



Argumentação em discursos de ódio no *Facebook*: uma categorização contributiva à Linguística Forense e à Linguística Computacional

Argumentation in hate speech on Facebook: a contributive categorization to Forensic Linguistics and Computational Linguistics

Welton Pereira e Silva

Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro,
Rio de Janeiro / Brasil

weltonp.silva@hotmail.com

<https://orcid.org/0000-0002-4693-3151>

Resumo: este artigo objetiva analisar estratégias argumentativas empregadas na veiculação de discursos de ódio contra minorias étnicas, de gênero/orientação sexual e religiosas no Facebook. Foi feita a coleta manual de comentários potencialmente ofensivos, sendo o *corpus* composto por 225 comentários de ódio com caráter lgbtfóbico, 194 comentários com teor racista e 181 comentários que apresentam intolerância religiosa, somando 600 comentários. Os dados foram analisados, levando-se em conta alguns trabalhos que tratam da argumentação como prática discursiva (AMOSSY, 2018; CHARAUDEAU, 2008, 2010), bem como trabalhos sobre o discurso de ódio (FORTUNA; NUNES, 2018). Ao final, apresentamos algumas categorias argumentativas para o discurso de ódio, visando à aplicação a análises periciais no âmbito da Linguística Forense e a tarefas de anotação de *corpus* no âmbito da Linguística Computacional.

Palavras-chave: argumentação; discurso de ódio; violência verbal.

Abstract: this article aims to analyze argumentative strategies used in the dissemination of hate speech against ethnic, gender/sexual orientation and religious minorities on Facebook. Manual collection of potentially offensives comments was carried out in a way that the *corpus* is composed by 225 comments with lgbtphobic aspect, 194 comments with racista content and 181 comments which presents possible religious

intolerance, adding up 600 comments. The data was analyzed through works that considers argumentation as a discursive practice (AMOSSY, 2018; CHARAUDEAU, 2008, 2010), as well as works about hate speech (FORTUNA; NUNES, 2018). In the end, we present some argumentative categories of hate speech, which can be applied to judicial expertise, in the scope of Forensic Linguistics, and to *corpus* annotation tasks in the ambit of Computational Linguistics.

Keywords: argumentation; hate speech; verbal violence.

Recebido em 11 de março de 2021

Aceito em 17 de maio de 2021

1 Introdução

Neste trabalho, temos o objetivo de analisar a argumentação empreendida em comentários na rede social Facebook que podem ser considerados portadores de discurso de ódio. A argumentação, tal como compreendida neste trabalho, é justamente a tentativa de influenciar o outro, levando-o a aderir a uma tese levantada pelo argumentante (PERELMAN; OLBRECHTS-TYTECA, 2005), persuadindo-o, convencendo-o ou ainda despertando nele emoções diversas (CHARAUDEAU, 2008, 2010), podendo estar presente em textos dos mais variados tipos e gêneros (AMOSSY, 2018).

A argumentação também pode ser entendida com um sentido próximo ao de retórica, tal como compreende Amossy (2018), ou vista justamente como o objeto da Retórica, que tinha como finalidade o estudo de técnicas argumentativas e provas retóricas (ARISTÓTELES, 2005). Conforme explica Júnior (2005), a retórica parece ter surgido em um momento crucial na história dos povos antigos de Roma e da Grécia, sendo bastante empregada na Antiguidade, principalmente para a negociação e resolução de conflitos de caráter bélico e judicial.

Entretanto, como pretendemos mostrar neste artigo, fruto de uma pesquisa em nível de pós-doutorado, a argumentação e, por extensão, a retórica, não é apenas empregada na resolução de conflitos, mas também em sua instauração. Por meio da análise de 600 comentários que apresentam conteúdos que veiculam discurso de ódio em alguma medida, pretendemos mostrar que o embasamento argumentativo desses

textos pode ser agrupado em determinadas categorias mais ou menos padronizadas e recorrentes.

Recentemente, tem havido um grande embate acerca dos limites da liberdade de expressão, sendo difícil para alguns indivíduos perceber que determinados discursos, ao ofenderem, ameaçarem e atacarem o outro, são passíveis de tipificação jurídica, podendo levá-los a sofrer uma sanção penal (AGUIAR; FARIAS; SALGADO, 2021). Ao apontarmos essas bases argumentativas e retóricas recorrentes, podemos apresentar uma contribuição aos estudiosos do Direito e da Linguística Forense que procuram encontrar indícios que denotem ódio em escritos de redes sociais. Este trabalho ainda pode contribuir para estudiosos da Ciência da Computação e da Linguística Computacional no que concerne à anotação manual ou automática de discursos de ódio, já que uma das tarefas do processamento de linguagem natural (PLN) é justamente a de anotação linguística de *corpora* que apresentem teor odioso para o aprendizado de máquinas e detecção automática de discurso de ódio na internet (cf. BAUMGARTEN *et al.*, 2019; ROß *et al.*, 2016).

Na primeira seção deste artigo, apresentamos uma breve retomada de noções retóricas nos estudos discursivos atuais, nomeadamente na Teoria Semiolingüística do Discurso (CHARAUDEAU, 2008, 2010, 2015) e na Teoria da Argumentação no Discurso (AMOSSY, 2018). Em seguida, discorreremos sobre os procedimentos metodológicos levados a cabo na coleta e análise do *corpus* desta investigação. Por fim, tecemos a análise de estratégias argumentativas recorrentes nos diferentes grupos de comentários de ódio provindos do Facebook tomados como *corpus*.

2 A Retórica “redescoberta” no século XX

Nesta seção, partiremos dos estudos aristotélicos acerca das técnicas argumentativas e desembocaremos em estudos contemporâneos que se valem de algumas noções apresentadas pelo estagirita, reformulando-as e adaptando-as às necessidades epistemológicas atuais da Análise do Discurso.

Em sua Retórica, Aristóteles (2005) nos apresenta dois meios de prova: a prova técnica e a prova retórica. Enquanto a primeira consistiria em alguma prova material, como um punhal sujo de sangue sendo a principal arma de um crime, a segunda seria provinda da argumentação e do emprego do discurso. Assim, Aristóteles nos apresenta as provas

retóricas do *logos*, do *ethos* e do *pathos*. O *logos* estaria relacionado ao dizer propriamente dito, ao pensamento lógico – palavra cuja etimologia já evoca o termo grego *logos*. O *ethos* estaria relacionado ao caráter do orador, se ele seria digno de nota e de fé. O *pathos*, por sua vez, dizia respeito a comover o auditório ao provocar, nele, diferentes emoções.

Ao contrário de outros filósofos da Antiguidade, como Cícero, que questionava a validade do emprego de argumentos pautados na emoção, dando preferência aos argumentos racionalizantes, Aristóteles separou uma importante parte de sua Retórica para o tratamento das paixões, como a cólera, a ira e o medo.

Durante o percurso pela Idade Média, a disciplina Retórica continuou tendo bastante espaço nas universidades europeias, compondo, ao lado da gramática e da lógica, o chamado *Trivium* – conjunto de três disciplinas relacionadas à expressão e à linguagem. Conforme nos explica Plantin (2013), entretanto, no percurso histórico, a lógica acabou suplantando a retórica, tomando para si a tarefa de estudar os argumentos. Essa mudança de paradigma se deveu ao forte impacto do positivismo de Comte, encontrando em René Descartes e em seu racionalismo cartesiano a cisão entre a matéria (*res extensa*) e a mente (*res cogitans*). As emoções, como sendo experienciadas pelo corpo, estariam em um patamar inferior ao raciocínio lógico, relacionado à capacidade mental que separa o homem dos outros animais.

