfarçado de elogio se pode revelar um padrão que se insere na categoria de normatividade, entendida aqui como um sistema de hierarquia que determina práticas e crenças sobre o que é “aceitável” e “normal” num determinado contexto social (Toomey et al., 2012). Com frequência, os participantes afirmaram que estes supostos elogios implicam conformidade com essas normas sociais, o que significa que, desde que uma pessoa LGBTQ+ se comporte como uma pessoa heteronormativa (Toomey et al., 2012), pode ser aceite, como ilustram as seguintes citações: “o que mais elogiam é o que mais se assemelha à norma em detrimento de algo que não seja considerado a norma. Esses elogios envolvem discursos como: ‘és trans, mas nem sequer reparei, assim está tudo bem’” (P1-LGBTQ+). Além disso, P4-LGBTQ+ afirma: “‘é tão bonito o vosso amor’ [referindo-se à manifestação de afeto de um casal transexual], ‘até pensei que ela não fosse trans’ [uma pessoa transexual], por exemplo, são coisas que as pessoas podem pensar como uma forma de elogio, mas não é”.

O participante P5-LGBTQ+ (um homem que se identifica como pansexual) questiona este fenómeno, pois muitas vezes implica superioridade masculina e discriminação contra as mulheres, e não apenas contra uma pessoa LGBTQ+. Um dos comentários frequentes que recebem é: “és gay, mas pelo menos és homem como deve ser” ou “és gay, mas não és efeminado”. Dentro do grupo LGBTQ+, os elogios manifestam-se também em termos de atração ou atratividade sexual, como explica P3-LGBTQ+: “também tive outra situação em que uma pessoa [um homem heterossexual] me desejou os parabéns [pelos meus anos] e disse: ‘como é que eu te saltava para cima se gostasses de homens!’”.

Ainda considerando a normatividade como um veículo de DO disfarçado de elogio, este resultado foi também evidente no GF das pessoas de etnia cigana ou Roma. Entre outros, os estereótipos antigiganos giram em torno da criminalidade, da preguiça e do apoio, sob a forma de benefícios sociais imerecidos, por parte do Estado (Sam Nariman et al.,

2020). Além disso, os ciganos são considerados inferiores tanto em calor humano/afeição como em competência (Grigoryev et al., 2019), perigosos e desprezíveis (Hadarics & Kende, 2019), ou mesmo vagabundos, sujos e imorais (Liégeois, 2007). Sempre que as pessoas de etnia Roma não se enquadram nesses estereótipos negativos, quase uma espécie de normal social no contexto português, ou expectativas acima mencionadas, são geralmente vistas como exceções.

Neste sentido, P1-Roma já experienciou este tipo de estigmatização no trabalho: “trabalhei noutra local em que ouvia, quando souberam que eu sou cigana, comentários como: ‘não pareces nada cigana, és tão calminha’, ‘nós não te vemos como cigana’”.

Adicionalmente, P2-Roma questiona a ideia de estar associado a um comportamento esperado: “o que me deixa mesmo irritado é quando dizem: ‘tu não pareces cigano’, mas porquê é que eu não pareço cigano? Eu continuo a ser cigano independentemente do que estiver a fazer”. P2-Roma também relatou que, ao responder educadamente a um comentário aleatório nas redes sociais, normalmente recebe de volta o seguinte comentário: “‘parabéns, se responde assim, você é diferente’, isso é uma forma de racismo”, o que nos levou a conceitualizar os elogios como uma forma eficaz de difundir o DO encoberto ou indireto.

Para concluir esta subsecção dos resultados, apresentamos a Tabela 1, que apresenta alguns exemplos ilustrativos fornecidos pelos participantes afrodescendentes, Roma e LGBTQ+, respetivamente, codificados como elogios.

	AFRO	ROMA	LGBTQ+
	“Tu para mim é como se fosses branco.” (P4-Afro)	“Tu não pareces cigano.” (P2-Roma)	“É tão bonito o vosso amor, até pensei que ela não fosse trans.” (P4-LGBTQ+)
	“Tu para preto és muito inteligente.” (P4-Afro)	“Não pareces nada cigana, és tão calminha”, “nós não te vemos como cigana.” (P1-Roma)	“És trans, mas nem sequer reparei, assim está tudo bem.” (P3-LGBTQ+)
	“‘Tu és muito mais inteligente do que eu alguma vez imaginei’ [risos], é um pouco, ‘Eu nunca imaginei que tu chegasses onde tu estás’, é um bocado isso.” (P5-Afro)	“Parabéns, se responde assim [educadamente], você é diferente.” (P1-Roma)	“‘Desperdício!’. É do género: ‘que desperdício estares com uma rapariga quando podias estar comigo!’” (P5-LGBTQ+)
Citações codificadas como “elogios”	“A pessoa quando começa a elogiar eu já estou, de certa forma, bloqueado. Ou o elogio é realmente muito bom — o que nunca acontece, porque vai sempre tocar no cabelo ou nesses elogios do: ‘apesar de seres preto e tal...’” (P7-Afro)		“Como é que eu te saltava para cima se gostasses de homens!” (P3-LGBTQ+)
	“Tu até falas muito bem para preto”, ou “tu até te portas bem.” (P2-Afro)		
	“Ah, não achava que fosses tão inteligente, que lesses tanto ou que fosses tão culto.” (P6-Afro)		

Tabela 1. Citações codificadas como “elogios” extraídas do NVivo — Afro, Roma e LGBTQ+

7.2. SEGUNDO TEMA: DISCURSO DE ÓDIO DISFARÇADO DE HUMOR

Um tema prevalente em todos os grupos focais é que a mensagem do DO encoberto assume frequentemente a forma de humor. Todos os grupos concordaram que este tipo de DO disfarçado assume este formato, muitas vezes no local de trabalho, nos meios de comunicação e redes sociais, contribuindo para perpetuar a discriminação contra os grupos-alvo e a normalização do DO na sociedade. Além disso, o humor é uma forma eficaz de invalidar as identidades sociais dos grupos marginalizados. Os participantes do nosso estudo associaram todas estas camadas ao DO, defendendo formas de desenvolver estratégias de contranarrativa baseadas precisamente no humor, porque reconhecem o poder deste dispositivo retórico.