Foi somente na segunda metade do século XX, no pós-Guerra, que a Retórica voltou novamente a ocupar um lugar na academia. Por meio dos estudos das técnicas argumentativas copiladas exaustivamente por Perelman e Olbretchs-Tyteca, observamos o retorno da retórica enquanto disciplina que se ocupa do estudo das técnicas argumentativas. O interesse da argumentação, no *Tratado da Argumentação: a nova Retórica*, era justamente o de levar o auditório a aderir às teses levantadas pelo orador (PERELMAN; OLBRETCHS-TYTECA, 2005). Conforme explica Lima:

O interesse por ela [argumentação] ressurgiu nos anos de 1960 após o longo jejum decorrente da desconfiança em torno dos estudos ligados à retórica. Todavia, esse interesse ainda não foi suficiente para angariar adeptos à sua causa. Apenas nos anos de 1990, os estudos da argumentação ganharam maior força, destacando obras já produzidas nas décadas posteriores, como as

de Perelman & Olbrechts-Tyteca, Toulmin e Ducrot, bem como introduzindo novas idéias com a atual safra de pesquisadores do assunto, entre os quais eu poderia destacar, por exemplo, Christian Plantin, Ruth Amossy, Van Eemeren e Ekkhard Eggs (LIMA, 2006, p. 84).

Além dos autores elencados por Lima (2006), no âmbito da Análise do Discurso, encontramos na Teoria Semiolinguística do Discurso, proposta pelo linguista Patrick Charaudeau, bem como na Teoria da Argumentação no Discurso, idealizada por Ruth Amossy, um lugar privilegiado reservado a uma releitura das provas retóricas do *ethos* e do *pathos*. Em seu livro sobre o discurso político, Charaudeau (2015) categoriza algumas imagens discursivas possíveis de serem produzidas pelo enunciador que procura convencer um determinado público, como o *ethos de virtude*, o *ethos de inteligência* e o *ethos de potência*, dentre outros. A noção de *ethos* como caráter, portanto, é estendida, sendo agora compreendida como a imagem que o enunciador constrói de si mesmo em seu discurso, pois a compreensão dessa imagem pelo destinatário é dependente de seu próprio universo de crença (CHARAUDEAU, 2015), noções também compartilhadas por Maingueneau (2008). Este último faz a distinção entre o *ethos prévio* – construído anteriormente ao discurso –, o *ethos dito* – imagem explicitamente verbalizada –, o *ethos mostrado* – constituído em conjunto com as ações do locutor – e o *ethos efetivo* – a imagem percebida pelo destinatário. Assim, um palestrante pode procurar construir para si a imagem de inteligente, mas alguém na plateia, pelo contrário, considerá-lo pedante ou arrogante.

No que concerne ao *pathos*, encontramos também na Semiolinguística a reformulação dessa prova aristotélica. Charaudeau (2010) apresenta a noção de *patemização* como sendo a capacidade que certos discursos apresentam de despertar determinadas emoções no interlocutor. Novamente, o despertar de uma ou outra emoção estará sempre condicionada ao universo de crença do interlocutor. Assim, ao contar uma piada, o enunciador pode esperar despertar um sentimento de alegria no destinatário, que por sua vez pode achar a piada de mau gosto e experienciar, ao invés, um sentimento de descontentamento. Conforme adverte o linguista:

A análise do discurso não pode se interessar pela emoção como realidade manifesta, vivenciada por um sujeito. Ela não possui os meios metodológicos. Em contrapartida, ela pode tentar estudar o processo discursivo pelo qual a emoção pode ser estabelecida, ou seja, tratá-la como um *efeito visado* (ou *suposto*), sem nunca ter a garantia sobre o *efeito produzido* (CHARAUDEAU, 2010, p. 34).

A patemização está na base do que entendemos aqui como discurso de ódio, visto ser este empregado com a visada discursiva de *fazer-sentir*, ofendendo, injuriando, ameaçando e deslegitimando o grupo-alvo do ataque. Charaudeau (2008) compreende que todo fazer discursivo se dá no interior de uma situação de comunicação do qual participam os protagonistas, o sujeito enunciador e o sujeito destinatário no circuito interno – espaço do dizer –, e os participantes da interação, o sujeito comunicante e o sujeito interpretante no espaço externo – espaço do fazer. O sujeito comunicante, enquanto ser empírico, ao enunciar, toma a vez de sujeito enunciador. Todo uso linguístico é também direcionado a um sujeito destinatário, aquele para quem é dirigido o discurso. O sujeito interpretante, por sua vez, é todo indivíduo que entre em contato com o discurso produzido pelo sujeito comunicante.

Outra noção basilar na Teoria Semiociológica do Discurso é a de contrato de comunicação. Toda troca discursiva obedece a uma série de regras que ditam o que pode ou não ser dito naquela interação. Ao lado do espaço de restrições, entretanto, há o espaço de manobra, em que o sujeito lança mão de estratégias discursivas para atingir a finalidade de sua interação. Essas estratégias são descritas por Charaudeau (2008) como aquelas relacionadas à legitimidade do sujeito enunciador, que procurará demonstrar que tem o direito à palavra, a estratégia de credibilidade, por meio da qual se tenta levar o destinatário a crer naquilo que se diz, e a captação, que consiste em persuadir o auditório, ou sujeito destinatário, por meio, por exemplo, da patemização. O contrato de comunicação dos comentários no Facebook apresenta as regras próprias de seu dizer, como a possibilidade de emprego informal da escrita. Porém, há também um conjunto de restrições do que pode ou não ser dito, visto que o próprio Facebook apresenta algumas regras de conduta para os usuários, como a proibição de comentários que incentivem a violência, o *bullying*, o uso de drogas e o discurso de ódio.

No que concerne à argumentação, Charaudeau (2008) explica que temos três elementos necessários a essa prática: uma proposta sobre

o mundo que seja capaz de despertar um questionamento quanto à sua legitimidade – ou seja, ser passível de refutação; um sujeito que se posicione em relação a esse questionamento e desenvolva um raciocínio, visando a estabelecer uma tese acerca dessa proposta; e um sujeito alvo da argumentação. No processo argumentativo, visando à persuasão ou convencimento do alvo, estratégias relacionadas ao *pathos* e ao *ethos* são empregadas.

Já no âmbito da Teoria da Argumentação no Discurso, Amossy diferencia os discursos que apresentam a intenção de ser persuasivos – caracterizam-se por uma *orientação argumentativa* – dos discursos que influenciam formas de ver o mundo, mesmo quando o discurso não seja explicitamente argumentativo – caracterizam-se por uma *dimensão argumentativa*. De acordo com Amossy (2018):

Há, porém, discursos que não se apresentam como ações de persuasão e nos quais a argumentação não aparece como resultado de uma intenção declarada, muito menos de uma programação: ela não está nem aparente, nem implícita e, às vezes, é até negada pelo locutor (como em um artigo de informação, por exemplo). Foi com o objetivo de designar a orientação involuntária ou subrepticamente impressa no discurso, a fim de projetar certa luz sobre aquilo de que ele trata, que escolhemos falar de dimensão argumentativa (AMOSSY, 2018, p. 273).

Tais noções apresentadas por Charaudeau e Amossy subsidiarão nossas análises e reflexões, pois compreendemos que todo enunciado pode ser argumentativo se procura influenciar em maior ou menor grau o destinatário, como é o caso dos discursos de ódio. Estes, por se valerem de uma polarização de pontos de vista e opiniões dissidentes, situam-se no interior do que Amossy chama de discurso polêmico, isto é, “ao choque muitas vezes brutal de opiniões contraditórias que acentuam as diferenças em vez de procurar um consenso viável voltado para a ação comum” (AMOSSY, 2017, p. 29).

Em nossas análises, procuraremos evidenciar como a argumentação é empreendida em discursos de ódio no Facebook, tomando como pano de fundo a noção de que a argumentação é uma prática discursivamente orientada, conforme propõem Charaudeau e Amossy. Na próxima seção, teceremos algumas considerações acerca da metodologia empregada para a coleta do *corpus*.

3 Metodologia

Para esta investigação, procedemos à coleta de comentários potencialmente ofensivos em postagens do Facebook cujas temáticas incidissem sobre questões de raça/etnia, orientação sexual/ identidade de gênero e crença religiosa. O Facebook foi escolhido por ser uma das redes sociais mais usadas pelos brasileiros, sendo também de fácil acesso. Além disso, concordamos com Amossy (2017), quando afirma ser o Facebook a nova praça pública dos debates retóricos, e com Biar e Paschoal (2020, p. 1066), que salientam como os “algoritmos e bolhas [do Facebook] tendem a nos confinar em discussões monofônicas bem pouco democráticas, que tendem ao extremismo e transformam as plataformas em terreno fértil para ascensão escalar de discursos de ódio”. As redes sociais e o Facebook em particular são, portanto, profícuo campo de investigações que se propõem a analisar embates argumentativos, em parte, dada a facilidade de acesso e à relativa sensação de impunidade que provocam.

A coleta do *corpus* foi realizada manualmente, procurando por palavras-chave como “racismo” “homofobia”, “candomblé” e “umbanda”. Nas postagens, optou-se pelos comentários “mais recentes”, pois o Facebook apresenta uma política de excluir comentários ofensivos, embora seja ainda uma tarefa de difícil realização (FORTUNA; NUNES, 2018), bem como os próprios usuários podem denunciar os *posts* e comentários que contenham discurso de ódio e violência verbal.¹ Como a coleta foi manual e não automática, realizamos uma pré-seleção dos textos que compõem o *corpus*, em um total de 600 comentários (11.925 *tokens*,² isto é, itens lexicais). O *corpus* conta com comentários produzidos no ano de 2019 e de 2020.³

¹ O Facebook apresenta sua própria política de uso da rede social, que desincentiva comportamentos violentos. Para mais informações, ver o item *Comportamento violento e criminoso*, no site oficial. Disponível em: https://www.facebook.com/communitystandards/violence_criminal_behavior. Acesso em: 10 mar. 2021.

² A contagem foi realizada por meio de ferramenta de análise estatística de textos disponibilizada pelo Grupo de Linguística e Computação Cognitiva da Insite. Disponível em: <http://linguistica.insite.com.br/>. Acesso em: 14 fev. 2021.

³ Os comentários coletados estão disponíveis no Facebook e não necessitaram de autorização prévia de seus autores para utilização para fins de estudo. Reiteramos, contudo, que as identidades dos autores foram preservadas.

Foram selecionados enunciados que apresentavam diretamente ou implicitamente um teor ofensivo a grupos minoritários, fazendo uso de termos de que, naqueles determinados contextos, poderiam ser tidos como depreciativos, como “preto”, “macaco”, “bicha”, “boiola”, “macumba”, “idolatria” ou comentários que incitavam a violência contra grupos minoritários. Ao mesmo tempo em que essa coleta nos impede de lançar mão de um vasto número de dados, visto não ter se dado por meio de *softwares* automáticos, ela também se mostra mais efetiva no que concerne à seleção do material analisado, pois selecionamos justamente os comentários cujo teor poderia ser considerado um discurso ofensivo, depreciativo e violento.

Levou-se em conta para a coleta discursos que (i) discriminavam um indivíduo ou um grupo devido à sua identidade racial e/ou étnica, fazendo uso de termos ofensivos ou enunciados que insultavam o indivíduo ou o grupo; (ii) discriminavam um indivíduo ou um grupo devido à sua orientação sexual ou identidade de gênero, empregando termos ou enunciados ofensivos que insultavam o indivíduo ou o grupo; (iii) discriminavam um indivíduo ou um grupo devido à sua crença religiosa, empregando termos e enunciados ofensivos que insultavam o indivíduo ou o grupo.

Os comentários foram separados em três categorias, vale saber: enunciados lgbtfóbicos (225 comentários; 4.791 *tokens*); enunciados racistas (194 comentários; 3.847 *tokens*) e enunciados com intolerância religiosa (181 comentários; 3.287 *tokens*). Cada grupo foi separado em três categorias, sendo estas apresentadas na tabela abaixo, juntamente ao número de comentários recolhidos em cada uma:

TABELA 1 – Categorias de separação do *corpus*

Lgbtfobia		Racismo		Intolerância religiosa	
Homofobia	94	Negros	101	Matriz afro	61
Transfobia	97	Índios	80	Islamismo	100
Lesbofobia	34	Judeus	80	Cristianismo	20
Total	225		194		181

Fonte: Elaboração própria.

O material foi obtido por meio de *prints* da tela do computador ou *smartphone*, através da conta pessoal do autor no Facebook, e, posteriormente, os comentários foram transcritos, visando à otimização da análise. Idiossincrasias textuais e ortográficas foram mantidas de acordo com o original. É importante salientar que a identidade dos autores dos comentários foi preservada, de modo a garantir seu anonimato. Desse modo, tanto rostos, quanto nomes e demais formas de identificação foram suprimidos.

É necessário pontuar que, para este texto, por questões de espaço inerentes à composição do gênero textual artigo científico, apresentamos as análises quantitativas referentes a todo o *corpus* de 600 comentários de ódio, mas a análise qualitativa ficou restrita a alguns enunciados representativos de cada categoria.

No processo de revisão de literatura, uma das etapas metodológicas, notamos que pesquisas que se debruçam sobre a produção de discursos de ódio em redes sociais ainda são bastante escassas no Brasil (NASCIMENTO, 2019) e em Portugal (FORTUNA; NUNES, 2018). Mesmo escrevendo em língua portuguesa, alguns autores preferem voltar seu olhar para algum idioma estrangeiro, como o inglês (cf. NASCIMENTO, 2019). Dessa forma, a língua portuguesa em suas variedades ainda carece de um estudo mais aprofundado acerca das características linguístico-discursivas do discurso de ódio produzidos nessa língua. Este trabalho não tem a intenção de sanar a lacuna, mas apenas de contribuir para que se possa melhor compreender a natureza do discurso de ódio produzido em uma rede social, no caso, o Facebook.

Na próxima seção, apresentamos as categorias que encontramos por meio da análise do *corpus*, mostrando como poderão ser empregadas em investigações no âmbito da Linguística Forense e da Linguística Computacional.

4 Categorias argumentativas do discurso de ódio

Nesta seção, apresentamos algumas categorias encontradas por meio da análise dos 600 comentários de ódio que compõem o *corpus*. Iremos nos guiar pela definição de discurso de ódio apresentada por Fortuna e Nunes (2018), que a elaboraram a partir de diversas outras definições advindas das redes sociais Facebook, Twitter, do YouTube, da Comissão da União Europeia, bem como do discurso científico. Para eles, que falam a partir da Ciência da Computação:

Discurso de ódio é o uso linguístico que ataca ou diminui, incita violência ou ódio contra grupos, baseado em características específicas como aparência física, religião, descendência, origem étnica ou nacional, orientação sexual, identidade de gênero, ou outras, podendo ocorrer com diferentes estilos linguísticos, mesmo de forma sutil ou quando humor é usado⁴ (FORTUNA; NUNES, 2018, p. 5, tradução minha).

Fortuna e Nunes (2018) ainda afirmam que, recentemente, houve um aumento do interesse na construção de ferramentas computacionais para o aprendizado de máquina (*machine learning*) no rastreamento de ódio, de modo a tornar o ambiente das redes sociais mais seguro. Segundo eles apontam, há tanto um interesse comercial quanto político por trás disso, já que a disseminação de discurso de ódio pode provocar danos emocionais e psicológicos a suas vítimas, bem como trazer consequências políticas internacionais no caso de discursos xenofóbicos. Em caso recente, no Brasil, um comentário de caráter xenofóbico por parte do então ministro da educação, Abraham Weintraub, provocou um desconforto nas relações diplomáticas entre Brasil e China, por exemplo.

Pelo fato de tais discursos serem tidos como portadores de ódio e, por extensão, violência verbal, há um interesse da Linguística Forense em estudá-los. A Linguística Forense, conforme ensinam Sousa-Silva e Coulthard (2016), consiste em uma área interdisciplinar que busca empregar determinados achados da linguística na resolução de problemas de cunho jurídico. Para os autores, há três grandes áreas na linguística forense: (i) a linguagem jurídica, cujos estudos se debruçam sobre o discurso realizado em práticas judiciais, evidenciando relações assimétricas de poder e as dificuldades de compreensão do “juridiquês”, por exemplo; (ii) a interação em contextos judiciais, cujos estudos se voltam para a forma como a língua é empregada em determinadas práticas jurídicas, como sessões de interrogatório, audiências de conciliação etc.; (iii) e a área chamada de linguagem como evidência, entendida como a Linguística Forense *stricto sensu*, pois se preocupa em estudar a língua empregada em um contexto de crime, podendo ajudar a inocentar ou culpar o produtor de determinado texto.

⁴ “Hate speech is language that attacks or diminishes, that incites violence or hate against groups, based on specific characteristics such as physical appearance, religion, descent, national or ethnic origin, sexual orientation, gender identity or other, and it can occur with different linguistic styles, even in subtle forms or when humour is used”.