Por exemplo, P1-Roma conceptualizou o humor como um disfarce para o DO capaz de reforçar estereótipos negativos acerca da população cigana, frequentemente difundidos pelos meios de comunicação social tradicionais. O participante exemplificou como os ciganos são geralmente retratados em programas de televisão portugueses:

uma personagem vestida de preto, com barba, mal-arrumado, com chapéu e com um sotaque forçado. [Os meios de comunicação social] fazem questão de forçar os estereótipos e prolongar aquilo que é negativo à pessoa cigana. O humor é o discurso de ódio de forma camuflada.

Adicionalmente, P3-Roma sublinha a normatividade associada ao conceito de “humor negro”:

sou ativista, voluntária e faço parte de outros grupos. Num desses grupos de que faço parte aconteceu estar com um rapaz que gosta de humor negro e que enquanto eu estive presente, ele conteve-se, mas a partir do momento em que eu saí, ele disse as piadas todas que quis. O humor negro está tão naturalizado que quem ouviu a piada, mais tarde, fez questão de me contar. A piada era: “sabes porque é que as crianças com síndrome de Down não podem casar com ciganos? Porque têm todos um problema”, ou seja, a síndrome de Down é um problema e associam os ciganos a um problema também.

Por outro lado, P4-LGBTQ+ realça a importância de capacitar os grupos-alvo do DO, fornecendo-lhes mecanismos para contrariar o discurso encoberto transmitido sob a forma de humor, tendo em consideração o quão prejudicial isso pode ser para as vítimas, especialmente num contexto familiar, onde as pessoas podem ter relações próximas, porém frágeis. Para destacar a relevância de se pensar na elaboração de contranarrativas eficazes, P4-LGBTQ+ questiona: “como é que desmantelamos uma piada deste tipo [vinda de um familiar]?”. Em relação a essa dificuldade de abordar o DO em contextos familiares, P2-LGBTQ+ acrescenta:

na noite de Natal ou nos velórios de família onde toda a gente contava anedotas sobre negros, sobre ciganos, eu passava-me. Eu lembro-me sempre de uma cena do *Filadélfia* [referência ao filme norte-americano, de 1993, realizado por Jonathan Demme], quando ele [referência ao protagonista do filme,

um homem gay] está na sauna com os colegas e alguém diz: “sabes como é que um gay simula um orgasmo? Deita um iogurte morno para cima das costas”, e ele cala-se. Eu acho que nós não nos devemos calar, independentemente de sermos LGBT ou não. É como as palavras, eu posso apropriar-me de uma expressão ou palavra negativa e torná-la num empoderamento, por exemplo, o [a palavra] *queer* era um termo negativo e na brincadeira eu posso dizer “paneleiro”, mas isso sou eu.

Os participantes do grupo de afrodescendentes também mencionaram vários exemplos de como o humor conota o DO mesmo nos principais meios de comunicação social, depreciando políticos ou jornalistas negros em Portugal ou no local de trabalho. Para P4-Afro, vários programas de entretenimento televisivo em Portugal podem funcionar como uma forma de disseminar o DO encoberto, como ilustra a citação abaixo:

é quase indissociável [os limites entre humor e o DO]. Quer de opinião, quer de humor, quer de ironia, vai se misturar com tudo. O Rui Sinel de Cordes², por exemplo, que é um humorista português — vamos chamar-lhe humorista —, tudo o que ele faz é um discurso de ódio em forma de piada, contra mulheres, contra homossexuais, contra negros, sobre a capa de: “posso dizê-lo, eu estou em cima de um palco. Isto não são as minhas opiniões, são piadas”, mas sim, aquilo são as opiniões deles, basta ver quem é o público do Rui Sinel de Cordes. Iniciou toda uma escola de humor em que quase todos os humoristas, são homens, brancos, héteros, *cis*, e repetem aquelas piadas *ad nauseam*. Não é uma piada, aquilo é repetição de um preconceito, só que dito em *punch line*. Isto eu posso dizê-lo em televisão ou, se calhar, até o próprio Fernando Rocha, fazendo de uma forma menos inteligente com personagens estereotipadas com personagens como o Tibúrcio e a Matumbina. É muito difícil destrinçar onde começa o discurso de ódio dos seus vários tipos, aquele “Preto, vai para a tua terra”, ou se o *Levanta-te e Ri*, todas as noites na SIC — do Tibúrcio e da Matumbina [ver Figura 4].

² Rui Sinel de Cordes é um ator português, comediante e apresentador de televisão, um pioneiro do “humor negro” em Portugal, e autor de vários programas de televisão e espetáculos de *stand-up* de sucesso.



Fernando Rocha Portugal a Rir 4 - Gago Vendedor

Figura 4. As personagens Tibúrcio e Matumbina mencionadas pelo P4-Afro aparecem no centro desta imagem

Fonte. Retirado de *Fernando Rocha Portugal a Rir 4 - Gago Vendedor*, por sancastroo07 [@sancastroo07], 2011, YouTube.
(<https://www.youtube.com/watch?v=IQBs7ITL18U&t=25s>)

Em consonância com P4-Afro, P2-Roma também referiu que o trabalho de Rui Sinel de Cordes mencionado acima é problemático porque “querem rebelar-se, dizer que estão contra o sistema e a opressão, mas depois o que fazem é destilar ódio e não piadas”. Como P1-Roma também apontou, os principais meios de comunicação social exploram estereótipos negativos, disfarçando-os de humor, para perpetuar sentimentos negativos e hostilidade em relação aos grupos-alvo.

Lembrei-me que utilizam muito o termo “lelos” [termo pejorativo usado para pessoas de etnia cigana] para se referirem a ciganos. No programa *Malucos do Riso*, já utilizaram esse termo. Sempre que há uma cena em que querem representar uma pessoa cigana, associam a uma personagem vestida de preto, com barba, mal-arrumada, com chapéu e com um sotaque forçado.

Embora os participantes reconheçam que o humor pode ser utilizado como um disfarce para a disseminação do DO e um reforço dos estereótipos negativos, também destacam a importância de identificar a identidade social e o ponto de vista do emissor da mensagem. Por exemplo, P6-Afro, residente em Inglaterra, menciona como os humoristas racializados contam as suas próprias histórias, o que lhe proporciona efetiva diversão, em contraste com humoristas que simplesmente repetem e reforçam estereótipos.

Outra coisa que eu acho muito interessante, é que o humor cá [Inglaterra] é muito diferente do humor que ainda se vive em Portugal. Cá os humoristas só fazem humor das próprias experiências, por exemplo, se é uma pessoa negra, só faz humor sobre aquilo que passou, não vai buscar a história de outra pessoa e fazer humor disso. O criar humor é contar a própria história, mas de uma forma engraçada, porque é tão ridículo que mete piada. Acho que é esse tipo de humor que eu gosto de ver.

Por outro lado, alguns dos participantes consideram que o humor produzido por pessoas marginalizadas pode ter um efeito negativo nas suas próprias comunidades, pois pode reforçar estereótipos. P4-LGBTQ+ explica:

eu acho que em momento nenhum devem ser utilizadas [as piadas]. A Hannah Gadsby [comediante de nacionalidade australiana, que aborda tópicos como o sexismo, homofobia, e temas relacionados com a sua sexualidade] diz num dos seus espetáculos, que existem tantas formas de fazer humor que não se deve fazer humor porque somos isto ou aquilo. Há outros humoristas que dizem, muitas vezes: “eu sou gordo e sempre fiz piadas de gordo, até perceber que isso era uma falta de amor próprio muito grande”, ou “eu sou lésbica e sempre fiz piadas sobre o facto de ser lésbica, depois comecei a perceber que não o deveria fazer porque isto era legitimar que fizessem piadas comigo”.

De forma geral, os resultados evidenciam que o humor é visto pelos participantes como um modo prejudicial de perpetuar e normalizar estereótipos negativos associados a identidades sociais, tais como os ciganos, tornando difícil para estes indivíduos encontrar emprego, alojamento e navegar na sua vida quotidiana. Além disso, os participantes destacam a importância de distinguir entre os conceitos de “ferir” e “prejudicar”. P3-Roma, pertencente a um dos grupos étnicos mais estigmatizados em Portugal, partilha esta perspetiva e enfatiza a gravidade do impacto do humor na vida das pessoas pertencentes a grupos marginalizados em contraposição a grupos dominantes.

Para além do senso comum — porque cada um tem o seu e é muito perigoso quando deixamos na mão de cada um —, o que difere o politicamente correto do discurso de ódio, é quando as pessoas utilizam estas piadas para justificar os seus atos de discriminação e racismo contra ciganos. Atrás de uma piada que é inofensiva e não faz mal, está a intenção. Em Portugal há muita gente que faz piada dos alentejanos, dizem que são preguiçosos e não gostam de trabalhar, mas é uma piada que não prejudica a vida real de um alentejano. Se um alentejano procurar trabalho, ninguém lhe vai dizer: “não te dou trabalho porque os alentejanos são preguiçosos”.

Em conclusão, todos os grupos envolvidos no estudo concordam que não há uma distinção clara entre o humor e a disseminação de DO. Como afirmou P1-Roma: “o que é proferido no discurso de ódio é o mesmo que é proferido no humor, uma vez que se encontram preconceitos, estereótipos, racismo e insultos; apenas o sarcasmo e a ironia são adicionados”.

8. DISCUSSÃO E CONCLUSÃO

As mensagens indiretas ou implícitas de ódio são frequentemente mencionadas na literatura como “camufladas”, “subtis”, “soft”, “encobertas” ou “disfarçadas” de DO (Baider & Constantinou, 2020; Bhat & Klein, 2020; Jha & Mamidi 2017; Rieger et al.,

2021; van Dijk 1992, 1993; Wodak & Reisigl, 2015). Todas estas definições metafóricas implicam a ideia de algo oculto que precisa de ser desvendado para ser plenamente compreendido e, assim, adequadamente abordado ao nível da sua deteção e implementação de políticas públicas. Neste sentido, o presente estudo procurou trazer à luz alguns aspetos que podem ajudar a compreender o DO encoberto a partir da perspetiva dos principais grupos-alvo em Portugal.

A Figura 5 apresenta uma representação visual que resume os principais resultados desta investigação, na qual as perceções dos grupos-alvo do DO reforçam a ideia de que os estereótipos podem ser antecedentes ou servir de base para o preconceito (Brigham, 1971). A imagem também destaca a convicção de que a deteção do DO depende fortemente do reconhecimento de estereótipos sociais (Davani et al., 2021). Embora algumas definições do DO incluam explicitamente a noção de estereótipos negativos (European Commission against Racism and Intolerance, 2020), argumentamos que os estereótipos positivos também podem alimentar o DO encoberto. De facto, as conclusões deste estudo indicam que os elogios não se baseiam apenas em estereótipos negativos (e.g., os ciganos são problemáticos: “não pareces nada cigana, és tão calma”, “não te vemos como um cigano”), mas também positivos (e.g., no que se refere à atração física: “desperdício!”, no sentido de ‘que desperdício estar com uma rapariga quando se pode estar comigo’”).