Quando algum crime é cometido exclusivamente ou principalmente por meio da língua, como os chamados crimes contra a honra (calúnia, injúria e difamação), a ameaça, o assédio, a apologia a práticas criminosas, o plágio, a extorsão e o estelionato, por exemplo, estamos diante do que Shuy (2005) chama de crimes de linguagem. Pelo fato de tais empregos linguístico-discursivos serem passíveis de tipificação, isto é, serem enquadrados em um tipo penal descrito no Código Penal, sendo seus autores passíveis de uma sanção penal, reunimos esses textos sob o rótulo de *discurso infrator*.

Partindo então para a análise argumentativa propriamente dita, faz-se importante ressaltar que, neste trabalho, a exemplo de Amossy (2016), tomamos os termos retórica e argumentação como intercambiáveis. Não nos debruçamos necessariamente sobre uma das provas retóricas apresentadas por Aristóteles (2005) e recuperadas recentemente por autores da Análise do Discurso (AMOSSY, 2018; CHARAUDEAU, 2008, 2010, 2015; MAINGUENEAU, 2008), vale saber, as estratégias do *logos*, voltadas para o raciocínio e uso da língua, do *ethos*, voltadas para construção de imagens discursivas de si, e do *pathos*, relacionadas ao despertar de determinadas emoções no alvo.

Apesar de podermos nos valer desses termos em nossas análises, nosso olhar partiu do embasamento retórico levado a cabo pelos sujeitos que produzem discursos de ódio. Isto é, interessa-nos não apenas classificar os enunciados, mas entender as possíveis motivações e fundamentações sociais e políticas que embasam tais construções argumentativas, desde que se apresentem na superfície textual. Notamos que há algumas categorias argumentativas recorrentes, o que nos permitiu fazer um levantamento quantitativo dessas ocorrências.

A seguir, apresentamos as categorias encontradas por meio da análise dos 600 comentários de ódio que compuseram nosso *corpus*. Foram encontradas 20 categorias, sendo que algumas aparecem em todos os grupos de comentários (lgbtfobia, racismo e intolerância religiosa) e algumas são características de apenas um dos grupos. É possível notar que algumas categorias podem ser agrupadas sob o rótulo de *argumentos ad personam* e *argumentos de autoridade*, já trabalhados por autores da argumentação, como Schopenhauer (2014) e Perelman e Olbrechts-Tyteca (2005). Entretanto, visto emergirem temáticas específicas e recorrentes que especificam os empregos de tais estratégias argumentativas, empreendemos uma categorização mais específica,

visando efetivamente contribuir para a compreensão dos diferentes tipos de argumentos empregados em discurso de ódio, e não no fenômeno argumentativo como um todo:

- 1) *Ad baculum*. Descrito por Schopenhauer (2014), o *argumento ad baculum* (literalmente, argumento pelo báculo, porrete) diz respeito ao emprego de ameaças ou argumentos fundamentados sobre a agressão física.
- 2) *Ad personam baseado em suposta enfermidade*. O argumento *ad personam* consiste no ataque direto ao argumentante, pois “a pessoa então será sujeita a humilhações, maldades, afrontas e grosserias”. (SCHOPENHAUER, 2014, p. 77). O *ad personam* baseado em suposta enfermidade consiste na classificação do comportamento do grupo ou indivíduo alvo como doença e enfermidade física ou psicológica.
- 3) *Ad personam baseado em suposta incapacidade*. É um tipo de *argumento ad personam* que ataca as condições físicas ou psicológicas do grupo ou indivíduo alvo, classificando-o como incapaz de realizar alguma atividade.
- 4) *Ad personam baseado em suposta selvageria*. É um tipo de argumento *ad personam* característico dos discursos de ódio contra indivíduos ou grupos indígenas brasileiros, categorizando-os como selvagens e atrasados.
- 5) *Ad personam baseado em suposta vadiagem*. É um tipo de *argumento ad personam* característico dos discursos de ódio contra indivíduos ou grupos indígenas brasileiros, categorizando-os como vadios e preguiçosos.
- 6) *Ad personam baseado em suposta vitimização*. É um tipo de *argumento ad personam* característico dos discursos de ódio contra indivíduos ou grupos negros brasileiros, categorizando seus apelos e lutas como vitimização.
- 7) *Ad personam baseado em ofensa estética*. É o ataque direto à aparência física do indivíduo, podendo se estender a todos os demais indivíduos de seu grupo no caso do julgamento de valor negativo acerca de uma característica étnica e racial.
- 8) *Ad personam baseado em suposto engodo*. Consiste no ataque ao grupo-alvo, acusando-o de engodo, mentira ou enganação.

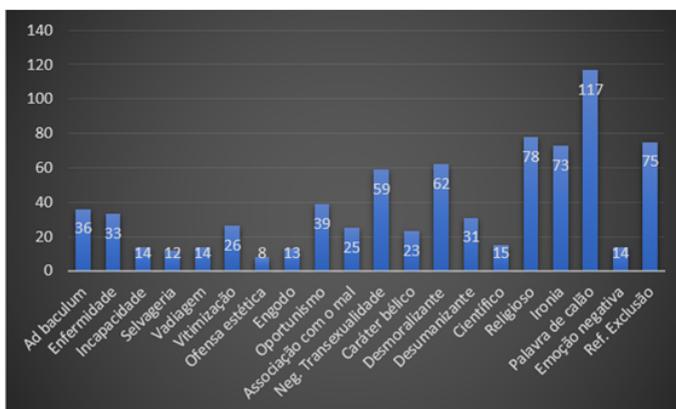
- 9) *Ad personam baseado em suposto oportunismo.* É um tipo de argumento *ad personam* característico dos discursos de ódio contra indivíduos ou grupos indígenas brasileiros. Consiste em uma estratégia que procura deslegitimar o indivíduo ou grupo-alvo, acusando-o de oportunismo, esperteza e vigarice.
- 10) *Ad personam baseado na associação com o mal.* É a categorização do grupo-alvo como maligno ou que faz ações que causam o mal.
- 11) *Ad personam baseado na negação da transexualidade.* Por meio desta estratégia, o enunciador deslegitima a luta e a própria condição transexual, negando sua existência.
- 12) *Ad personam de caráter bélico.* O *ad personam* de caráter bélico é o argumento que ataca o sujeito muçulmano, baseando-se em uma aproximação entre a religião islâmica e o terror da guerra.
- 13) *Ad personam desmoralizante.* É uma espécie de argumento *ad personam* que procura deslegitimar o alvo do discurso de ódio, destituindo-o de algum aspecto moral.
- 14) *Ad personam desumanizante.* É a aproximação do alvo ao animalesco, ao bestial, destituindo-o de sua natureza humana na tentativa de legitimar ódio e intolerância.
- 15) *Argumento de autoridade científico.* É o apelo à ciência ou a um suposto posicionamento científico para embasar o preconceito e intolerância.
- 16) *Argumento de autoridade religioso.* É o argumento fundamentado em preceitos religiosos. Por meio do apelo à religiosidade, principalmente cristã, o argumentante procura categorizar o indivíduo ou grupo-alvo como pecador.
- 17) *Emprego de ironia.* Nem sempre o ataque é direto e linguisticamente explícito, podendo o enunciador, por vezes, valer-se da ironia como uma estratégia para tornar o comentário de ódio implícito. O emprego da ironia pode se sobrepor a outra categoria argumentativa, entretanto, para fins metodológicos, houve a separação em nossas análises.
- 18) *Emprego de palavrões e palavras de calão.* Consiste no emprego de palavrões e palavras de calão para atacar, ofender e injuriar o indivíduo ou grupo-alvo. É descrita por Silva (2020) como uma estratégia de patemização; isto é, uma estratégia argumentativa

que procura despertar emoções no destinatário. Foi a estratégia mais recorrente, sendo empregada em todos os grupos do *corpus*.

- 19) *Expressão de emoção negativa*. Consiste na explicitação linguística da emoção negativa sentida em relação ao grupo ou indivíduo algo.
- 20) *Reforço da exclusão*. É a apologia à exclusão do grupo-alvo, reforçando a separação da sociedade em grupos distintos e relegando ao grupo-alvo um lugar apartado dos demais.