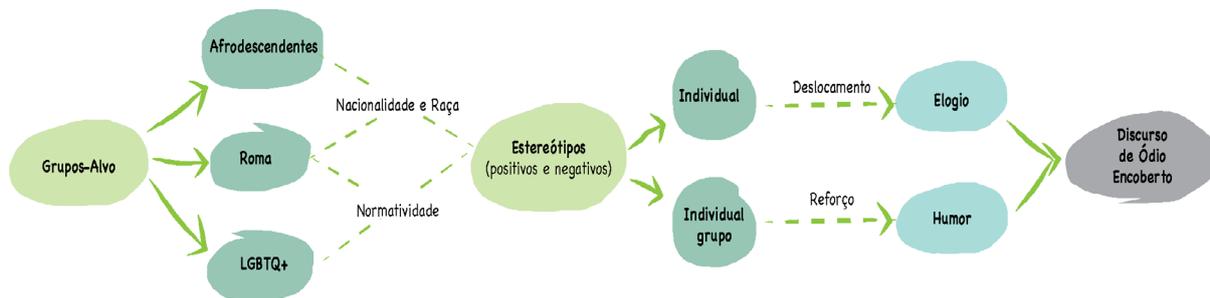


Figura 5. Representação visual dos processos subjacentes à materialização do discurso de ódio encoberto com base em elogios e humor

No que diz respeito aos elogios, o alvo (geralmente um membro específico de uma comunidade) é deslocado do grupo ao qual pertence (i.e., o “exogrupo”) para o grupo dominante (i.e., o “endogrupo”). Por outro lado, no humor, há um reforço dos estereótipos (geralmente negativos) associados ao grupo-alvo, o que reforça a sua posição no “exogrupo”. As operações de deslocamento e reforço subjacentes, respetivamente, aos elogios estereotipados e ao humor, estão também representadas na Figura 5. Mais detalhadamente, analisando a figura da esquerda para a direita, as comunidades afrodescendente e cigana estão ligadas por linhas pontilhadas porque ambas partilham a experiência de que o elogio se baseia em estereótipos positivos que estão enraizados em noções de nacionalidade e “raça”. Neste contexto, e recorrendo ao trabalho de Kilomba (2018), estes resultados dialogam com a questão provocadora levantada pela autora: é possível uma pessoa não-branca ser portuguesa? Ou a pessoa negra é sempre a estranha, a estrangeira que é posta fora da nação? Para a autora, no escopo da fantasia colonial, a

etnia, cor da pele ou “raça” é construída dentro de limites nacionais específicos, e a nacionalidade nos termos da “raça”. Para os participantes deste estudo, formas encobertas de DO como elogio incorporam essa noção trazida por Kilomba, como no exemplo dado por P4-Afro: “tu pareces muito português” combinado com “tu para mim é como se fosses branco”.

Na Figura 5, em contrapartida, a comunidade LGBTQ+ está ligada à comunidade cigana porque, para elas, o elogio deriva do cumprimento de uma certa norma social (e.g., um homossexual que parece heterossexual, ou uma pessoa de etnia cigana que parece ser calma). O elemento central da Figura 5 refere-se aos estereótipos, um mecanismo basilar na construção do DO encoberto, que atua a nível individual no que concerne aos elogios e a nível individual ou de grupo no que diz respeito ao humor.

É importante contrastar esta articulação dos estereótipos como estruturas basilares do DO com a literatura existente, uma vez que estudos anteriores, na área de psicologia social, demonstraram que os “elogios estereotípicos” se baseiam em generalizações de estereótipos positivos construídos unicamente sobre a filiação em grupo, sem uma formação individual adequada (Czopp, 2008; Czopp et al., 2015). Contudo, o nosso estudo sugere algo diferente. Do ponto de vista do emissor da mensagem, o alvo merece ser elogiado porque não se enquadra nos estereótipos negativos associados ao grupo social a que pertence, o que motiva a sua separação (ou deslocamento) da posição grupal (e.g., “sendo negro, és muito inteligente”; “és trans, mas eu nem reparei, por isso está tudo bem” ou “nem pareces cigana”). Do ponto de vista do alvo, este tipo de elogio é prejudicial porque revela opressões estruturais profundas, como, por exemplo, supremacia branca ou branquitude (Fanon, 1952/2017; Maeso, 2021), racismo contemporâneo (Kilomba, 2018) ou heteronormatividade (Toomey et al., 2012), que o torna impotente e sem saber como reagir. Isto é evidenciado por diversos estudos sobre como os alvos reagem quando recebem “estereótipos elogiosos” e sobre a dificuldade de enfrentar ou responder a estereótipos positivos que pretendem ser elogios (Alt et al., 2019; Czopp & Monteith, 2006; Siy & Cheryan, 2013).

Assim como nos elogios, o humor que visa grupos vulneráveis ou marginalizados muitas vezes é construído com base em estereótipos, principalmente negativos, com o objetivo de ridicularizar o grupo ou o indivíduo como membro desse grupo. No entanto, ao contrário dos elogios, que buscam particularizar e deslocar o indivíduo das características do grupo, o humor, mesmo quando direcionado a um indivíduo específico, procura generalizar o estereótipo e o preconceito associado à sua comunidade. Nesse sentido, a função do humor é reforçar o estereótipo, contribuindo para sua normalização.