O Gráfico 1 apresenta, em número de ocorrência, o total de estratégias argumentativas encontradas em nosso *corpus*.

GRÁFICO 1 – Estratégias argumentativas por número de ocorrência



Fonte: Elaboração própria.

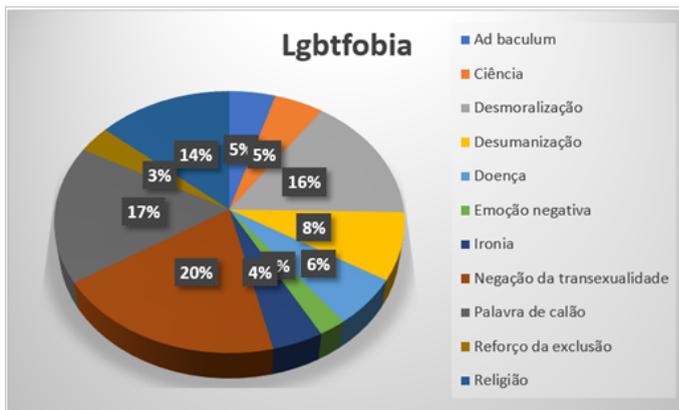
Como podemos perceber, a estratégia relacionada ao *emprego de palavras e palavras de calão* foi a mais empregada no *corpus* (117 ocorrências), seguida pelo *argumento de autoridade religioso* (78 ocorrências) e pelo *reforço da exclusão* (75 ocorrências). Dentre as estratégias elencadas, *argumento ad baculum*, *argumento de autoridade científico*, *argumento ad personam desmoralizante*, *argumento ad personam desumanizante*, *ad personam baseado em suposta enfermidade*, *expressão de emoção negativa*, *ad personam baseado em suposto engodo*, *emprego de ironia*, *ad personam baseado na associação como mal*, *emprego de palavras e palavras de calão*, *reforço da exclusão* e *argumento de autoridade religioso* foram compartilhadas por diferentes

grupos de comentários. Por sua vez, algumas estratégias foram encontradas apenas em grupos específicos: *ad hominem de caráter bélico* (muçulmanos); *ad personam baseado em suposta incapacidade* (negros); *negação da transexualidade* (transexuais e lésbicas); *ofensa estética* (negros); *ad personam baseado em suposto oportunismo* (indígenas); *ad personam baseado em suposta selvageria* (indígenas); *ad personam baseado em suposta vadiagem* (indígenas) e *ad personam baseado em suposta vitimização* (negros). Assim, os comentários de ódio contra grupos étnicos minoritários apresentam maior emprego de estratégias específicas, dadas as características de cada grupo. Embora tenha havido constantes empregos do *argumento ad personam*, a fundamentação para cada categoria é respaldada em diferentes valores, bem como as estratégias linguístico-discursivas se diferenciam de um grupo-alvo para outro, fatores que justificam a separação em categorias argumentativas diferenciadas.

Apresentaremos primeiramente a análise quantitativa e qualitativa dos comentários de ódio/intolerância contra a comunidade LGBTQI+. Classificamos esse grupo como comentários lgbtfóbicos, sendo ainda subdividido em: comentários homofóbicos (107 argumentos nos 94 comentários), quando o alvo eram homens gays ou bissexuais; comentários transfóbicos (144 argumentos nos 97 comentários), quando o alvo eram homens e mulheres trans; e comentários lesbofóbicos (46 argumentos nos 34 comentários), quando o alvo eram mulheres lésbicas. A bifobia não apareceu linguisticamente em nosso *corpus*; assim, mesmo que o *post* falasse sobre indivíduos bissexuais, os ataques evocavam termos e argumentos que se referiam à homossexualidade. Por questões metodológicas, mas sem pretendermos cair no erro do reducionismo, inserimos os comentários homofóbicos e bifóbicos na mesma categoria. O número de comentários e de argumentos em cada grupo não é condizente, visto a extensão de cada comentário ser diferente, normalmente abarcando mais de uma estratégia argumentativa.

No Gráfico 2, apresentamos os argumentos encontrados nessas categorias, por porcentagem de ocorrência, levando em conta o total de argumentos (297 argumentos em 225 comentários).

GRÁFICO 2 – Comentários com teor lgbtfóbico



Fonte: Elaboração própria.

O *ad personam* baseado na *negação da transexualidade* representa 20% do total de argumentos de ódio, sendo a categoria mais comum neste grupo, embora tenha aparecido uma vez no grupo de comentários lesbofóbicos e 58 vezes no grupo de comentários transfóbicos, representando a maior ocorrência neste subgrupo. Fazem parte dessa categoria os argumentos que procuram deslegitimar a natureza trans dos indivíduos e grupos atacados, sejam homens trans, mulheres trans, travestis, sejam não binários, conforme transparece no comentário abaixo:

- a) Mulher trans ou é mulher ou é homem. Trans não define gênero. Aliás não define porra nenhuma. Palhaçada.

Nessa estratégia, é comum a negação da paternidade e da maternidade por parte de indivíduos trans, bem como a não aceitação do nome social, adquirido após constantes lutas judiciais por parte da comunidade LGBTQI+. Por meio dessa estratégia, o argumentante procura deslegitimar o indivíduo trans, atacando-o para justificar seu pensamento contrário. Nota-se que o sujeito enunciador procura atingir a comunidade LGBTQI+ por meio de um argumento patemizante que teria a capacidade de despertar em determinados sujeitos interpretantes, a depender de seu próprio universo de crença, emoções como indignação e revolta.

O enunciado acima, que defende a tese de que o termo “trans” não configura um gênero, também evidencia outra estratégia bastante utilizada, o *emprego de palavrões e palavras de calão* (porra), sendo a segunda estratégia argumentativa mais recorrente, surgindo em 17% das vezes. Por meio do *emprego de palavrões e palavras de calão*, é possível levar o destinatário a experimentar diversas emoções, apresentando alto teor patêmico (SILVA, 2020) e, em nosso *corpus*, dada a natureza ofensiva e humilhante dos discursos de ódio, é possível que emoções como a raiva e a indignação sejam experienciadas pelos sujeitos-alvo. Se estes se sentirem lesados moralmente e/ou psicologicamente, tais sujeitos podem recorrer à justiça, o que faz com que os comentários contendo ódio e intolerância sejam tidos como discursos passíveis de tipificação.

A terceira estratégia mais recorrente foi o *Argumento ad personam baseado na desmoralização*, ocorrendo 16% das vezes. Como vimos anteriormente, por meio dessa estratégia, o sujeito enunciator procura deslegitimar a vítima ou o grupo-alvo. Ocorreu no conjunto de comentários com teor lgbtfóbico, principalmente, para construir a imagem discursiva do alvo como pessoas sem moral, sem escrúpulos, voltadas a obscenidades e sexo desregrado, atitudes tidas como condenáveis pelo enunciator, como no exemplo abaixo:

- b) Falta de vergonha na cara, vai direto morar com o capeta, ao invés de procurar a salvação, fica procurando depravação, francamente, este mundo tem que pegar fogo e ser consumido

Ao categorizar o outro como imoral, o sujeito enunciator procura construir para si um *ethos de virtude* (CHARAUDEAU, 2015), de alguém que não coaduna com práticas “depravadas”, em seu ponto de vista. O teor altamente patêmico dos enunciados apresentados em (b) é percebido por meio do emprego de termos relacionados à moralidade, como “vergonha na cara”, “depravação” e à religiosidade, como “capeta” e “salvação”.