Sustentamos que, neste caso, o que está em causa não é o humor afiliativo (Martin et al., 2003), mas o humor agressivo, o qual se apoia geralmente no uso de sarcasmo, ironia e sátira, com o objetivo de normalizar o estereótipo negativo e reforçar a barreira entre aqueles que se encontram numa posição de poder, e geralmente apreciam este tipo de humor, e as comunidades vulneráveis, que são simultaneamente objetos e receptores (indiretos) de ridicularização. É importante destacar que o humor agressivo pode ter graves implicações para as comunidades visadas, perpetuando a discriminação e reforçando estereótipos prejudiciais, e como tal deve ser combatido, por exemplo, através de contranarrativas que tirem igualmente partido deste poderoso mecanismo retórico (Baider, 2023).

Independentemente das intenções dos que produzem ou reproduzem elogios e humor baseados em estereótipos positivos ou negativos, o indivíduo ou grupo-alvo é fortemente afetado pela força retórica destas estratégias, assumindo que podem ser mais prejudiciais do que o DO direto, e inibir ou anular a sua capacidade de resposta ou reação. Embora o humor afiliativo possa ter um impacto positivo nos indivíduos, reforçando potenciais laços sociais e culturais, tal observa-se principalmente nos casos em que os indivíduos pertencem ao mesmo grupo étnico (Gogová, 2016) e partilham a mesma identidade social. De facto, como afirma Holmes (2000), nas interações em que o poder é particularmente desequilibrado, as funções de humor na construção de relações são frequentemente mais complexas.

Assim, este estudo destaca a importância de examinar a diversidade de estratégias retóricas subjacentes ao DO encoberto e revela a fragilidade da maioria das atuais abordagens de deteção do DO, que frequentemente negligenciam o DO encoberto em geral e o humor em particular, e não têm em consideração a identidade social dos grupos-alvo (Baider & Constantinou, 2022).

Compreender o DO encoberto na forma de elogios e humor, seja online ou offline, tem implicações na análise e deteção do DO, frequente em conteúdos gerados pelos utilizadores e publicados nas redes sociais e outros meios digitais, incluindo os meios de comunicação social tradicionais, no caso específico do humor, como mencionado pelos participantes deste estudo.

Concluimos este artigo com a expectativa de que esta investigação contribua para o desenvolvimento informado de ferramentas de deteção automática do DO, cujo desempenho depende fortemente da capacidade de reconhecer estratégias retóricas e discursivas complexas, tais como a ironia e o humor. Além disso, esta investigação contribui para o desenvolvimento de estudos futuros, no sentido em que aponta para novas linhas de investigação, pois ao recuperarmos o conceito de estereótipos positivos, clássico no campo da psicologia social, mas pouco explorado na comunicação, contribuimos para conceptualizar formas encobertas de DO como elogios. Esta articulação conceptual representa uma contribuição significativa para a compreensão do DO encoberto, podendo impactar tanto a investigação quanto o desenvolvimento de sistemas de deteção automática de DO online. Por fim, os resultados deste estudo podem ter implicações que transcendem o âmbito académico, informando a criação de políticas de conduta acerca do DO adotadas pelas redes sociais.

Tradução: Cláudia Silva e Paula Carvalho

AGRADECIMENTOS

As autoras agradecem a todos os participantes que tornaram este estudo possível e também aos revisores anónimos, cujas contribuições foram valiosas para a melhoria do manuscrito original. Salientamos que esta investigação foi apoiada pelo ITI-LARSyS (ITI-LX UIDP/50009/2020 - IST-ID) e INESC-ID (UIDB/50021/2020) e integra o projeto *HATE COVID-19.PT* (projeto 759274510), financiado pela Fundação para a Ciência e a Tecnologia.

REFERÊNCIAS

- Albelda Marco, M. (2022). Rhetorical questions as reproaching devices. *Journal of Language Aggression and Conflict*. <https://doi.org/10.1075/jlac.00077.alb>
- Alt, N. P., Chaney, K. E., & Shih, M. J. (2019). “But that was meant to be a compliment!”: Evaluative costs of confronting positive racial stereotypes. *Group Processes & Intergroup Relations*, 22(5), 655–672. <https://doi.org/10.1177/1368430218756493>
- Assimakopoulos, S., Baider, F. H., & Millar, S. (2017). *Online hate speech in the European Union*. Springer. https://doi.org/10.1007/978-3-319-72604-5_1
- Baider, F. (2022). Covert hate speech, conspiracy theory and anti-semitism: Linguistic analysis versus legal judgement. *International Journal for the Semiotics of Law*, 35, 2347–2371. <https://doi.org/10.1007/s11196-022-09882-w>
- Baider, F. (2023). Accountability issues, online covert hate speech, and the efficacy of counter-speech. *Politics and Governance*, 11(2). <https://doi.org/10.17645/pag.v11i2.6465>
- Baider, F., & Constantinou, M. (2020). Covert hate speech: A contrastive study of Greek and Greek Cypriot online discussions with an emphasis on irony. *Journal of Language Aggression and Conflict*, 8(2), 262–287. <https://doi.org/10.1075/jlac.00040.bai>
- Bakowski, P. (2022). *Combating hate speech and hate crime in the EU*. [https://www.europarl.europa.eu/RegData/etudes/ATAG/2022/733520/EPRS_ATA\(2022\)733520_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/ATAG/2022/733520/EPRS_ATA(2022)733520_EN.pdf)
- Bardzell, S., & Bardzell, J. (2011). Towards a feminist HCI methodology: Social science, feminism, and HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 675–684). Association for Computing Machinery.
- Ben-David, A., & Fernández, A. M. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, 10, 1167–1193. <https://ijoc.org/index.php/ijoc/article/view/3697>
- Bhat P., & Klein, O. (2020) Covert hate speech: White nationalists and dog whistle communication on Twitter. In G. Bouvier & J. Rosenbaum (Eds.), *Twitter, the public sphere, and the chaos of online deliberation* (pp. 151–172). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-41421-4_7
- Billig, M. (2001). Humour and hatred: The racist jokes of the Ku Klux Klan. *Discourse & Society*, 12(3), 267–289. <https://doi.org/10.1177/0957926501012003001>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Breazu, P., & Machin, D. (2022a). Using humor to disguise racism in television news: The case of the Roma. *Humor*, 35(1), 73–92. <https://doi.org/10.1515/humor-2021-0104>
- Breazu, P., & Machin, D. (2022b). Racism is not just hate speech: Ethnonationalist victimhood in YouTube comments about the Roma during COVID-19. *Language in Society*, 1–21. <https://doi.org/10.1017/S0047404522000070>
- Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin*, 76, 15–38. <https://doi.org/10.1037/h0031446>
- Brown, A. (2017). What is hate speech? Part 1: The myth of hate. *Law and Philosophy*, 36, 419–468. <https://doi.org/10.1007/s10982-017-9297-1>
- Buturoiu, D. R., & Corbu, N. (2020). Exposure to hate speech in the digital age. Effects on stereotypes about Roma people. *Journal of Media Research*, 13(2), 5–26. <https://doi.org/10.24193/jmr.37.1>