A estratégia argumentativa que busca aproximar a homossexualidade e transexualidade à enfermidade também foi bastante recorrente (6%):

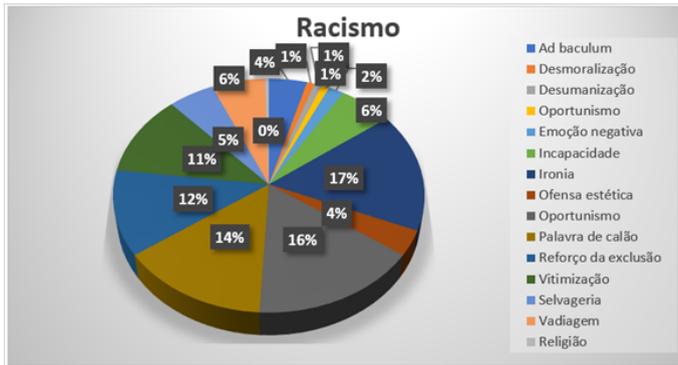
- c) Temos que entender que são pessoas doentes... tem que buscar tratamento o quanto antes

No comentário acima, percebemos a construção do *ethos* do enunciador como alguém supostamente preocupado com a saúde da população LGBTQI+. Até recentemente, a homossexualidade figurava entre as doenças psicológicas, justamente o sentido denotado pelo sufixo *-ismo* em “homossexualismo”; entretanto, no Brasil, desde 1985, o Conselho Federal de Psicologia deixou de considerar o comportamento homossexual como doença. Em 1990, a Organização Mundial da Saúde retirou a homossexualidade da lista de doenças mentais. Mesmo com todos esses avanços, em 2011, o deputado federal João Campos (PSDB) sugeriu um Projeto de Decreto Legislativo (PDL 234/11), apresentando o que ficou conhecido como “Cura Gay” – um conjunto de práticas que, contrariando a resolução do Conselho Federal de Psicologia, procurava reorientar o comportamento sexual de indivíduos gays. Após sua aprovação, o projeto foi arquivado, não sem deixar claro, assim como os comentários encontrados em nosso *corpus*, que ainda há muita desinformação e preconceito acerca do comportamento homossexual.

Passemos à análise qualitativa e quantitativa dos argumentos encontrados no conjunto de comentários classificados como possível⁵ racismo (228 argumentos em 194 comentários). Os comentários foram subdivididos em: teor racista contra negros (116 argumentos em 101 comentários); teor racista contra indígenas brasileiros (98 argumentos em 80 comentários); e teor racista contra judeus (14 argumentos em 13 comentários). Dos diversos comentários coletados, escolhemos alguns mais ilustrativos do que pretendemos demonstrar neste artigo, pois há limite de espaço para explanarmos sobre uma quantidade maior de modo qualitativo e aprofundado.

⁵ A modalização se faz necessária, visto que há diferenças entre os tipos penais do racismo e da injúria racial, bem como o crime só será tido como tal após o trânsito em julgado.

GRÁFICO 3 – Comentários com teor racista



Fonte: Elaboração própria.

Como podemos perceber no gráfico 3, o argumento pautado no *emprego de ironia* foi o mais recorrente, representando 17% do total de argumentos. A categorização de argumentos como irônicos foi feita principalmente por meio de escolhas lexicais, pontuação, emoticons e alguma pista linguística que denotasse a prosódia impressa sobre o comentário, visto estarmos lidando com textos escritos. Por meio da ironia, o sujeito argumentante denota seu ódio e desprezo pelo indivíduo ou grupo-alvo, mas de modo implícito, talvez uma tentativa de se defender de alguma possível acusação de racista – rótulo não desejado. Entretanto, em vias legais, um linguista forense pode ser chamado para averiguar o significado de um determinado comentário de ódio implícito, caso o juiz considere que o texto apresentado levante dúvidas (SILVA, 2020). Essa área é conhecida na Linguística Forense como Determinação de Significado (SOUZA-SILVA; COULTHARD, 2016). O emprego de ironia no discurso de ódio também foi atestado por Van Hee, Lefever e Hoste (2018), que afirmam que 34% dos tweets irônicos que analisaram em sua investigação são ataques que procuram ridicularizar ou causar dano a alguém. Assim, conforme Muñoz (2020, p. 243), e aqui atestamos, “a violência verbal pode se realizar de modo explícito ou implícito”.

No comentário a seguir, essa questão fica salientada:

d) Índio? Ainda existe isso?? Acho obsoleto demais né.

Nesse comentário, observamos uma demonstração de insatisfação quanto à existência dos povos indígenas, bem como em relação às suas lutas e direitos garantidos. Por meio das perguntas retóricas, o autor do comentário procura expressar uma tentativa de humor, dizendo que os índios são obsoletos, isto é, atrasados. Ocorre, portanto, não apenas o emprego da ironia que denota uma insatisfação, mas também uma desumanização do indivíduo indígena por meio do pronome demonstrativo “isso” usado na referenciação. A estratégia argumentativa do *ad personam desumanizante* foi a segunda mais recorrente, representando 16% do total.

No conjunto de comentários com teor racista contra negros, destacou-se a prevalência de argumentos que os classificavam como vitimistas (26 oc.), como o que segue:

- e) O racismo só vai acabar quando os negros pararem de ficar se vitimizando e ir em busca de seu próprio futuro

O comentário acima é um dos que podem gerar discussão quanto ao entendimento acerca de o texto consistir em um comentário de ódio ou uma livre expressão de opinião. Entretanto, ao relacionar toda a luta e reivindicação por igualdade por parte do povo negro a uma suposta vitimização, ocorre uma tentativa de silenciamento. Ir em busca de seu próprio futuro parece retomar intertextualmente o discurso da meritocracia, entretanto, mesmo negros bem-sucedidos não deixam de denunciar casos explícitos de racismo, o que invalida o argumento acima apresentado. No processo, o sujeito enunciador parece pretender construir para si um *ethos de virtude*, de alguém preocupado com o fim do racismo. A evocação do termo “vitimização” inserido pelo verbo “parar”, que introduz o pressuposto de que os negros se vitimizam, apresenta forte teor patêmico capaz de levar alguns sujeitos interpretantes a experienciar sentimentos relacionados à indignação e à raiva. Por outro lado, a depender de seus imaginários e crenças, outros sujeitos interpretantes seriam levados a um sentimento de satisfação, por concordarem com o argumentante.

Outro argumento que aparece apenas no conjunto de comentários destinados a negros foi ao *ad personam baseado na ofensa estética*, ocorrendo 8 vezes:

- f) Maior cara de bandido esses malucos tem

Esse comentário acerca de dois homens negros evidencia o julgamento estético do negro como criminoso, algo já bastante discutido pela Teoria do Etiquetamento Social (*Label Approach*), vinculada à Criminologia Crítica (cf. BARATTA, 1999). Os dois rapazes em questão eram estagiários do Núcleo de Esportes da Globo, pretos e vestidos casualmente. O que o comentário evidencia é, portanto, uma ofensa aos indivíduos negros simplesmente pela sua aparência (cara de bandido), além da aproximação de um grupo inteiro com a criminalidade.

Quanto ao conjunto de comentários racistas dirigidos a indivíduos ou grupos indígenas, observou-se um destaque maior para os argumentos pautados sobre o *suposto oportunismo*, representando 16% do total de argumentos, mas aparecendo apenas dentre os comentários contra indígenas (37 oc.):

- g) Índios não existem há séculos! Hoje são bandidos aproveitadores vagabundos que não gostam de trabalhar e se passam por índios, para sugar o governo, e fazer pedágio em rodovias para roubar pessoas. E nem flechas usam mais, estou de pistolas! Kkk

Evidencia-se, no comentário em apreço, uma tentativa de apresentar os indígenas brasileiros como oportunistas, sendo chamados de “aproveitadores vagabundos que não gostam de trabalhar”, preferindo “sugar o governo” e “roubar pessoas”, enunciados que funcionam como argumentos para comprovar a tese de que “índios não existem há séculos”. Além de oportunistas, os indígenas também são classificados como vadios e vagabundos, o que representa 6% do total (14 oc.). Esses comentários são fundamentados sobre um imaginário sociodiscursivo intolerante acerca dos povos indígenas como atrasados, que nada produzem para o avanço capitalista do País.