- Cádima, F. R., Baptista, C., Martins, L. O., Silva, M. T., & Lourenço, R. (2020). *Monitoring media pluralism in the digital era: Application of the media pluralism monitor in the European Union, Albania & Turkey in the years 2018-2019. Country report: Portugal*. European University Institute.
- Chovanec, J. (2021). 'Re-educating the Roma? You must be joking...': Racism and prejudice in online discussion forums. *Discourse & Society*, 32(2), 156–174. <https://doi.org/10.1177/0957926520970384>
- Comissão para a Igualdade e Contra a Discriminação Racial. (2020). *Relatório anual 2019: Igualdade e não discriminação em razão da origem racial e étnica, cor, nacionalidade, ascendência e território de origem*. ACM; CIDR; República Portuguesa.
- Condon, L., Bedford, H., Ireland, L., Kerr, S., Mytton, J., Richardson, Z., & Jackson, C. (2019). Engaging gypsy, Roma, and traveller communities in research: Maximizing opportunities and overcoming challenges. *Qualitative Health Research*, 29(9), 1324–1333. <https://doi.org/10.1177/1049732318813558>
- Council of Europe. (2022). *Recommendation CM/Rec (2022)16 of the Committee of Ministers to member States on combating hate speech*. https://search.coe.int/cm/Pages/result_details.aspx?ObjectId=0900001680a67955
- Czopp, A. M. (2008). When is a compliment not a compliment? Evaluating expressions of positive stereotypes. *Journal of Experimental Social Psychology*, 44(2), 413–420. <https://doi.org/10.1016/j.jesp.2006.12.007>
- Czopp, A. M., Kay, A. C., & Cheryan, S. (2015). Positive stereotypes are pervasive and powerful. *Perspectives on Psychological Science*, 10, 451–463. <https://doi.org/10.1177/1745691615588091>
- Czopp, A. M., & Monteith, M. J. (2006). Thinking well of African Americans: Measuring complimentary stereotypes & negative prejudice. *Basic and Applied Social Psychology*, 28, 233–250. https://doi.org/10.1207/s15324834basp2803_3
- Daniels, J. (2008). Race, civil rights, and hate speech in the digital era. In A. Everett (Ed.), *Learning race and ethnicity: Youth and digital media* (pp. 129–154). The MIT Press. <https://doi.org/10.1162/dmal.9780262550673.129>
- Davani, A. M., Atari, M., Kennedy, B., & Dehghani, M. (2021). *Hate speech classifiers learn human-like social stereotypes*. <https://doi.org/10.48550/arXiv.2110.14839>
- De Cillia, R., Reisigl, M., & Wodak, R. (1999). The discursive construction of national identities. *Discourse & Society*, 10(2), 149–173. <https://doi.org/10.1177/0957926599010002002>
- Delcour, C., & Hustinx, L. (2017). The Roma as ultimate European minority and ultimate outsider? In J. Chovanec & K. Molek-Kozakowska (Eds.), *Representing the other in European media discourses* (pp. 259–280). John Benjamins. <https://doi.org/10.1075/dapsac.74>
- Dynel, M. (2018). *Irony, deception and humour. Seeking the truth about overt and covert untruthfulness*. De Gruyter Mouton. <https://doi.org/10.1515/9781501507922>
- ElSherief, M., Ziems, C., Muchlinski, D., Anupindi, V., Seybolt, J., De Choudhury, M., & Yang, D. (2021, 7–11 de novembro). *Latent hatred: A benchmark for understanding implicit hate speech* [Apresentação de conferência]. The 2021 Conference on Empirical Methods in Natural Language Processing, Punta Cana, República Dominicana. <https://doi.org/10.48550/arXiv.2109.05322>
- European Commission. (s.d.). *The EU Code of conduct on countering illegal hate speech online: The robust response provided by the European Union*. https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en
- European Commission against Racism and Intolerance. (2016). *ECRI general policy recommendation n°15 on combating hate speech*. <https://rm.coe.int/ecri-general-policy-recommendation-no-15-on-combating-hate-speech/16808b5b01>

- European Commission against Racism and Intolerance. (2020). *Annual report on ECRI's activities covering the period from 1 January to 31 December 2019*. Council of Europe. <https://rm.coe.int/ecri-annual-report-2019/16809ca3e1>
- Fanon, F. (2017). *Pele negra, máscaras brancas* (A. Pomar, Trad.). Letra Livre. (Trabalho original publicado em 1952)
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902. <https://doi.org/10.1037/0022-3514.82.6.878>
- Fortuna, P., & Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)*, 51(4), 1–30. <https://doi.org/10.1145/3232676>
- Given, L. M. (Ed.). (2008). *The SAGE encyclopedia of qualitative research methods*. SAGE publications.
- Gogová, L. (2016). Ethnic humour in a multicultural society. *Ars Aeterna*, 8(2), 12–24. <https://doi.org/10.1515/aa-2016-0006>
- Grigoryev, D., Fiske, S. T., & Batkhina, A. (2019). Mapping ethnic stereotypes and their antecedents in Russia: The stereotype content model. *Frontiers in Psychology*, 10, Artigo 1643. <https://doi.org/10.3389/fpsyg.2019.01643>
- Hadarics, M., & Kende, A. (2019). Negative stereotypes as motivated justifications for moral exclusion. *The Journal of Social Psychology*, 159(3), 257–269. <https://doi.org/10.1080/00224545.2018.1456396>
- Haladzhun, Z., Harmatiy, O., Bidzilya, Y., Kunanets, N., & Shunevych, K. (2021). Hate speech in media towards the representatives of Roma ethnic community. In N. Sharonova, V. Lytvyn, O. Cherednichenko, Y. Kupriianov, O. Kanishcheva, T. Hamon, N. Grabar, V. Vysotska, A. Kowalska-Styczen, & I. Jonek-Kowalska (Eds.), *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems* (pp. 755–768). Colins AI.
- Holmes, J. (2000). Politeness, power and provocation: How humour functions in the workplace. *Discourse Studies*, 2(2), 159–185. <https://doi.org/10.1177/1461445600002002002>
- Jha, A., & Mamidi, R. (2017). When does a compliment become sexist? Analysis and classification of ambivalent sexism using twitter data. In D. Hovy, S. Volkova, D. Bamman, D. Jurgens, B. O'Connor, O. Tsur, and A. S. Doğruöz (Eds.), *Proceedings of the Second Workshop on NLP and Computational Social Science* (pp. 7–16). Association for Computational Linguistics.
- Jussim, L. J., & Rubinstein, R. (2012). *Stereotypes*. Oxford University Press.
- Kilomba, G. (2018). *Memórias da plantação: Episódios de racismo cotidiano*. Orfeu Negro.
- Krobová, T., & Zàpotocký, J. (2021). “I am not racist, but...”: Rhetorical fallacies in arguments about the refugee crisis on Czech Facebook. *Journal of Intercultural Communication*, 21(2), 58–69. <https://doi.org/10.36923/jicc.v21i2.14>
- Liégeois, J. P. (2007). Roma education and public policy: A European perspective. *European Education*, 39(1), 11–31. <https://doi.org/10.2753/EUE1056-4934390101>
- MacAvaney, S., Yao, H.-R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PLoS ONE*, 14(8), Article e0221152. <https://doi.org/10.1371/journal.pone.0221152>
- Maeso, S. R. (Ed.). (2021). *O estado do racismo em Portugal: Racismo antinegro e anticiganismo no direito e nas políticas públicas*. Tinta-da-China.

- Magu, R., & Luo, J. (2018). Determining code words in euphemistic hate speech using word embedding networks. In D. Fišer, R. Huang, V. Prabhakaran, R. Voigt, Z. Waseem, & J. Wernimont (Eds.), *Proceedings of the 2nd Workshop on Abusive Language Online* (pp. 93–100). Association for Computational Linguistics.
- Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the humor styles questionnaire. *Journal of Research in Personality*, 37(1), 48–75. [https://doi.org/10.1016/S0092-6566\(02\)00534-2](https://doi.org/10.1016/S0092-6566(02)00534-2)
- Matamoros-Fernández, A., & Farkas, J. (2021). Racism, hate speech, and social media: A systematic review and critique. *Television & New Media*, 22(2), 205–224. <https://doi.org/10.1177/1527476420982230>
- Matsuda, M. J. (1989). Public response to racist speech: Considering the victim's story. *Michigan Law Review*, 87(8), 2320–2381. <https://doi.org/10.2307/1289306>
- McInroy, L. B., & Craig, S. L. (2017). Perspectives of LGBTQ emerging adults on the depiction and impact of LGBTQ media representation. *Journal of Youth Studies*, 20(1), 32–46. <https://doi.org/10.1080/13676261.2016.1184243>
- Paz, M. A., Montero-Díaz, J., & Moreno-Delgado, A. (2020). Hate speech: A systematized review. *SAGE Open*, 10(4), 1–12. <https://doi.org/10.1177/2158244020973022>
- Poletto, F., Basile, V., Sanguinetti, M., Bosco, C., & Patti, V. (2021). Resources and benchmark corpora for hate speech detection: A systematic review. *Language Resources and Evaluation*, 55(2), 477–523. <https://doi.org/10.1007/s10579-020-09502-8>
- Reynders, D. (2022). *5th evaluation of the Code of Conduct* [Fact sheet]. European Commission. <https://commission.europa.eu/system/files/2022-12/Factsheet%20-%207th%20monitoring%20around%20of%20the%20Code%20of%20Conduct.pdf>
- Richardson-Self, L. (2018). Woman-hating: On misogyny, sexism, and hate speech. *Hypatia*, 33(2), 256–272. <https://doi.org/10.1111/hypa.12398>
- Rieger, D., Kümpel, A. S., Wich, M., Kiening, T., & Groh, G. (2021). Assessing the extent and types of hate speech in fringe communities: A case study of Alt-Right communities on 8chan, 4chan, and Reddit. *Social Media+ Society*, 7(4), 1–14. <https://doi.org/10.1177/20563051211052906>
- Sam Nariman, H., Hadarics, M., Kende, A., Lášticová, B., Poslon, X. D., Popper, M., Boza, M., Ernst-Vintila, A., Badea, C., Mahfud, Y., O'Connor, A., & Minescu, A. (2020). Anti-Roma bias (stereotypes, prejudice, behavioral tendencies): A network approach toward attitude strength. *Frontiers in Psychology*, 11, Article 2071. <https://doi.org/10.3389/fpsyg.2020.02071>
- Sancastroo07. (2011, 4 de setembro). *Fernado Rocha Portugal a Rir 4 - Gago Vendedor* [Vídeo]. YouTube. <https://www.youtube.com/watch?v=IQBs7ITL18U&t=25s>
- Sanguinetti, M., Poletto, F., Bosco, C., Patti, V., & Stranisci, M. (2018). An Italian Twitter corpus of hate speech against immigrants. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odiijk, S. Piperidis, & T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation* (pp. 1768–1775). European Language Resources Association.
- Schlesinger, A., Edwards, W. K., & Grinter, R. E. (2017). Intersectional HCI: Engaging identity through gender, race, and class. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 5412–5427). Association for Computing Machinery.
- Schmid, U. K., Kümpel, A. S., & Rieger, D. (2022). How social media users perceive different forms of online hate speech: A qualitative multi-method study. *New Media & Society*. Advance online publication. <https://doi.org/10.1177/14614448221091185>