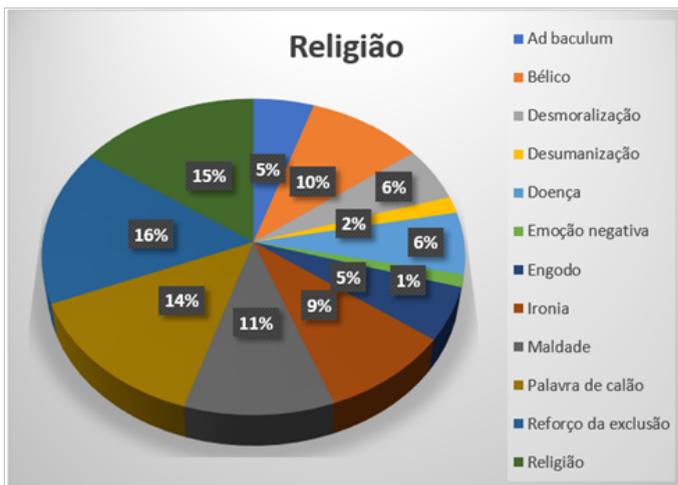
Quanto aos comentários antissemitas, observou-se a prevalência de argumentos fundamentados sobre o *reforço da exclusão* (5 oc.) O número é baixo em relação ao total, mas representativo em relação aos comentários de ódio contra o povo judeu, pois apenas 13 comentários foram encontrados nessa categoria:

- h) Judeu é um povo que plantou o que colheu. Eles não aceitam outros q não sejam judeus. Não encostam em quem não são judeus. Não conversam com quem não são judeus. Eu morei em londres e vi com meus próprios olhos estas atitudes.

Há, no comentário acima, uma tentativa de evidenciar as diferenças entre o “nós” e o “eles”, sendo “eles” o povo judeu. Há ainda a justificativa para as atrocidades cometidas contra o povo judeu no curso da História, atrocidades que, na visão do enunciador, são justificadas pelas próprias ações cometidas por esse povo. Uma estratégia argumentativa ainda empregada pelo enunciador, mas não classificada por nós por não veicular necessariamente discurso de ódio, foi a narrativa de uma situação vivida para fundamentar a argumentação, prática descrita por Charaudeau (2008). O sujeito procura legitimar seu argumento por meio da narrativa de que, morando em Londres, presenciou *in loco* as atitudes que ele atribui aos judeus: a autoexclusão.

A seguir, apresentamos os argumentos encontrados no grupo de comentários de ódio contra religiosos. É importante salientar que os judeus consistem não apenas em um grupo religioso, mas também étnico, motivo pelo qual os inserimos no grupo analisado anteriormente. No âmbito dos comentários de ódio contra religiosos, separamos os comentários que tinham como grupo-alvo os fiéis de religiões de matriz africana, cristãos e muçulmanos. Apesar de os cristãos serem maioria no Brasil, lembramos que o discurso de ódio não se fundamenta apenas contra grupos minoritários, embora esses sejam os grupos preferíveis e que necessitam de maior proteção por parte do Estado.

GRÁFICO 4 – Comentários de ódio com caráter de intolerância religiosa



Fonte: Elaboração própria.

Como podemos perceber a partir da leitura do gráfico apresentado, a estratégia vinculada ao *reforço da exclusão* foi a mais representativa, tendo 38 ocorrências, o que corresponde a 16% do total. Nesse tipo de argumento, situam-se aqueles que objetivam excluir o indivíduo ou grupo-alvo como não fazendo parte do grupo ao qual o enunciador pertence, o que justificaria o ódio. Por exemplo:

i) Não aceitem é macumba!

O termo “macumba” aparece como uma metonímia de qualquer religião de matriz africana, seja candomblé, seja umbanda ou outra religião afro-brasileira. No enunciado acima, a temática era referente aos doces e comidas distribuídas por fiéis no dia de São Cosme e Damião, uma tradição em diversas localidades no Brasil. Ao solicitar enfaticamente, representado pelo sinal de exclamação, que as pessoas não aceitem os doces ofertados por serem “macumba”, isto é, algo negativo, o enunciador reforça a exclusão entre os grupos, entre o “nós” e o “eles”.

A estratégia mais empregada no grupo de comentários de ódio contra indivíduos ou grupos que professam alguma fé de matriz africana foi o argumento pautado na premissa de que tais religiões são malignas (19 oc.), como no exemplo a seguir:

j) Palhaçada! Uma religião que sacrifica os animais. Deveria ser banida. Religião prega amor e não a maldade ao próximo.

No exemplo acima, há evidentemente uma aproximação entre as religiões de matriz africana e a maldade, tanto em relação a animais, quanto em relação “ao próximo” – termo bastante empregado no discurso religioso. Há ainda a apologia ao banimento dessa religião na tese defendida pelo sujeito enunciador, o que também contraria os Direitos Humanos de livre expressão religiosa.

Outra estratégia bastante comum foram os argumentos baseados em preceitos religiosos, funcionando como um argumento de autoridade. Essa categoria ocorreu 36 vezes, o que representa 15% do total. Por meio dessa estratégia, o sujeito enunciador ataca o indivíduo ou grupo-alvo como sendo satanista, pecador, idólatra, merecedor do Inferno e como não fazendo a vontade de Deus. Como vimos, essa estratégia foi bem recorrente no grupo de comentários com teor lgbtfóbico, mas também é bastante evidenciada no grupo de comentários contendo intolerância

religiosa, sendo a mais recorrente no grupo de comentários destinados a indivíduos ou grupos islâmicos (23 oc.). Isso acontece pelo fato de que cada religião apresenta sua própria doutrina e conjunto de dogmas, tidos pelos fiéis como os únicos verdadeiros e corretos. Tudo o que foge a esse ponto de vista espiritual seria taxado como um desvio dos ensinamentos de sua própria religião e, por isso, tido como errado:

- k) Religião de sanatás, dos destruidores, que explodem aviões, não respeitam as mulheres e nem a religião dos outros, que ensinam as crianças a serem bombas

O argumento acima é dirigido à religião islâmica e seus seguidores. Como pode ser notado, há um argumento fundamentado em critérios religiosos, defendendo que o Islã é a “religião de Satanás”, de acordo com aquele ponto de vista. Como já mencionamos, um mesmo comentário pode abarcar estratégias argumentativas diferentes, como neste, em que há também o que chamamos de *argumento ad personam de caráter bélico*, encontrado apenas no conjunto de comentários cujo indivíduo ou grupo-alvo pertence à religião islâmica. Por meio desse argumento, o grupo é relacionado à guerra e à destruição, a exemplo de sua denominação como “destruidores” “que explodem aviões” e “que ensinam as crianças a serem bombas”, argumentos altamente patemizantes que podem despertar emoções diferentes, levando-se em conta o universo de crença do sujeito interpretante: se este conhecer a religião islâmica e souber que apenas algumas facções terroristas cometem tais atrocidades, pode experimentar um sentimento de indignação pela difamação contra esse grupo religioso; por outro lado, caso não conheça bem a doutrina islâmica, o sujeito interpretante pode ser levado a se indignar justamente contra o Islã e seus praticantes. Há também uma tentativa de desmoralização da cultura e religião islâmica, ao afirmarem que ela não respeita as mulheres nem as demais religiões.

Por fim, o terceiro argumento mais representativo, seguindo a tendência geral do *corpus*, foi o *emprego de palavras e palavras de calão*, ocorrendo 33 vezes, o que corresponde a 14% do total, como observamos em:

- l) É por conta disso q eu acho q a Bíblia tinha que ser banida e enfiada no cu de cada um q segue esse livro de merda, a única coisa q a Bíblia consegue trazer é o caos... o mundo era ótimo antes disso existir

O comentário acima não nomeia um grupo específico, mas ao se referir à Bíblia e a “cada um q segue esse livro de merda”, o sujeito enunciativo se refere explicitamente aos cristãos, visto que judeus e muçulmanos, apesar de seguirem em partes alguns textos bíblicos, apresentam sua própria compilação de textos sagrados, vale saber, o Torá e o Corão, respectivamente. O teor odioso é denotado pelos trechos “enfada no cu” e “livro de merda”, configurando uso da violência verbal. Há uma tomada de posição não cristã por parte do sujeito enunciativo, que afirma, inclusive, que a Bíblia deveria ser banida, o que atenta contra os Direitos Humanos de livre expressão e manifestação religiosa. Entretanto, por vezes, um grupo religioso cristão pode dirigir ofensas odiosas a outro grupo cristão, como protestantes *versus* católicos.

Para este texto, como já evidenciado, apresentamos as análises quantitativas referentes a todo o *corpus* de 600 comentários de ódio, mas a análise qualitativa ficou restrita a alguns enunciados representativos de cada categoria.

A seguir, tecemos algumas considerações finais.