- Schmidt, A., & Wiegand, M. (2019). A survey on hate speech detection using natural language processing. In L.-W. Ku & C.-T. Li (Eds.), *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media* (pp. 1–10). Association for Computational Linguistics.
- Seiter, E. (1986). Stereotypes and the media: A re-evaluation. *Journal of Communication*, 36(2), 14–26. <https://doi.org/10.1111/j.1460-2466.1986.tb01420.x>
- Silva, T. M. (2021). *Discurso de ódio, jornalismo e participação das audiências. Enquadramento, regulação e boas práticas*. Almedina; ERC.
- Siy, J. O., & Cheryan, S. (2013). When compliments fail to flatter: American individualism and responses to positive stereotypes. *Journal of Personality and Social Psychology*, 104(1), 87–102. <https://doi.org/10.1037/a0030183>
- Taecharungroj, V., & Nueangjamnong, P. (2015). Humour 2.0: Styles and types of humour and virality of memes on Facebook. *Journal of Creative Communications*, 10(3), 288–302. <https://doi.org/10.1177/0973258615614420>
- Toomey, R. B., McGuire, J. K., & Russell, S. T. (2012). Heteronormativity, school climates, and perceived safety for gender nonconforming peers. *Journal of Adolescence*, 35(1), 187–196. <https://doi.org/10.1016/j.adolescence.2011.03.001>
- Trebbe, J., Paasch-Colberg, S., Greyer, J., & Fehr, A. (2017). Media representation: Racial and ethnic stereotypes. In P. Rössler (Ed.), *The International encyclopedia of media effects* (pp. 1–9). <https://doi.org/10.1002/9781118783764.wbiemeo146>
- van Dijk, T. A. (1992). Discourse and the denial of racism. *Discourse & Society*, 3(1), 87–118. <https://doi.org/10.1177/0957926592003001005>
- van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, 4(2), 249–283. <https://doi.org/10.1177/0957926593004002006>
- Weaver, S. (2013). A rhetorical discourse analysis of online anti-Muslim and anti-Semitic jokes. *Ethnic and Racial Studies*, 36(3), 483–499. <https://doi.org/10.1080/01419870.2013.734386>
- Wodak, R., & Reisigl, M. (2015). Discourse and racism. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis* (pp. 576–596). John Wiley & Sons.
- Woodzicka, J. A., Mallett, R. K., Hendricks, S., & Pruitt, A. V. (2015). It's just a (sexist) joke: Comparing reactions to sexist versus racist communications. *Humor*, 28(2), 289–309. <https://doi.org/10.1515/humor-2015-0025>
- Zannettou, S., Caulfield, T., Blackburn, J., De Cristofaro, E., Sirivianos, M., Stringhini, G., & Suarez-Tangil, G. (2018). On the origins of memes by means of fringe web communities. In *Proceedings of the Internet Measurement Conference 2018* (pp. 188–202). Association for Computing Machinery.

NOTAS BIográficas

Cláudia Silva (doutorada em média digitais) é professora auxiliar convidada no Instituto Superior Técnico da Universidade de Lisboa, no Departamento de Engenharia Informática. É também investigadora integrada no Instituto de Tecnologias Interativas/Laboratório de Robótica e Sistemas de Engenharia. Anteriormente, lecionou como professora auxiliar convidada no Departamento de Comunicação da Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa. A sua pesquisa situa-se no uso das novas tecnologias de informação e comunicação para o empoderamento de comunidades

marginalizadas e desfavorecidas através de metodologias participativas e inclusivas.

ORCID: <https://orcid.org/0000-0002-5334-3424>

Email: claudiasilva01@tecnico.ulisboa.pt

Morada: Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

Paula Carvalho (doutorada em linguística) é investigadora no Instituto de Engenharia de Sistemas e Computadores — Investigação e Desenvolvimento, em Lisboa, e professora auxiliar convidada no Instituto Superior Técnico, Universidade de Lisboa. Nos últimos anos, tem integrado as equipas de vários projetos de investigação interdisciplinar, focados no desenvolvimento de recursos, métodos e aplicações para análise de diversos fenómenos linguísticos no âmbito das humanidades digitais e das ciências sociais. Entre outros tópicos, a sua investigação tem incidido na análise e deteção automática de sentimento, ironia e, mais recentemente, do discurso de ódio online e teorias da conspiração nos média sociais.

ORCID: <https://orcid.org/0000-0003-2884-1250>

Email: pcc@inesc-id.pt

Endereço: Rua Alves Redol, 9, 1000-029 Lisboa, Portugal

Submetido: 19/09/2022 | Aceite: 01/03/2023



Este trabalho encontra-se publicado com a Licença Internacional Creative Commons Atribuição 4.0.