Considerações finais

Neste artigo, partimos de um ponto de vista linguístico-discursivo, procurando evidenciar estratégias argumentativas empregadas em discursos de ódio produzidos em português brasileiro na rede social Facebook.

Os dados foram coletados manualmente, sendo obtidos 600 comentários ao total, que apresentassem comentários com teor lgbtfóbico, racista e de intolerância religiosa. Cada um desses grupos foi subdividido da seguinte maneira: comentários possivelmente lgbtfóbicos – homofóbicos/bifóbicos, transfóbicos e lesbofóbicos –; comentários possivelmente racistas – contra negros, índios e judeus –; comentários com possível intolerância religiosa – contra religiões de matriz africana; muçulmanos e cristãos.

Ao final da análise, pudemos categorizar vinte estratégias argumentativas empregadas para atacar, humilhar, diminuir e ofender os grupos-alvo, sendo algumas encontradas em todos os grupos, como o *emprego de palavras e palavras de calão*, enquanto outras são particulares a determinados grupos-alvo, como o *argumento ad personam baseado na negação da transexualidade* contra pessoas trans, o *argumento*

ad personam de caráter bélico contra os muçulmanos e a *argumento ad personam baseado na ofensa estética* dirigida a indivíduos negros.

Notamos, portanto, que há um conjunto de temáticas que subsidiam a argumentação pautada no discurso de ódio contra os grupos selecionados. Munido dessas vinte categorias por nós descritas e exemplificadas, linguistas computacionais e cientistas da computação que procuram anotar *corpora* com *tagsets* relacionados ao discurso de ódio podem generalizar, complementar ou refutar esses dados, visando a otimizar o aprendizado de máquinas (*machine learning*) no que concerne, principalmente, à detecção automática de discursos de ódio na internet.

Além disso, profissionais do Direito e linguistas forenses podem se valer de tais categorias analíticas para a resolução de dúvidas no que concerne à determinação de significado odioso em comentários em redes sociais ou mesmo em outros gêneros. Isto é, essas categorias podem ser empregadas para a resolução de casos em que o teor ofensivo de um comentário ou outro gênero textual seja passível de disputa judicial, o que pode ser considerado um passo na tarefa de apresentar pistas linguístico-discursivas que ajudem a delimitar os limites entre o discurso de ódio e a liberdade de expressão, tão necessária, mas tão mal-empregada nas redes sociais.

Salientamos ainda que trabalhos futuros podem ser desenvolvidos a partir dessa primeira categorização metodológica aqui apresentada, como: (i) a anotação de *corpora* maiores, visando a perceber se tais categorias argumentativas podem ser estendidas a outros gêneros textuais e a outros grupos-alvo de ataques de ódio, (ii) a análise da patemização possivelmente engendrada por esses discursos; (iii) o *ethos* construído pelos sujeitos que proferem discursos de ódio e (iv) o potencial didático desta investigação.

Referências

AGUIAR, L. R. L.; FARIAS, S. S.; SALGADO, A. A. R. T. Liberdade de expressão nas redes sociais: a resignificação do Direito Penal e Processual Penal à luz dos discursos de ódio. In: ALMEIDA, F. I. L.; VIEIRA, S. G. N.; CAVALCANTI, S. C. M. (org.). *Criminalidade na era digital*. João Pessoa: ADEPDEL, 2021. p. 123-134.

AMOSSY, R. *Apologia da polêmica*. São Paulo: Contexto, 2017.

AMOSSY, R. *A argumentação no discurso*. São Paulo: Contexto, 2018.

ARISTÓTELES. *Retórica*. 2. ed. Lisboa: Imprensa Nacional Casa da Moeda, 2005.

BARATTA, A. *Criminologia crítica e crítica do Direito Penal*. Introdução à Sociologia do Direito Penal. 3. ed. Trad. Juarez Cirino dos Santos. Rio de Janeiro: Editora Revan, 1999.

BAUMGARTEN, N. *et al.* Towards Balance and Boundaries in Public Discourse: Expressing and Perceiving Online Hate Speech (XPEROHS). *RASK: International Journal of Language and Communication*, Odensev. 1, n. 50, p. 87-108, 2019. Disponível em <https://www.researchgate.net/publication/337913890>. Acesso em: 28 fev. 2021.

BIAR, L. A.; PASCHOAL, F. V. C. “(Não) leia os comentários”: a disputa da notícia sobre o assassinato de Marielle Franco. *Trab. Ling. Aplic.*, Campinas, v. 59, n. 2, p. 1051-1069, 2020. DOI: <https://doi.org/10.1590/01031813679571620200330>

CHARAUDEAU, P. *Linguagem e discurso: modos de organização*. São Paulo: Contexto, 2008.

CHARAUDEAU, P. A patemização na televisão como estratégia de autenticidade. In: MENDES, E.; MACHADO, I. L. (org.). *As emoções no discurso*. Campinas: Mercado das Letras, 2010. v. II, p. 23-56.

CHARAUDEAU, P. *Discurso Político*. São Paulo: Contexto, 2015.

FORTUNA, P.; NUNES, S. A Survey on Automatic Detection of Hate Speech in Text. *ACM Comput. Surv.* Nova Iorque, v. 51, n. 4, p. 1-30, 2018. DOI: <https://doi.org/10.1145/3232676>

JÚNIOR, M. A. Origem da retórica e formação do sistema retórico. In: ARISTÓTELES. *Retórica*. 2. ed. Lisboa: Imprensa Nacional Casa da Moeda, 2005. p. 15-84.

LIMA, H. M. R. *Na tessitura do Processo Penal: a Argumentação no Tribunal do Júri*. 2006. 260f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2006.

MAINGUENEAU, D. A propósito do ethos. In: MOTA, A. R.; SALGADO, L. (org.). *Ethos discursivo*. São Paulo: Contexto, 2008. p. 11-29.

MUÑOZ, S. C. Es posible detección automática de la violencia lingüística em las redes sociales? *E-AESLA*, Vigo, n. 6, p. 241-252, 2020.

NASCIMENTO, R. M. F. *Classificação automática de discursos de ódio em textos do Twitter*. 2019. 48f. Trabalho de conclusão de curso (Bacharelado em Sistemas de Informação) – Curso de Bacharelado em Sistemas de Informação, Universidade Federal Rural de Pernambuco, Recife, 2019.

PLANTIN, C. *A Argumentação*. São Paulo: Parábola Editorial, 2013.

PERELMAN, C.; OLBRECHTS-TYTECA, L. *Tratado da Argumentação: A Nova Retórica*. São Paulo: Martins Fontes, 2005.

ROß, B. *et al.* Measuring the Reliability of Hate Speech Annotations: The Case of the European Refugee Crisis. In: WORKSHOP ON NATURAL LANGUAGE PROCESSING FOR COMPUTER-MEDIATED COMMUNICATION, 3., 2016, Bochum. *Anais [...]*. Bochum: Bochumer Linguistische Arbeitsberichte, 2016. p. 6-9. DOI: <https://doi.org/10.17185/dupublico/42132>

SCHOPENHAUER, A. *38 Estratégias para se vencer qualquer debate. A arte de ter razão*. São Paulo: Faro Editorial, 2014.

SHUY, R. W. *Creating Language Crimes: How Law Enforcement Uses (and Misuses) Language*. New York: Oxford University Press, 2005. DOI: <https://doi.org/10.1093/acprof:oso/9780195181661.001.0001>

SILVA, W. P. *Argumentação e patemização em cartas de ameaça: uma análise semiolinguística como contribuição à Linguística Forense*. 2020. 275f. Tese (Doutorado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020.

SOUSA-SILVA, R.; COULTHARD, M. Linguística Forense. In: DINIS OLIVEIRA, R. J.; MAGALHÃES, T. (org.). *O que são as Ciências Forenses? Conceitos, abrangência e perspectivas futuras*. 1. ed. Lisboa: Pactor, 2016. p. 137-144.

VAN HEE, C.; LEFEVER, E.; HOSTE, V. Exploring the Fine-Grained Analysis and Automatic Detection of Irony on Twitter. *Lang Resources & Evaluation*, Ghent, v. 52, p. 707-731, 2018. DOI: <https://doi.org/10.1007%2Fs10579-018-9414-2>